# Refining the Model of Plant Disease Resistance Protein Activation

by

Emma de Courcy-Ireland BBiotech(Hons)

A thesis submitted for the

Degree of Doctor of Philosophy

at

The School of Biological Sciences,

Flinders University

August 2014

# Table of Contents

# Abstract

Plant disease resistance (R) proteins act as specific pathogenic sensors within each and every plant cell. Upon recognition of their cognate pathogenic protein (effector or avirulence protein), R proteins initiate the pathogen-specific branch of plant immunity called effector triggered immunity, or ETI. Activation of ETI most often culminates in a form of cell death known as the hypersensitive response. Sacrifice of infected plant cells can save a plant from pathogenic colonisation.

R proteins have a tripartite architecture that consists of an N-terminal effector domain, a C-terminal sensor domain and a central activation domain and can be separated into two main classes on the basis of the form of their effector domain. These domains are either coiled-coil (CC) structures, or domains that share considerable homology with *Drosophila* Toll and mammalian interleukin-1 receptors and are thus called TIR domains. The central activation domain is a nucleotide binding site (NBS) domain that has been proven, in some R proteins, to bind adenosine nucleotides and catalyse an ATP hydrolysis reaction.

The molecular mechanisms that drive R protein activation remain poorly understood. The classical model of activation for these proteins describes a molecular switch that is turned on through exchange of an ADP molecule for ATP and switched off by hydrolysis of that ATP molecule. More recent evidence from the flax M TIR disease resistance protein suggests that, at least for this protein, activation of the HR maybe dependent on the hydrolysis of ATP. The aim of this study was to refine, redefine and/or expand the current model of R protein activation, with a focus on the M protein. Expression of R proteins other than M was also trailled so that their activation mechanisms could be studied too.

Expression of the *Arabidopsis thaliana* R proteins RPS2 and RPP1B was trialled with the *Pichia pastoris*-based system developed for the M protein. RPS2 expression was also tested in a baculovirus/insect cell system. Despite testing a number of truncated expression constructs, no RPS2 or RPP1B expression was detected in *P. pastoris*; however, insoluble full-length RPS2 protein was detected in the baculovirus/insect cell system. Additional work is warranted and may result in the production of soluble RPS2 protein.

In order to further investigate the mechanism of M protein activation, mutations were made within the highly conserved acidic residues of the Walker B motif within the NBS domain. These acidic residues are believed to play roles in coordinating and catalysing the ATP hydrolysis reaction. The functional consequence of each mutation was established with an *in planta Agrobacterium tumefaciens*-mediated transient

expression assay. Mutants that had a function dissimilar to that of the M protein were subsequently recombinantly expressed in *P. pastoris* and purified using affinity chromatography. The identity and quantity of any nucleotide bound within these mutant proteins was determined with a bioluminescent ATP quantification assay. ATP hydrolysis assays were also conducted with a subset of mutants. Results from this study corroborate the proposition that ATP hydrolysis is required for M to activate the HR.

Is a simple ATP hydrolysis reaction performed by the M protein, or does it have other biochemical functions? To investigate this possibility the M protein was tested for its potential to act as a nucleotide phosphatase and for its ability to perform protein phosphorylation. A functionally inactive mutant was also subjected to the same assays. A potential phosphorylation signal that was both time and M protein dependent was observed in SDS-PAGE after reaction of the M protein with [γ-$^{32}$P]ATP. The intensity of this signal was reduced in reactions containing the inactive mutant that has an impaired nucleotide binding capacity. A similar signal was not observed in reactions of M proteins with [α-$^{32}$P]ATP, suggesting that M is not a nucleotide phosphatase. Further study is required for the exact identity of this signal to be determined, but its presence does hint that the mechanism of M protein activation is more complex that first hypothesised and may involve protein phosphorylation.

It is evident from the results presented in this study that the mechanisms underlying R protein activation are likely to be more complex than first proposed. It is possible that some R proteins can perform biochemical reactions other than simple β-γ ATP hydrolysis, and that such activity may be necessary for the initiation of an effective immune response.

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

Emma deCourcy-Ireland

# Acknowledgements

To my supervisor Dr Peter Anderson, thank you for your guidance, patience and eternal optimism and for reminding me that even though R protein biochemistry is hard at times, the rewards of the research are worthwhile.

To my colleagues and friends Simon Williams, Nick Warnock, Pradeep Sornaraj, Hayden Burdett and Adam Bentham thank you for sharing your ideas and advice and for your seemingly endless support and encouragement. Doing science with you has been an absolute pleasure. Simon and Nick, I thank you especially for the frequent discussions that have helped me to plan my experiments and come to terms with science and PhDs.

Chevaun, Bianca, Lexi, Mel, Bart, Patrick and Drew thank you for your support and for all of the science talks, and most importantly laughs that we have had over the years together. Sharing the PhD experience with all of you made it a more enjoyable and bearable experience.

To the past and present students and staff of the Anderson, Soole and Stangoulis labs thank you for your help and for all of the laughs.

To all of my friends outside of science, especially my netball team, tennis and book clubs, thank you for putting up with all of the science talk, even though you probably could not have cared less! Letting me get it all out and helping me to live a life outside of the lab has been therapeutic and invaluable.

I also need to thank the Grains Research and Development Corporation of Australia for their generous funding of this research via their Grains Industry Research Scholarship.

To my family, Mum, Dad, Nicole, Brock, Rebecca, Jamie, Gran and Grandma, I owe you all so much for your emotional (and financial) support and for putting up with me not wanting to talk about the ‘P’ and ‘T’ words for the last few years. Even though you may not have understood all of the science, you supported me every day and I really appreciate it.

I really, truly could not have finished this PhD without every single one of you and I cannot thank you enough.

# 1 Introduction

## 1.1 Mechanisms of Plant Disease Resistance

Plants must continually battle to survive the persistent assault of their many bacterial, fungal, oomycete, viral, nematode and insect pathogens. To combat this pathogenic attack plants employ many different mechanisms of defence. The most simplistic of these mechanisms are passive, preformed resistance barriers like waxy cuticular surfaces, antimicrobial compounds and rigid cells walls (Jones & Takemoto, 2004). Pathogens that manage to penetrate these passive barriers of resistance and enter a plant's intracellular space are then faced with more specialised and sophisticated defensive manoeuvres.

### 1.1.1 MAMP- or PAMP-triggered immunity

At the cellular level plants possess a primary or basal immune system that has some similarities to the mammalian innate immune system. Called microbe or pathogen-associated molecular pattern-triggered immunity, MAMP-TI or PAMP-TI, this system serves as a non-specific method of pathogen perception. Once activated this system acts to inhibit the spread of a pathogen after it has successfully infected and manifested disease symptoms in the host plant (Chisholm *et al.*, 2006; Jones & Dangl, 2006).

PAMPs or MAMPs are conserved microbial features (i.e. lipopolysaccharide, flagellin and chitin) that are unique to microbes and essential for their function (Jones & Takemoto, 2004). They are first perceived by the plant at the external face of a cell via transmembrane pattern recognition receptors (PRRs). This recognition event or stimulation induces many resistance responses in the infected cell. These include the induction of protein kinase signalling cascades, transcriptional activation of pathogen-responsive genes, production of reactive oxygen species and the deposition of callose at infection sites. Together these responses act to limit the spread of an invading pathogen before it can successfully take hold of the entire plant (Belkhadir *et al.*, 2004; Chisholm *et al.*, 2006; Dodds & Rathjen, 2010).

While this primary line of defence can provide a plant with effective resistance, pathogens have continued to evolve to subvert, and even bypass, this first tier of defence. Vast repertoires of effectors are released into a host by a pathogen in an

attempt to undermine basal plant defences, re-direct the nutrient supply and ultimately enhance their virulence.

### 1.1.2 Effector-triggered immunity (ETI)

Plants have responded to this evolutionary advancement of pathogens with a more specific means of pathogen perception. This second tier of resistance centres on the detection of the secreted race-specific effectors and the translation of this perception into an appropriate resistance response (Chisholm *et al.*, 2006; Dangl & Jones, 2001; Takken *et al.*, 2006).

This gene-for-gene theory of resistance was first proposed by Flor during work on the flax/flax rust (*Linum usitatissimum*/*Melampsora lini*) system. A specific interaction was genetically characterised to occur between a plant carrying a specific *R* gene and a pathogen carrying a corresponding race-specific Avirulence (*Avr*) gene (Flor, 1956; Flor, 1971; Lawrence *et al.*, 2007). The protein products of these *Avr* genes are known as Avr proteins or effectors. This *R* gene-mediated resistance, also known as effector-triggered immunity (ETI), is a plant's secondary line of innate immunity.

Extensive research has identified that signalling pathways employed by the PTI and ETI systems are very similar, and appear to converge at certain points, such as at the mitogen-activated protein kinase (MAPK) cascade (Jones & Takemoto, 2004). While the pathways that lead to the formation of a resistance response may be similar for both systems, *R* gene-mediated resistance is more vigorous and guarantees a halt to infection (Dodds & Rathjen, 2010). This form of resistance can often be characterised by areas of localised cell death at the sites of infection. This response is called the hypersensitive response (HR) and is a type of programmed cell death that restricts the growth of biotrophic pathogens in infected cells. However, a HR is not always required, or observed during ETI (Goodman & Novacky, 1994; Heath, 2000).

Manifestation of ETI and the HR includes, but is not limited to, the following events: rapid calcium and ion fluxes; the production of reactive oxygen species (ROS) and nitric oxide; and protein kinase activation (including the MAPKs) (Belkhadir *et al.*, 2004; Chisholm *et al.*, 2006; Jones & Dangl, 2006). Transcriptional reprogramming within and around infection sites is also a major part of the formation of a resistance

response. WRKY and TGA families of transcription factors are largely responsible for these changes in gene expression (Eulgem, 2005).

Other genes that are upregulated during the defence response include those that encode or synthesise molecules crucial for the establishment of systemic acquired resistance (SAR) and induced systemic resistance (ISR). Those molecules include salicylic acid (SA), and ethylene and jasmonic acid; the key molecules in the establishment of SAR and ISR respectively (Hammond-Kosack & Parker, 2003). SAR and ISR work in combination with similarly induced pathogenesis-related (PR) and PR-like genes to enhance the resistance of the rest of the plant. Together they prevent the possibility of future pathogenic attack, promote cell wall strengthening and lignification and ultimately act to inhibit pathogenic growth (Chisholm *et al.*, 2006; Ferreira *et al.*, 2007).

## 1.2 Avirulence Proteins (Effectors)

### 1.2.1 Bacterial effectors

Phytopathogenic bacteria, such as *Pseudomonas syringae* (Dangl & Jones, 2001), typically employ a type-III secretion system (TTSS) to deliver their effectors directly into the host cell cytosol. They do this for two main reasons; to evade detection at the host cell plasma membrane and to enhance their pathogenicity (Chisholm *et al.*, 2006; Staskawicz *et al.*, 2001). Pathogenicity is generally assisted by the delivery of virulent effector proteins whose targets include host proteins that are involved in physiological pathways (Dangl & Jones, 2001). Through these targets, pathogens possess the ability to redirect their host's energy supplies and use them to sustain their own growth during colonisation. More often than not, detection of type-III effectors by R proteins occurs in resistant plants before a pathogen becomes established. Subsequent R protein activation results in the restriction of pathogenic colonisation.

### 1.2.2 Fungal and oomycete effectors

To date, no TTSS or similar secretory system has been identified in biotrophic fungal phytopathogenic organisms. Instead, these pathogens, which are typified by rust

fungi, form specialised feeding structures called haustoria from haustorial mother cells (HMC). These haustoria invaginate a mesophyll cell between its cell wall and plasma membrane (Figure 1.1). It is postulated that avirulence proteins are then secreted across the haustorial membrane by the endomembrane secretory pathway into the extrahaustorial matrix (EHM), the space between the haustorial cell wall and the host plasma membrane (Dodds *et al*., 2004, Ellis *et al.*, 2006). In the case of the flax rust *M. lini*, many features of its encoded Avr proteins support this model. Firstly, the presence of signal peptides at the N termini of these Avr proteins, which are expected to direct the trafficking of the proteins through the rust's secretory pathway (Ellis *et al.*, 2007). Secondly, that *Avr* gene expression is abundant in haustoria relative to ungerminated spores, and thirdly, recognition of the Avr proteins by flax R proteins occurs intracellularly (Catanzariti *et al.*, 2006; Dodds *et al.*, 2004; Takemoto *et al.*, 2012).

Supporting evidence for this model also comes from the first fungal avirluence protein shown to be translocated into the host cell; rust transferred protein 1 (RTP1) from *Uromyces fabae*. Although the mechanism used by this protein to cross the EHM and accumulate in the host cell is not known, RTP1 is found within the host cell's cytoplasm, and even within its nucleus (Kemen *et al.*, 2005).

Biotrophic fungal effector proteins that enter the host cell have also been demonstrated to subvert the metabolism of the living plant cell to favour their own growth and reproduction (Chisholm *et al.*, 2006; Hammond-Kosack & Jones, 1997). How the Avr effectors enter the host cell from the EHM remains a mystery, but it is postulated that rust-encoded Avr proteins may enter the host cell via a host-encoded system (Catanzariti *et al.*, 2006). Indeed, the uptake of the *M. lini* effectors AvrL567 and AvrM can occur independently of the rust fungus itself, but it is dependent on unknown signals in their N-terminal regions (Rafiqi *et al.*, 2010).

How oomycete effectors achieve translocation into the host cell is a little better understood. Two motifs, the RxLR and dEER, have been identified in the N-terminal region of many oomycete Avr proteins, including many effectors from *Hyaloperonospora arabidopsidis*. The RxLR motif is similar to the host-cell targeting motif of the malarial (*Plasmodium* species) parasite, RXLX(E/Q) (Birch *et al.*, 2006).

These motifs are located after the haustorial secretion motif and mutation within this region impairs the host-cell translocation of many oomycete effectors (Birch *et al.*, 2006; Chou *et al.*, 2011; Whisson *et al.*, 2007).

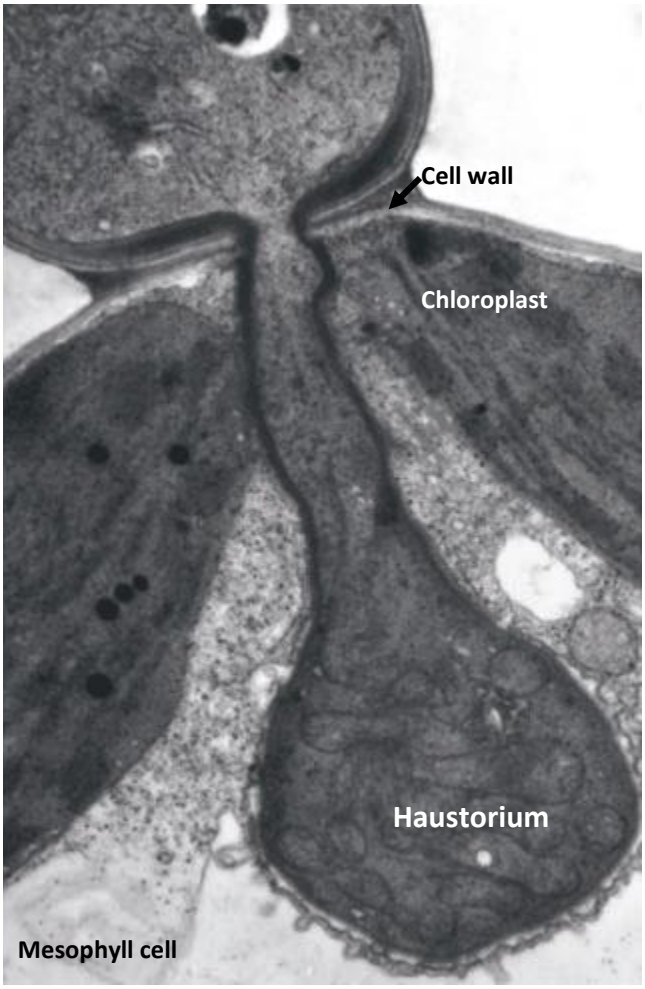



**Figure 1-1: A *M. lini* (flax rust) haustorium that has formed in the mesophyll cell of an infected flax plant (Ellis *et al.*, 2007)**

This electron micrograph shows a flax rust haustorium that has formed in a mesophyll cell of an infected flax leaf. The fungus entered the plant through the stomata and once in the sub-stomatal space infection hyphae spread and form haustorial mother cells (HMC). These HMC then invaginate haustorial feeding structures between the cell wall and plasma membrane of a mesophyll cell (Lawrence *et al.*, 2007). The haustoria will give the invading rust pathogen access to the host plant's energy supplies and it is believed that Avr proteins are secreted across the haustorium/cell interface into the host cell cytoplasm. Once inside a plant cell disease resistance or susceptibility is determined by the presence or absence of the corresponding R protein (Ellis *et al.*, 2007).

## 1.3 Plant Resistance Proteins

*Resistance* (*R*) genes are classified according to the nature of the domains that their protein products are predicted to contain. The majority of cloned *R* genes belong to three major classes; the serine/threonine protein kinases, the extracellular LRR (eLRR) class and the nucleotide binding site-leucine rich repeat classs (NB-LRRs). The first resistance gene cloned, *HM1* from Maize, is a member of the eLRR class of *R* genes (Johal & Briggs, 1992).

Genes belonging to the NB-LRR class are the most prevalent (Chisholm *et al.*, 2006; Meyers *et al.*, 1999). Their protein products typically have tripartite domain architecture and are expressed intracellularly. Many have been shown to have peripheral membrane associations with the plasma membrane, endomembrane system or other organelles. For example, the flax NB-LRRs M and L6 are located at the tonoplast and Golgi body, respectively (Ellis *et al.*, 2007; Takemoto *et al.*, 2012). Others, like Rx and Mla10, from *Solanum tuberosum* (potato) and *Hordeum vulgare* (barley) respectively, shuttle between the cytosol and the nucleus (Bonardi *et al.*, 2012).

### 1.3.1 STAND classification

Plant NB-LRR proteins share significant architectural homology with nucleotide oligomerisation domain-leucine-rich repeat proteins (NOD-LRRs) or NOD-like receptors (NLRs), which are integral members of the mammalian innate (& adaptive) immune system (Maekawa *et al.*, 2011b). The NLR class includes the mammalian and *Caenorhabditis elegans* cell death pathway regulators, Apaf-1 (apoptotic protease-activating factor-1) and CED-4 (*C. elegans* death-4) (Inohara & Nunez, 2003).

R proteins and NLR proteins are members of the STAND (signal transduction ATPase with numerous domains) sub-class of the AAA-ATPase superfamily classification. The STAND class brings together proteins with varying types of N-terminal signalling domains (known as effector domains) that are involved in downstream signalling. They also contain a central regulatory switch region that is usually tethered to a C-terminal sensor domain, which is involved in protein-protein interactions and binds the effector/inducers (Danot *et al.*, 2009). It is this multi-domain structure that makes

STAND proteins a hub for perception, signalling and regulation. NB-LRR proteins, as well as Apaf-1 and CED-4, belong to the AP-ATPase clade. NLRs such as NOD1 and NOD2, are classified in the NACHT (after NAIP, CIIA, HET-E and TP1) clade (Zurek *et al.*, 2011b; Leipe *et al.*, 2004).

The STAND protein family is characterised by enzymes that most commonly hydrolyse the β-γ phosphate bond of a bound nucleoside triphosphate (NTP) (most often adenosine triphosphate (ATP)). STAND are molecular switches that are regulated by nucleotide binding (Leipe *et al.*, 2004; Marquenet & Richet, 2007). The ADP-bound form is believed to be the resting or 'off' state of the proteins, whilst the active state is induced by pathogen perception and is ATP-bound. ATP hydrolysis is the mechanism that re-sets them back to their dormant state and thus controls the molecular switch (Danot *et al.*, 2009; Marquenet & Richet, 2007; Takken & Tameling, 2009; Williams *et al.*, 2011).

The conserved, multi-domain regulatory switch domain carries the ATPase activity and is referred to as the STAND module (Leipe *et al.*, 2004). This region is defined by a number of sequence characteristics: a conserved proton-abstracting acidic residue in the Walker B motif; a hhGRExE motif ahead of the P-loop motif; a GxP or GxxP motif near the C-terminus of the NTPase domain and a less conserved sequence, including an arginine residue ahead of the Gx(x)P motif (Leipe *et al.*, 2004).

Most STAND members also possess a region of ~200 amino acids between the Gx(x)P and the C-terminus, or the start of another domain, which has been termed the HETHS domain (helical third domain of STAND proteins). This helical bundle region contains a conserved serine and a HD motif and is predicted to act like a lever to transmit ATP hydrolysis-generated conformational changes to the signalling domain(s) (Leipe *et al.*, 2004). In NLRs this region contains the sensor 2 motif, which may be analogous to the MHD motif of plant NB-LRRs.

While there are many homologies between R proteins and other NB-LRR proteins from other STAND classes, there are also differences, and mechanistic modes of action may not be conserved across the classes. Therefore, while biochemical investigations of these proteins help to guide R protein experiments, definitive

conclusions about molecular mechanisms of activation should not be made without first testing them in each and every plant NB-LRR protein.

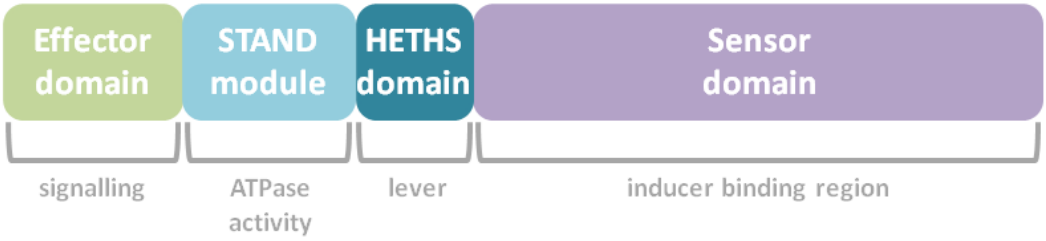



**Figure 1-2: Schematic representation of the domain architecture of a STAND protein**

The STAND classification brings together proteins with a wide variety of N-terminal domains that are all involved in signalling. The C-terminal domain plays the role of sensing ligands or inducers and also varies between different STAND proteins. The central regulatory switch STAND module and the HETHS domains are predicted to play crucial roles in protein activation, regulation and the communication of intramolecular conformational changes. These roles are carried out via the mechanisms of a nucleotide switch and NTP hydrolysis (Danot *et al.*, 2009; Leipe *et al.*, 2004; Marquenet & Richet, 2007).

### 1.3.2 The N-terminal domain

There are two main types of N-terminal domains in R proteins; Toll Interleukin-like Receptor (TIR) domains and coiled coil (CC) domains. The TIR domain possesses considerable homology with the cytoplasmic domains of the *Drosophila* Toll and mammalian interleukin-1 receptors. Whereas CC domains are mostly $\alpha$-helical rich and form coiled-coil structures (Maekawa *et al.*, 2011a; Meyers *et al.*, 1999; Rairdan & Moffett, 2007).

Some N-termini have been demonstrated to be crucial for the activation of the downstream signalling pathways. The TIR domain from the flax NB-LRR protein L6 has been demonstrated to be both necessary and sufficient for activation of the immune response (Bernoux *et al.* 2011b). This ability to activate effector-independent HR has also been reported for an extended CC region of Mla10, the TIR domain of RPP1-WsB, (Recognition of *Peronospora parasitica*1-WsB) fused to a wild-type green fluorescence protein (Krasileva *et al.*, 2010), also RPP1-WsA and RPS4 ((Maekawa *et al.*, 2011a; Krasileva *et al.*, 2010; Mestre & Baulcombe, 2006; Swiderski *et al.*, 2009; Weaver *et al.*, 2006). Additionally, this action has been demonstrated by the requirement of the TIR-NB-LRRs for a particular set of downstream genes; *enhanced disease susceptibility 1* (*EDS1*) and the CC-NB-LRR subclass for another; *non-race-specific disease resistance 1* (*NDR1*) (Aarts *et al.*, 1998; Swiderski *et al.*, 2009).

The recently solved crystal structures of the TIR domain from L6 and the CC domain from Barley's Mla10 have revealed that these N-terminal domains can homodimerise (Bernoux *et al.* 2011a; Maekawa *et al.*, 2011a; Ve *et al.*, 2011). Mla10 exists as a homodimer in the absence of its effector AvrA10 (Maekawa *et al.*, 2011a), whilst the TIR domains of two L6 proteins homodimerise in response to the detection of their cognate effector AvrL567. This evidence *of in vitro* dimerisation has been supported by *in planta* and yeast-two hybrid experiments for L6 and *in planta* data for Mla10. In addition, mutations at the dimer interfaces of both proteins disrupt homodimerisation and signalling. The self-association of N-terminal domains has also been reported for the TIR-NB-LRR *Nicotiana tabacum* (tobacco) N and the CC-NB-LRRs RPS5 and Prf (Ade *et al.*, 2007; Gutierrez *et al.*, 2010; Mestre and Baulcombe, 2006).

It is generally considered that the LRR domain of an R protein is responsible for determining specificity towards the pathogen effector protein. However, research has shown that amino acids in regions of the TIR domain predicted to interact with the LRR are also under diversifying selection to evolve and may be co-evolving with the LRR to provide specificity (Luck *et al.*, 2000). Also, by physically interacting with the host protein target of an Avr effector, it has been demonstrated that the TIR domains of some R proteins are involved in mediating the interaction between an R protein and its Avr effector. Using the yeast two-hybrid system such an interaction has been demonstrated between the R protein RPM1 and the host target of its cognate *P. syringae* Avr effector, the RPM1 interacting protein 4 (RIN4) (Ade *et al.*, 2007; Mackey *et al.*, 2002; Mackey *et al.*, 2003).

### 1.3.3 The nucleotide-binding site or NB-ARC domain

The nucleotide binding site (NBS) domain is a common functional element of many enzymatic proteins in plants and animals. Extensive research has determined that this region is required for NTP-binding in many of these proteins (Marquenet & Richet, 2007; Traut, 1994; Saraste *et al.*, 1990; Walker *et al.*, 1982). ATP and adenosine diphosphate (ADP) binding are linked with the proposed method of activation for R proteins and for this reason the NBS is also often called the activation domain. A functional nucleotide binding pocket has been found within six R proteins to date; I-2 & Mi-1 from tomato (Tameling *et al.*, 2002), N from tobacco (Ueda *et al.*, 2004), Mla27 from barley (Maekawa *et al.*, 2011a) and M and L6 from flax (Williams *et al.*, 2011). It is worth noting that the biochemical analysis of I-2, Mi-1 and N was all performed on truncated versions of the proteins.

Due to their STAND classification, the NBS domain of R proteins can be extended to include a carboxy-terminal extension called the ARC domain (due to its conservation in Apaf-1, R proteins and CED-4) and is therefore also often referred to as the NB-ARC domain (van der Biezen & Jones, 1998a; Leipe *et al.*, 2004; Takken *et al.*, 2006). The crystallisation of Apaf-1 has given insights into the possible structure of the R protein NB-ARC domain (Riedl *et al.*, 2005). Structural models based on this crystal structure propose that the R protein NB-ARC domain forms three structural sub-domains. These have been designated the nucleotide-binding (NB), ARC1 and ARC2 sub-

domains (Albrecht & Takken, 2006; van Ooijen *et al.*, 2008). The NB is defined by a fold that makes a nucleotide binding and hydrolysing pocket. The ARC1 sub-domain is a helical domain (HD) and has been predicted to be involved in mediating an intra-molecular interaction between the LRR domain and the N-terminal domain (Rairdan & Moffett, 2006, Slootweg *et al.*, 2013). The most C-terminal sub-domain, the ARC2, is predicted to form an extended winged helix domain (WHD). This region is thought to help maintain R proteins in an autoinhibited state. It may consequently also be involved in their activation when it communicates with the rest of the protein the predicted conformational changes that occur as a consequence of effector recognition (Rairdan & Moffett, 2006; Riedl *et al.*, 2005; Slootweg *et al.*, 2013).

#### 1.3.3.1 Conserved motifs within the NBS

Eight highly conserved motifs have been identified within the NB-ARC domain of plant NB-LRR proteins and Table 1-1 illustrates their consensus sequences. These motifs are the phosphate binding loop (P-loop; Walker A or Kinase 1a), the RNBS-A (resistance nucleotide binding site), Walker B (or Kinase 2a), RNBS-B (some homology to kinase-3 motif), RNBS-C, GLPL, RNBS-D motifs and the MHD (Met-His-Asp) (Leipe *et al.*, 2004; Meyers *et al.*, 1999; Meyers *et al.*, 2003; Saraste *et al.*, 1990; Traut, 1994; van der Biezen & Jones, 1998a). Some of the above conserved motifs are analogous to those found in related ATP/GTP-binding proteins and also STAND proteins with ATPase activity. They are thus predicted to have defined roles in nucleotide binding and/or ATP hydrolysis and therefore in R protein function (Traut, 1994; Saraste *et al.*, 1990; Walker *et al.*, 1982).

The first of these motifs with a defined role in nucleotide binding is the P-loop. Located within the NB sub-domain of the NB-ARC, this motif is the most commonly conserved within the NBS of all ATP/GTP-binding proteins (Saraste *et al.*, 1990; Traut, 1994). An intact P-loop is required for R protein function as has been shown in a number of R proteins, including L6, M, N (tobacco), RPS2 and RPM1 (Dinesh-Kumar *et al.*, 2000; Howles *et al.*, 2005; Tao *et al.*, 2000; Tornero *et al.*, 2002; Williams *et al.* 2011). The invariant lysine residue in this motif (GVGKTT) is predicted to coordinate the positioning of the $\beta$- and $\gamma$-phosphates of the bound nucleotide (Saraste *et al.*, 1990) and R proteins containing lysine mutations have been shown to be inactive and

purified with significantly lower levels of bound nucleotide (Tameling *et al.*, 2002; Williams *et al.*, 2011).

The Walker B motif of the NB sub-domain also contains residues that play important roles in R protein function. The first invariant aspartate of this motif is believed to be responsible for coordinating the metal ion co-factor (often $Mg^{2+}$) that is required for phospho-transfer reactions (Traut, 1994). The Walker B also contains a second highly conserved acidic residue that is thought to act as the catalytic base for ATP hydrolysis, due to its role in coordinating the hydrolytic water molecule that is required for the ATP hydrolysis event (Leipe *et al.*, 2004; Muneyuki *et al.*, 2000). In NB-LRR R proteins there is strong conservation of an aspartate residue at this position (Tameling *et al.*, 2006). Mutations of either residue in other STANDS have resulted in a wide variety of functional consequences, including loss-of-function, pathogen-independent activation and no change (Marquenet & Richet, 2007; Tameling *et al.*, 2006). Some TIR-NB-LRR proteins have additional acidic residues in this region, but their roles in R protein function are as yet unknown (for more see chapter 4).

The integrity of the histidine and aspartate residues of the ARC2 sub-domain's MHD motif is also important for R protein function. Mutation of either residue can cause autoactivity, which is defined as pathogen-independent R protein activation (de la Fuente van Bentem *et al.*, 2005; Gabriëls *et al.*, 2007; Howles *et al.*, 2005; Williams *et al.*, 2011). The imidazole ring of the histidine of the analogous LHD motif in Apaf-1 forms a hydrogen bond with the β-phosphate of the bound ADP molecule (Riedl *et al.*, 2005). It is postulated that the MHD motif is analogous to the sensor II motif in the AAA+ ATPase superfamily (ATPases associated with diverse cellular activities; Proell *et al.*, 2008; Takken *et al.*, 2006). Due to these factors, the MHD motif is thought to mediate nucleotide-dependent conformational changes (Riedl *et al*., 2005).

Although highly speculative, it is thought that mutational destabilisation of the histidine-ADP interaction causes autoactivity because it allows the protein to mimic the ATP-bound active state that is achieved after the release of ARC2-mediated autoinhibition; as would normally be triggered by effector-induced activation. Indeed, an autoactive mutant of the M flax rust resistance protein has been found to

have a preference for binding ATP over ADP, and thus more often than not, it may mimic M's active state (Williams *et al.*, 2011).

Based on functional and biochemical information gained from mutations in regions described above, and on structural information from related STAND proteins, it is hypothesised the NB-ARC domain of an R protein has a switch-like function. An R protein is switched on or off by conformational changes that occur when nucleotides bind and dissociate (or are hydrolysed) from this region (Leipe *et al.*, 2004; Marquenet & Richet, 2007; Takken *et al.*, 2006). According to this hypothesis an R protein's off or resting state is ADP-bound, while they are active when bound to ATP.

**Table 1-1: The amino acid consensus sequences of the eight highly conserved motifs within the NBS of TIR- and CC-NB-LRR proteins**

The P-loop, RNBS-B, RNBS-C and GLPL are commonly conserved between nearly all the NBS sequences, whilst the Walker B, RNBS-A, RNBS-D and MHD motifs have varying consensus sequences between the TIR- and CC-NB-LRR sub-classes (adapted from Meyers *et al.*, 1999 & 2003).

| Motif | CC-NB-LRR | TIR-NB-LRR |
|---|---|---|
| P-loop | GVGKTT | |
| RNBS-A | FDLxAWVCVSQxF | FLENIRExSKKHGLEHLQK KLLSKLL |
| Walker B | LLVLDDVW | LLVLDDVD |
| RNBS-B | GSRIIITTRD | |
| RNBS-C | YEVxxLSEDEAWELFCKxAF | |
| GLPL | CGGLPLA | |
| RNBS-D | CFLYCALFPED | FLDIACFF |
| MHD | VKMHDVVREMALWIA | MHNLLQQLGREIV |
| x indicates any amino acid residue | | |

#### 1.3.3.2 Nucleotide binding and ATPase activity in the NB-ARC

Other than the M protein (see section 1.5), there is only one other report specifically detailing the nucleotide occupancy of an NB-LRR protein and this comes from the barley Mla27 CC-NB-LRR protein. This full-length protein was expressed in an insect cell expression system and during its purification by gel filtration it was found that the protein was directly associated with an amount of ADP that equated to an occupancy of ~43%. Free ADP found in later gel filtration fractions was thought to have dissociated from Mla27 during purification and thus the overall ADP occupancy of this protein is reported as ~61% (Maekawa *et al.*, 2011a). Similar dissociation of ADP during purification has been observed in studies on the M protein (Williams, *et al.*, 2009). The ADP occupancy of Mla27 is higher than that obtained for the near full-length M protein, but ATPase activity of Mla27 has not yet been reported.

The intrinsic ATPase activity or turnover rate of full-length STAND proteins is often very low and thus difficult to measure (Danot *et al.*, 2009). Additionally, some STANDs, including the *Drosophila* ARK apoptotic factor, are not predicted to possess the ability to hydrolyse ATP because they lack critical functional residues (Riedl *et al.*, 2005). Others, like CED-4, despite containing such residues, do not have any detectable ATPase activity (Yan *et al.*, 2005). The ability to hydrolyse ATP to ADP and inorganic phosphate has only been demonstrated by the I-2, Mi-1 (Tameling *et al.*, 2002), tobacco N (Ueda *et al.*, 2006) and M plant NB-LRR proteins (Sornaraj, 2013). It is important to note that the ATPase activities of I-2, Mi-1 and N were measured with truncated versions of each protein (CC-NBS only for I-2 and Mi-1, NB-LRR for N) and therefore may not represent the true ATPase activity of a fully regulated NB-LRR. The M protein ATPase assays used a near-full length version of the protein that only had the first 21 amino acids removed (see 1.5 for more; Sornaraj, 2013).

In I-2 and M, mutation of the invariant P-loop lysine results in a significant decrease in the rate of ATP hydrolysis. This observation is attributed to the inability of these mutants to bind a nucleotide (Sornaraj, 2013; Tameling *et al.*, 2002; Williams *et al.*, 2011). Biochemical analysis of autoactive I-2 mutants suggests that ATP hydrolysis is not required for the initiation of downstream signalling, as these mutants display a decreased rate of ATP hydrolysis compared to the un-mutated I-2. Instead, this work

suggests that the ATPase activity allows the protein to cycle back to its inactive state and therefore acts as a 'reset' switch (Tameling *et al.*, 2006). For the tobacco N protein, which is ATP-bound, interaction with its tobacco mosaic virus (TMV) effector p50 increased its intrinsic ATPase activity (Ueda *et al.*, 2006). This account with tobacco N is the only report of an R protein bound to ATP in its resting state and ATP hydrolysis being triggered by effector interaction.

ATPase activity implies that the energy released during ATP hydrolysis is used to perform a mechanochemical function. This often requires a conformational change and in R proteins this is postulated to involve the ARC2 (or HETHS) domain (Takken *et al.*, 2006). Interestingly, an autoactive ARC2 mutant of M, $M^{D555V}$, has recently been shown to have a higher maximal rate of ATP hydrolysis than wild-type, un-activated, M (Sornaraj, 2013) and, as mentioned previously, it is co-purified with both bound ADP and ATP, but displays a preference for ATP. It is proposed that this mutation allows the protein to escape ARC2 mediated-autoregulation and adopt an open or loose conformation that can easily assume the ATP-bound form, thereby presenting it in an activated conformation more often than not. An increased rate of ATP hydrolysis may mean that the protein can cycle back to the pre-activated form more quickly than wild-type M, which means that it is free to bind ATP and be activated again more quickly.

#### 1.3.3.3 Nucleotide phosphatase activity in the NB sub-domain

Fenyk *et al.* (2012) recently presented biochemical evidence that the NB-ARC domain of the *A. thaliana* CC-NB-LRR protein RPM1 possesses nucleotide phosphatase activity. This biochemical activity was also shown by the NB sub-domain of a CC-NB-ARC containing *Oryza sativa* (rice) protein Os02g_25900 and the NB-ARC domain of the *Zea mays* (maize) Pollen Signalling Protein, an orphan R protein that also contains a CC domain. The NB sub-domain of the rice protein was found to sequentially cleave terminal phosphates from adenine and guanidine nucleotide substrates to form a nucleoside product, but its preferred substrate was adenosine monophosphate (AMP). The mechanism behind this ability to cleave phosphodiester bonds is proposed to be nucleophilic attack by a divalent metal activated water molecule, an activity that is supported by the residues within the Walker B motif of the NB sub-

domain. This rice NB-ARC domain may act as an oligomer and mutations of the catalytically important P-loop lysine and Walker B aspartate residues significantly reduce its phosphatase activity (Fenyk *et al.*, 2012).

The impact of this research on other R protein studies that have previously shown ATPase activity is yet to be explored. Although, it is interesting to note that Tameling *et al.* (2002) performed competition assays with I-2 and showed it also has affinity for ADP, and also deoxy versions of ATP and ADP (dATP and dADP). Perhaps the suggestion that these R proteins are simple ATPases is somewhat unsophisticated and additional functions, such as nucleotide phosphatase activity, may also be performed by R proteins.

Alternatively, despite sequence similarities that suggest otherwise, it may possible that different R proteins have completely different modes of biochemical activation. Further investigations into the activation mechanisms of more R proteins are required before this will be known. This requirement has prompted the study presented here.

#### 1.3.3.4 Alternative activities/functions of the NB-ARC domain

Before nucleotide binding had been demonstrated by the NBS of any R proteins, it was postulated from a bioinformatics study that some R proteins possess a region within their NBS domain that is structurally homologous to the CheY family of response regulator receiver domains of His-Asp phosphotransfer signalling pathways (Rigden *et al*., 2000).

These pathways, also known as two-component signalling systems, involve a sensor histidine kinase and a response regulator that contains the receiver domain. In response to a signal, the sensor histidine kinase catalyses ATP-dependent autophosphorylation of a conserved histidine residue in the transmitter domain. The response regulator then catalyses transfer of the phosphoryl group to a conserved aspartate in the receiver domain (West & Stock, 2001). The phosphoryl signal is then transferred via intramolecular or intermolecular interactions of the phosphorylated receiver, which is often carried out by separate effector domains. Intrinsic or extrinsic phosphatase activity causes dephosphorylation and deactivation of the receiver

domain, and hence acts as an off switch for signal transduction. Rigden *et al.* (2000) suggest that the phosphorylated aspartate residue in the CheY receiver domain is analogous to the first conserved aspartate of the Walker B motif of R proteins. Phosphatases and kinases are often associated with the His-Asp phosphotransferase for regulatory reasons (West & Stock, 2001).

Despite that bound nucleotide has now been found in some R proteins, could these findings suggest a more sophisticated mechanism of self-regulation, activation and/or signalling? Histidine kinase two-component signal transduction systems are predicted to play large roles in the signalling pathways of plants (Urao *et al.*, 2000). In eukaryotes, protein involved in two-component signalling systems are found at the beginning of signal transduction pathways and they can interact with conventional signalling systems, like MAPK pathways. In fact, a member of the GHKL (gyrase, Hsp90, histidine kinase, MutL) ATPase/kinases superfamily, CRT1 (compromise recognition of turnip crinkle virus), has been found associated with the R proteins SSI4, RPS2, and Rx (Kang *et al.*, 2008) and it is suggested that this protein may be important in mediating the signal transduction pathways of TIR- and CC-NB-LRRs. In addition, a phosphatase has also been found associated with the I-2 CC-NB-LRR protein from tomato (de la Fuente van Bentem *et al.*, 2005).

Interestingly, a large-scale phosoproteomics study carried out by Nakagami *et al.,* (2010) identified phosphorylated residues in functionally significant NB-LRR motifs. Perhaps most interesting in the context of this study was the discovery of two phosphorylated residues in the NB-ARC domain (Nakagami *et al*., 2010). The authors of this study suggest that NB-LRR proteins could be regulated by conserved phosphorylation-dependent systems. Further research is clearly warranted to explore some of these possibilities and chapter 5 of this study represents some hints that M protein phosphorylation maybe an outcome of ATP hydrolysis.

### 1.3.4 The C-terminal LRR domain

The LRR domain is a feature of many proteins, both plant and animal and has long been hypothesised to be a key player in mediating the protein-protein and protein-ligand interactions of those proteins (Kobe & Deisenhofer, 1994). It is clear that the

LRR domain plays a major role in determining the recognition specificity of R proteins; this holds true for those proteins of the eLRR class (Cf-9 and Cf-4, for example) and from the NB-LRR class. As the focus of this thesis is on the NB-LRR class I will only focus on LRRs from these proteins. Comparative sequence analyses between closely related NB-LRR proteins show that specific residues within the LRR domain are under diversifying selection to evolve (Meyers *et al.*, 1999; Hwang & Williamson, 2003).

In the flax disease resistance genes, including those at the *L* and *P* loci, Ellis *et al.* (1999) have shown that the different specificities encoded by many of these genes can be attributed to residues within the LRR, especially those in the carboxy-terminal LRR units. The putatively solvent-exposed residues xxLxLxx (x = any amino acid, L = leucine) confined to the β-strand/β-turn regions of individual LRR units, predicted to be at the ligand-binding interface, have been shown to be under particular diversifying selection (Dodds *et al.*, 2001; Dodds *et al.*, 2006; Ellis *et al.*, 1999, Ellis *et al.*, 2007). These same researchers have also demonstrated, with physical evidence, this mediation of R protein recognition specificity by the LRR domain. They generated chimeric proteins with non-native LRR domains and native TIR and NBS regions and showed that the proteins retained the Avr protein recognition specificity of the LRR domain's native protein (Ellis *et al.*, 2007).

LRR regions of some R proteins have also been demonstrated to bind directly with their corresponding Avr proteins, implicating this region in the provision of recognition specificity through direct ligand binding. As an example, Jia *et al.* (2000) showed, using the yeast two-hybrid system and an *in* vitro binding assay, that the LRR domain of the R protein Pi-ta could bind to Avr-Pi-ta, independently of the rest of the protein.

It also predicted that this domain has other roles, and is particularly important for the autoregulation of R protein activation. Evidence to support this function comes from many fronts, including the observation that removal of the LRR domain from some R proteins, like RPS2 and RPP1-WsA, causes the constitutive activation of defence responses (Tao *et al.*, 2000; Weaver *et al.*, 2006). However, in other R proteins, for example I-2 and Rx, it does not (Moffett *et al.*, 2002; Tameling *et al.*, 2006). Interestingly, in Rx, the LRR domain is required for the autoactivity of the MHD motif

mutant, $Rx^{D460V}$ (Moffett *et al*., 2002). Together these observations suggest that the LRR domain of R proteins can have both positive and negative regulatory roles during R protein activation.

Interestingly, some NB-LRR proteins possess C-terminal domains in addition to the typical LRR. For example, Resistance to *Ralstonia solanacearum* 1 (RRS1) from A*. thaliana* has a putative transcriptional domain after its LRR domain which contains a WRKY motif (Deslandes *et al.*, 2002). Such atypical NB-LRRs may be able to directly influence the gene transcription controlling an immune response.

## 1.4 Resistance Protein Activation

### 1.4.1 Effector recognition

A great deal of debate surrounds whether recognition of an Avr protein by its corresponding plant R protein occurs by means of direct or indirect protein interactions (DeYoung & Innes, 2006). As more and more R proteins have been studied, examples of both types of recognition have been revealed and it has become clear that there is not one common rule across similar R protein and effector classes for a specific mode of recognition.

#### 1.4.1.1 Indirect recognition

When genome searches first began to identify R proteins, in comparison with the number of related mammalian immune receptors, a relatively low number were identified (Rairdan & Moffett, 2006). How then was it possible that such a limited number of R proteins could detect the vast and diverse number of pathogenic effectors (DeYoung & Innes, 2006)? To explain this conundrum the Guard hypothesis was developed. This hypothesis states that R proteins indirectly detect the presence of effector (Avr) proteins by monitoring the integrity of key host cellular targets (virulence targets or mimics of one) (van der Biezen & Jones, 1998b), thus explaining how a limited number of R proteins can activate a resistance response against the vast array of potential pathogens.

With continuing research into pathogenic effectors is has become apparent that some effector proteins play a role in the virulence of the pathogen. This function

enhances pathogen fitness and the host target of the effector is consequently deemed indispensable for the virulence function of the effector. It seems logical that the plant would respond to this evolutionary pressure by evolving a target 'mimic' that would function only in effector perception, possibly in R protein function, but not in plant defence pathways. This theory has coined a new hypothesis for indirect detection that is the Decoy model (van der Hoorn & Kamoun, 2008). It is possible that effector target gene duplication, or independent evolution of a decoy, occurs for this purpose. The role of either is to trap the pathogen into a recognition event that will then signal to activate the defence response. This decoy model is typified by the NB-LRR protein Prf and the host protein Pto, a Ser/Thr kinase which interacts with Prf and mediates detection of the AvrPto and AvrPtoB proteins of *P. syringae*. The Bait hypothesis is an extension of the Decoy model and suggests that an NB-LRR protein first interacts with a 'bait' or effector target protein to facilitate direct interaction between it and the effector (Collier & Moffett, 2009).

#### 1.4.1.2 Direct recognition

Though there are fewer reports that indicate direct R-Avr protein interaction, the examples that do exist boast *in vitro* evidence to support their claims. In particular, and of most relevance to this study, are the examples of the flax rust resistance proteins encoded at the *L* locus and their cognate effector, AvrL567. Thirteen different allelic specificities are encoded at the *L* locus and 12 at the *Avrl567* locus (Dodds *et al.*, 2006; Lawrence *et al.*, 2007). Using the yeast two-hybrid and *in planta* co-expression systems only those proteins that displayed a direct interaction in yeast developed plant resistance responses (Wang *et al.*, 2007). Interestingly, a functional nucleotide binding site domain is also necessary for effector interaction, suggesting that the LRR domain alone is not sufficient for effector binding (Dodds *et al.*, 2006).

Direct interaction *in planta* has also been demonstrated between the *A. thaliana* R protein RPP1-WsB and its cognate effector *Arabidopsis thaliana* Recognised 1 or ATR1 from the oomycete *Hyaloperonospora arabidopsidis*. Using co-immunoprecipitation this study demonstrated that the LRR region of RPP1-WsB meditates the interaction with ATR1 and that this interaction occurs in a race specific manner. RPP1-WsB only interacted with recognised ATR1 alleles. It was additionally

found that the LRR region alone can interact with ATR1, but this is not sufficient to trigger the HR (Krasileva *et al.*, 2010).

### 1.4.2 The molecular switch model of NB-LRR protein activation

Whether the R/Avr protein interaction occurs directly or indirectly, this event is widely believed to be the trigger for activation of the NB-LRR. The mechanisms behind their activation and the regulation of the on/off switch are the subjects of a great deal of research. Advancement in this area has been hindered by the complicated nature of recombinant R protein production and biochemical analysis. As mentioned above, despite more than 50 NB-LRRs being cloned (Lukasik & Takken, 2006), recombinant protein production and nucleotide binding have only been reported for six NB-LRRs; I-2, Mi-2, N (tobacco), M, L6 and Mla27 (Maekawa *et al.*, 2011a; Tameling *et al.*, 2002; Ueda *et al.*, 2006; Williams *et al.*, 2011).

The following activation model is the most widely accepted. It brings together models proposed by many researchers and in it R proteins are represented as signalling hubs that combine receptor, regulatory switch and signal transducer into one. It is based on the foundations of the crystal structures of the STAND AP-ATPase family members, Apaf-1 & CED-4, biochemical data from the I-2, Mi-1, M, L6 and Mla27 R proteins and the STAND MalT transcriptional activator from *E. coli*, and functional studies of Rx (Leipe *et al.*, 2004; Maekawa *et al.*, 2011a; Marquenet & Richet, 2007; Moffett *et al.*, 2002; Rairdan & Moffett, 2006; Riedl *et al.*, 2005; Tameling *et al.*, 2002; Tameling *et al.*, 2006; Williams *et al.*, 2011; Yan *et al.*, 2005).

In this model, activation is mediated by a molecular switch found within the NB-ARC domain, which is triggered by pathogen perception in the LRR and the subsequent switch of ADP for ATP. ATPase activity resets the protein back to its inactive state. Activation allows the TIR or CC domain, often through dimerisation, to initiate downstream signalling events. So far, no direct evidence has been provided to support the claim that an effector triggers nucleotide exchange in the R protein upon its recognition (Figure 1-3).

In the absence of their cognate effector protein, an R protein is resting in a closed, ADP-bound state. This autoinhibited conformation is regulated by the LRR and the

ARC2 (HETHS) sub-domain, and possibly the CC/TIR domains, in combination with bonds between residues of the NB-ARC domain and the bound ADP. Recognition of the effector by the LRR domain (or sensing of an interaction with associated proteins in cases of indirect detection) triggers a conformational change within the NB-ARC that is communicated by intramolecular interactions that exist between these regions. This conformational change releases the negative regulation and results in the exchange of the bound ADP molecule with ATP. Binding of ATP opens up the protein and facilitates a further change of conformation that releases the N-terminal domain. This allows it to interact with downstream partners and induce the resistance signalling pathways. For some R proteins this initiation of signalling involves homodimerisation of their N-terminal domains, a process by which a signalling and interaction platform is formed (Ade *et al.*, 2007; Bernoux *et al.* 2011a; Gutierrez *et al.*, 2010; Maekawa *et al.*, 2011a; Mestre and Baulcombe, 2006). ATP hydrolysis, facilitated by the NB sub-domain, returns the R protein to its inactive or resting state.

As noted above, different modes of activity have been demonstrated and/or suggested for the NBS. Therefore, whilst the molecular switch model is supported by some of the biochemical evidence to date, it most likely only represents the activation pathway of a limited number of R proteins, and maybe missing some important steps.

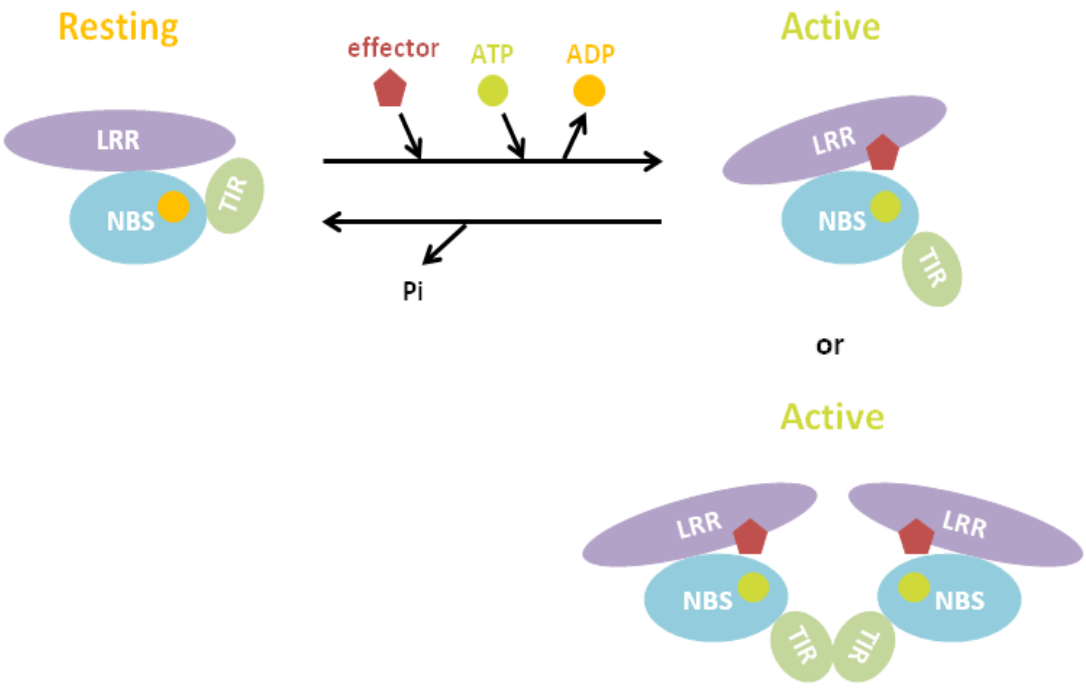



**Figure 1-3: The molecular switch activation model**

In this activation model, conformational changes upon effector recognition result in nucleotide exchange within the NB-ARC domain and this triggers further conformational changes that ultimately release the N-terminal domain for the initiation of signalling. Signal initiation may, or may not, involve dimerisation. The ATPase activity of the NBS domain cycles the activated protein back to its ADP-bound resting state, thereby terminating the signalling process, or perhaps facilitating further activation and signalling cycles.

## 1.5 The M flax rust resistance protein

The M protein from flax is a member of the TIR-NB-LRR class and it is located, due to an N-terminal signal anchor, in the tonoplast membrane (Takemoto *et al.*, 2012). In the Y2H system is has been shown to interact directly with its flax rust (*M. lini*) cognate effector AvrM (*AvrM-A*), but not the virulence effector avrM (Catanzariti *et al.*, 2010). Despite being 86% similar to the L6 flax rust resistance protein, which has a TIR domain that can form a dimer and is sufficient to elicit effector-independent HR (Bernoux *et al.*, 2011a), the TIR domain of M does not form a homodimer in the Y2H system and is not autoactive (M. Bernoux, personal communication, 2012; Figure 1-6).

Recombinant M is purified with an ADP molecule bound within its NB-ARC and this nucleotide binding is dependent on a functional P-loop. An autoactive mutant, $M^{D555V}$, is purified with bound ADP and ATP, but displays a preference for bound ATP (Williams *et al.*, 2011). This preference of an active protein for ATP was the first direct evidence that the ATP-bound form is the active state and thus provided much needed support for the molecular switch model of activation.

Recent work by Sornaraj (2013) has demonstrated that the M protein possesses intrinsic ATPase activity and that the autoactive $M^{D555V}$ mutant has a faster turnover rate ($K_{cat}$) than wild-type M. $M^{D555V}$ converts 2.63 molecules of ATP into ADP every minute, whilst M converts only 0.39 ATP molecules per minute. The inactive P-loop mutant, $M^{K286L}$, has both a decreased maximum hydrolysis rate and affinity for ATP compared to wild-type M (Sornaraj, 2013; Table 1-2).

Together, these results support the classical molecular switch model, in which M protein exists in an ADP-bound resting state and an ATP-bound active state. However, they also provide evidence that ATP hydrolysis most likely resets or recycles M, either to, turn signalling off, or enable further rounds of effector-mediated activation. Furthermore, the results suggest that the M protein may not exist in an inactive state in the absence of a pathogen, rather that it is constantly cycling between ADP- and ATP-bound states. Whether M's effector, AvrM, triggers activation by promoting nucleotide exchange to sustain the protein's ATP-bound state (Figure 1-4), or by

increasing the ATP hydrolysis rate has not yet been conclusively demonstrated (Figure 1-5). However, evidence that the autoactive mutant $M^{D555V}$ responds to AvrM *in planta* with an increased HR intensity and very preliminary evidence that it may have an increased hydrolysis rate in the presence of AvrM perhaps suggests the latter (Sornaraj, 2013).

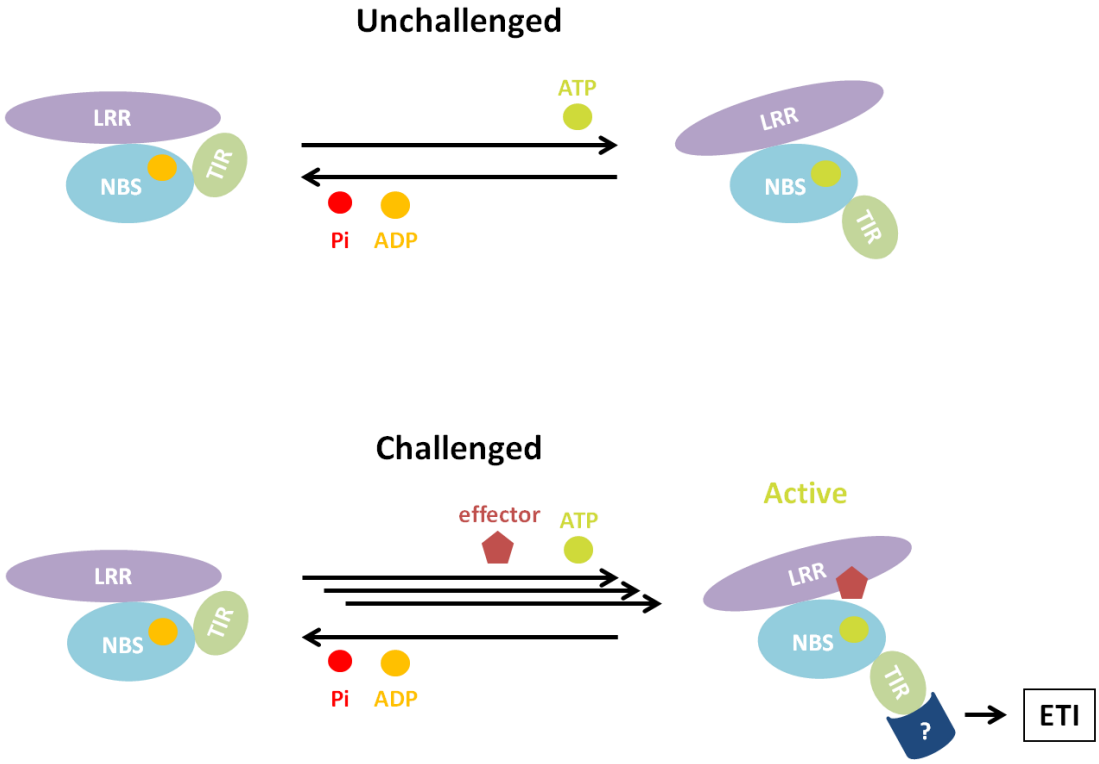



**Figure 1-4: The sustained ATP model of activation**

As a consequence of its intrinsic ability to hydrolyse ATP the M protein continuously cycles between ATP and ADP bound forms within an unchallenged cell. Interaction with AvrM at the LRR domain changes the nucleoide cycling equilibrium and holds the protein in the active ATP-bound state to trigger ETI activation. ATP hydrolysis is used as a re-sets switch only.

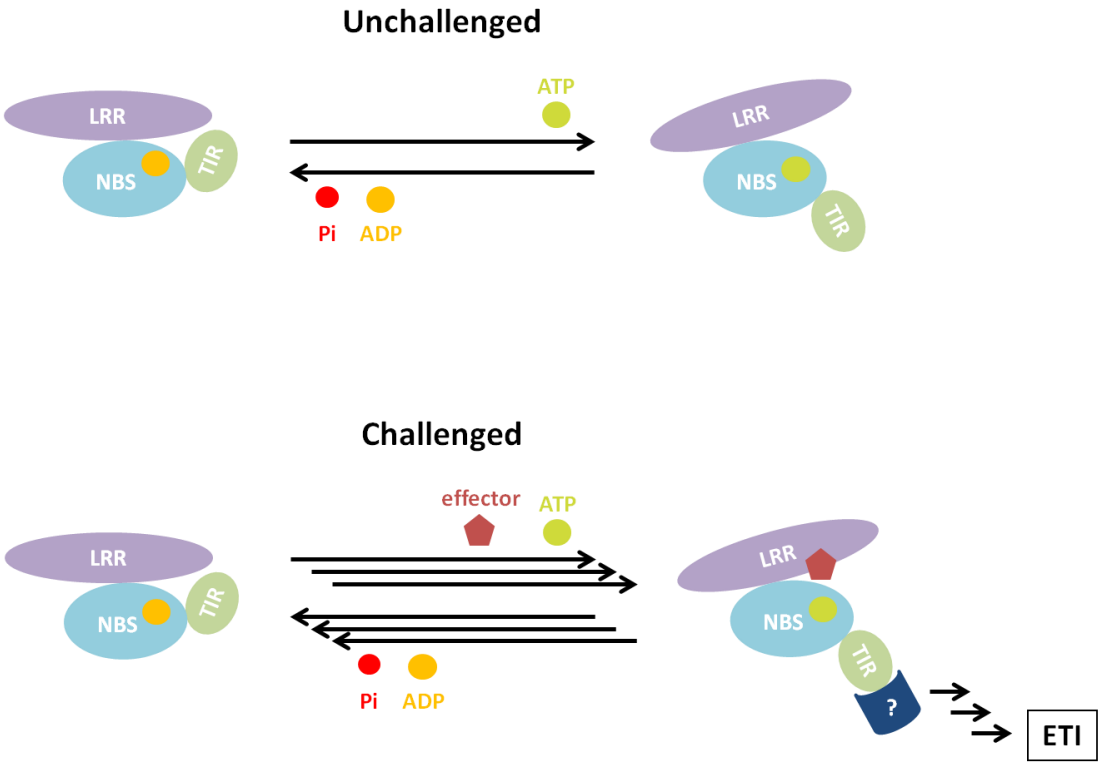



**Figure 1-5: The continuous cycling model of activation**

As a consequence of its intrinsic ability to hydrolyse ATP the M protein continuously cycles between ATP and ADP bound forms within an unchallenged cell. This background level of signalling is below an HR activation threshold. Interaction with AvrM at the LRR domain increases the ATP hydrolysis rate, breaking the activation threshold. The mechanochemical energy released from the ATP hydrolysis reaction drives a conformational change that releases the TIR domain so that it can interact with downstream protein partners, which ultimately results in activation of the HR pathways and ETI.

**Table 1-2: Kinetics of the ATPase activity of wild-type M, the inactive $M^{K286L}$ mutant and the autoactive $M^{D555V}$ mutant**

| Kinetic Properties | M | $M^{K286L}$ | $M^{D555V}$ |
|---|---|---|---|
| $K_m$ (µM) | 4.5 ± 1.4 | 28 ± 4.6 | 9.4 ± 1.4 |
| $V_{max}$ (pmol ATP/min/µg) | 2.4 ± 0.1 | 2.9 ± 0.2 | 10.1 ± 3.6 |
| $K_{cat}$ | 0.39 | n/a | 2.63 |

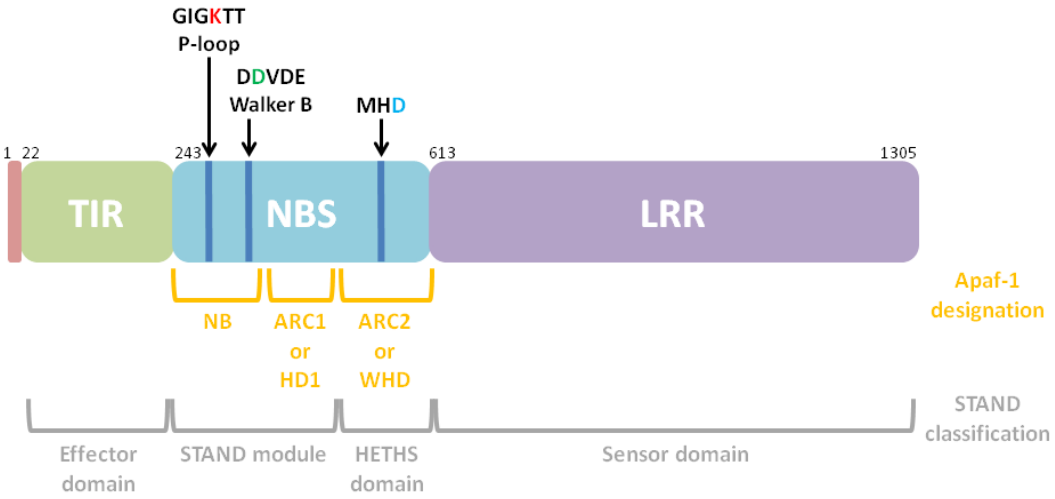



**Figure 1-6: Schematic representation of the M flax rust resistance protein**

The M protein contains an N-terminal 21 amino acid signal anchor (coloured red) that targets it to the tonoplast membrane. The amino acids that make up the highly conserved P-loop, Walker B and MHD motifs are highlighted. The invariant lysine of the P-loop (coloured red) when substituted with leucine has been shown to inactivate M and significantly reduce bound nucleotide levels. Substitutions of the second conserved aspartate in the Walker B, coloured green, with glutamate does not impact on M function; however, it does increase the amount of ADP found within the NBS of in vitro purified protein. Substitution of the aspartate (coloured blue) within the MHD motif with valine causes autoactivity and also increases the ATPase turnover rate ($K_{cat}$) (Sornaraj, 2013).

### 1.5.1 Aims

While it has been known for some time now that some plant NB-LRR proteins can bind nucleotides and perform ATP hydrolysis, the roles that these molecular activities play in NB-LRR activation remain poorly understood and controversial. Although the current molecular switch activation model is widely accepted, it is based on data from a limited number of R proteins and needs to be refined, redefined and/or expanded. It was the aim of this study to investigate the activation mechanisms of NB-LRR proteins by focusing on the M protein. It was also hoped that other plant NB-LRRs could also be recombinantly expressed and studied. As the study evolved and some experimental roadblocks were encountered, new aims were set. The broad objective was, however, to explore the role of functional domains in the NB pocket of an R protein and this was addressed with the following experimental aims, which form the foundation of the results chapters presented in this thesis:

1. To explore the heterologous production of plant NB-LRR proteins other than M using the *Pichia pastoris*-based heterologous expression system developed for M protein;
2. To characterise the roles of the acidic residues within the Walker B motif of the nucleotide binding site domain in order to gain a better understanding of the mechanism behind the ATP hydrolysis reaction and its role in the activation pathway;
3. To investigate biochemical activities of the M protein other than ATP hydrolysis.

# 2 Materials and Methods

## 2.1 Materials

### 2.1.1 Water

Milli-Q water supplied by the School of Biological Sciences at Flinders University was used in all experiments.

### 2.1.2 Chemicals and reagents

Yeast extract, tryptone and bacteriological peptone were obtained from Oxoid (Adelaide, South Australia, Australia). Glycine, Sodium Dodecyl Sulphate (SDS), Bacteriological Agar, D-glucose, Tween 20, Triton X-100 (TX-100) Tris-base and Tris ($Cl^-$) were all purchased from Amresco (Solon, Ohio, U.S.A.). All other chemicals were purchased from Sigma-Aldrich (Castle Hill, New South Wales, Australia).

### 2.1.3 Cloning and expression strains

- *Escherichia coli* strain DH10B
- *Pichia pastoris* X33 (Invitrogen; Mulgrave, Victoria, Australia)
- *Agrobacterium tumefaciens* strain GV3101.

### 2.1.4 Molecular biology reagents

Restriction enzymes, λ*HIND* III DNA ladder, Phusion® High-Fidelity DNA Polymerase, *Taq DNA* Polymerase and dNTPs were purchased from New England Biolabs Pty. Ltd. (NEB) (Ipswich, Massachusetts, U.S.A.). Agarose and ethidium bromide were purchased from Amresco.

### 2.1.5 Protein chemistry reagents

Prestained Protein Marker, Broad Range (7-175 kDa) was purchased from NEB. SuperSignal® West Pico Chemiluminescence Substrate was supplied from Thermo Fisher Scientific (Rockford, Illinois, U.S.A.). KODAK Medical X-ray film, General Purpose Blue, RP X-Omat LO Developer and Fixer were obtained from DLC Australia (Hoppers Crossing, Vic., Australia). Nitrocellulose Hybond ECL membrane was from GE Healthcare Life Sciences (Rydalmere, NSW, Australia).

### 2.1.6 Primers

All primers used in this research were manufactured and supplied in lyophilized form by Geneworks (Adelaide, S.A., Australia) and were resuspended in sterilised Milli-Q water to a concentration of one microgram per millilitre. Primers were used in PCR at a final concentration of 100 ng/µL and in sequencing reactions at 40 ng/µL. A list of all primers can be found in Table 2-1 and Table 2-2.

### 2.1.7 Protein expression, purification and biochemical assays

French press (Aminco, Urbana, IL, U.S.A), Ni Sepharose 6 Fast Flow resin (GE Healthcare Life Sciences; Rydalmere, NSW, Australia), 30 kDa MWCO concentrators (Millipore; Kilsyth, Vic., Australia), Sigma-Aldrich liquid chromatography column, 1250 luminometer (BioOrbit, Turku, Finland), Bio-Rad Molecular Imager, VersaDoc and protein assay reagent. *Strep*-Tactin® high capacity resin and strep regeneration buffer with HABA were supplied by Fisher Biotec (Wembley, Western Australia, Australia). ATP, [$\alpha$-$^{32}$P] and ATP, [$\gamma$-$^{32}$P] 10mCi/mL, 250 µCi (3000 Ci/mmol) were from Perkin Elmer (Glen Waverley, Vic., Australia). Thin layer chromatography Polyethylenimine (PEI) Cellulose F plastic plates from Merck, Darmstadt, Germany. Myosin, calcium activated from rabbit muscle was from Sigma-Aldrich.

## 2.2 General Buffers and Media

### 2.2.1 Molecular biology buffers

- TrisOAc EDTA (TAE): 40 mM TrisOAc, 1 mM EDTA($Na^+$) pH 8.0

### 2.2.2 SDS-PAGE, Western transfer and immunoblot

- 3X sample buffer: 62.5 mM Tris ($Cl^-$) pH 6.8, 20% (w/v) glycerol, 2% (w/v) SDS, 5% (v/v) β-mercaptoethanol, 0.5% (w/v) bromophenol blue
- running buffer: 25 mM Tris-base pH 8.3, 192 mM glycine, 0.1% (w/v) SDS
- transfer buffer: 25 mM Tris-base pH 8.3, 152 mM glycine
- blotto: 5 mM Tris-base, 0.1% (v/v) Tween, 150 mM NaCl and 5% (w/v) skim milk powder
- TBS-T: 5mM Tris-base, 0.1% (v/v) Tween, 150mM NaCl

- Coomassie blue: 0.1% (w/v) Coomassie blue, 40% (v/v) methanol, 10% (v/v) glacial acetic acid

## 2.3 Buffers and Media for *In Vitro* Protein Expression, Purification and Biochemical Assays

### 2.3.1 Cultures

*E. coli*

- half salt LB: 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl, 40 µg/mL zeocin

*P. pastoris*

- YPD: 1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) dextrose, 100 µg/mL zeocin
- BM(G/M)Y: 1% (w/v) yeast extract, 2% (w/v) peptone, 100 mM potassium phosphate buffer, pH 6.0, 1.34% (w/v) yeast nitrogen base, 4 x $10^{-5}$% (w/v) biotin, 1% (w/v) glycerol or 0.5% (v/v) methanol, 100 µg/mL ampicillin

### 2.3.2 Nickel affinity chromatography

- lysis buffer: 50 mM Tris-base pH 8.0, 100 mM NaCl, 10% (w/v) glycerol, 0.25 mM Triton X-100 (TX-100), 10mM MgOAc, 10 mM imidazole
- equilibration buffer: 20 mM Tris-base pH 8.0, 0.25 mM TX-100, 130 mM NaCl, 10 mM imidazole, 5 mM β-mercaptoethanol
- wash buffer: 20 mM Tris-base pH 8.0, 0.25 mM TX-100, 130 mM NaCl, 55 mM imidazole, 5 mM β-mercaptoethanol
- elution buffer: 20 mM Tris-base pH 8.0, 0.25 mM TX-100, 130 mM NaCl, 250 mM imidazole, 5 mM β-mercaptoethanol
- gel filtration buffer: 20 mM Tris-base pH 7.5, 150 mM NaCl, 10% (w/v) glycerol, 10 mM MgOAc, 1 mM DTT
- protease inhibitors: 100 mM PMSF, 100 mM benzamide, 100 mM benzamidine HCL and 200 mM ε-amino-n-caproic acid

### 2.3.3 *Strep*-Tactin® chromatography

- strep binding buffer: 100 mM Tris ($Cl^-$) pH 8.0, 150 mM NaCl
- strep elution buffer: 2.5 mM D-desthiobiotin, 50 mM Tris ($Cl^-$) pH 8.0, 50 mM NaCl, 10 mM $MgCl_2$, 10% (w/v) glycerol
- strep regeneration buffer: buffer R with HABA
- hydrolysis assay buffer B: 50 mM Tris ($Cl^-$) pH 8.0, 50 mM NaCl, 10 mM $MgCl_2$, 10% (w/v) glycerol

### 2.3.4 Nucleotide quantification assay

- pyruvate kinase buffer: 125 mM TrisOAc pH 7.4, 5 mM phosphoenolpyruvate, 2.5 mM $MgSO_4$, 31 units pyruvate kinase

### 2.3.5 ATP hydrolysis & autophosphorylation assays

- hydrolysis assay buffer A: 50 mM Tris ($Cl^-$) pH 7.5, 50 mM NaCl, 10 mM $MgCl_2$, 10% (w/v) glycerol
- Myosin assay buffer: 20 mM Tris ($Cl^-$) pH 6.8, 10 mM $CaCl_2$, 600 mM KCl

## 2.4 Buffers and Media For *In Planta* Protein Expression & Analysis

### 2.4.1 Cultures

*E. coli*

- pEG100 vectors: LB with 50 µg/mL kanamycin
- pTNSpec vectors: LB with 5 µg/mL tetracycline
- pRK2013 (helper strain): LB with 50 µg/mL kanamycin

*A. tumefaciens*

- pEG100 vectors: LB with 25 µg/mL rifampicin, 50 µg/mL gentamycin, 50 µg/mL kanamycin
- pTNSpec vectors: LB with 25 µg/mL rifampicin, 50 µg/mL gentamycin, 5 µg/mL tetracycline

### 2.4.2 Transient protein expression and extraction

- infiltration buffer: 10 mM MES pH 5.6, 10 mM $MgCl_2$, 200 µM acetosyringone
- extraction buffer: 0.24 M Tris ($Cl^-$) pH 6.8, 6% (w/v) SDS, 30% (w/v) glycerol, 16% (v/v) β-mercaptoethanol, 0.006% (w/v) bromophenol blue, 10 M urea.

## 2.5 General Molecular Biology Methods

### 2.5.1 Agarose gel electrophoresis

1% (w/v) agarose gels were made and run in 1X TAE buffer (40 mM TrisOAc, 1 mM EDTA($Na^+$) pH 8.0). Gels were run at 100 V for 40-60 minutes and stained in ethidium bromide at a concentration of one microgram per millilitre. Bands were visualised with ultraviolet light in a UV transilluminator 2000 (Bio-Rad; Gladesville, NSW, Australia).

### 2.5.2 Plasmid purification

The Wizard® *Plus* SV Minipreps DNA Purification System (Promega; Alexandria, NSW, Australia) was used to purify plasmids from transformed *E. coli* clones as per the manufacturer's instructions.

### 2.5.3 DNA sequencing and analysis

All sequencing reactions were performed by the Australian Genome Research Facility Ltd. (AGRF) and contained 200-500 ng of plasmid DNA and 40 ng of primer (refer to Table 2-2 for a list of all sequencing primers). Analysis of raw sequencing data was performed using Sequencer® (Genecodes Corporation; Michigan, U.S.A.).

## 2.6 General Protein Chemistry Methods

### 2.6.1 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were separated in 10% (w/v) polyacrylamide gels using a Bio-Rad Laboratories Mini-Protean® Tetra Cell gel electrophoresis unit. Electrophoresis was conducted at 170 V for one hour with running buffer (25 mM Tris-base pH 8.3, 192 mM Glycine, 0.1% (w/v) SDS) (Sambrook *et al.*, 1989). All protein samples were mixed with 3X sample buffer (62.5 mM Tris ($Cl^-$) pH 6.8, 20% (w/v) glycerol, 2% (w/v) SDS, 5% (v/v) β-mercaptoethanol, 0.5% (w/v) bromophenol blue) and denatured at 98°C for five minutes. Prestained Protein Marker, Broad Range (7-175 kDa) (NEB) was used as the reference for protein size.

### 2.6.2 Coomassie blue stain

Acrylamide gels were fixed in 10% (v/v) glacial acetic acid and 40% (v/v) methanol for 10-30 minutes, stained with a solution of Coomassie blue (0.1% (w/v) Coomassie blue, 40% (v/v) methanol, 10% (v/v) glacial acetic acid) for 60 minutes and destained in a solution of 10% methanol, 40% (v/v) glacial acetic acid for 30-60 minutes.

### 2.6.3 Western transfer and immunoblot

Proteins were transferred to Hybond-ECL nitrocellulose membrane (GE Healthcare Life Sciences) in Bio-Rad Mini Trans-Blot® apparatus at 60 V for 90 minutes in transfer buffer (25 mM Tris-base pH 8.3, 152 mM glycine). Membranes were blocked for one hour at room temperature or overnight (12-16 hours) at 4$^{o}$C with Blotto (5 mM Tris-base, 0.1% (v/v) Tween, 150 mM NaCl and 5% (w/v) skim milk powder). The appropriate primary antibody was diluted in Blotto and incubated with the membrane at room temperature for 45 minutes. This was followed by two 15-minute washes of the membrane with Blotto, one 45-minute incubation with the appropriate secondary antibody conjugated to horseradish peroxidase (HRP) and two 10-minute TBS-T (5mM Tris-base, 0.1% (v/v) Tween, 150mM NaCl) washes of the membrane. HRP signal was visualised with SuperSignal West Pico chemiluminescence substrate (Thermo Fisher Scientific) and membranes were exposed to KODAK Medical X-ray film, General Purpose Blue. Films were developed with a KODAK X-omat 1000 developer with KODAK RP X-omat fixer and developer.

## 2.7 Methods for *In Vitro* Protein Expression & Purification

The methods for nickel affinity purification, nucleotide quantification, and Sypro®Ruby protein estimation were carried out according to the methods described in Williams *et al.* (2011).

### 2.7.1 Expression constructs

All constructs used to express recombinant protein in the heterologous host *P. pastoris* were constructed in the pPICZa vector (Invitrogen; Mulgrave, Vic., Australia) in which expression is driven by the methanol-inducible *AOX1* promoter.

#### 2.7.1.1 M and L6 constructs

Protein expression constructs for nucleotide occupancy determination comprised the *M* or *L6* gene coding sequences minus the first 63 base pairs (21 amino acids) for M and the first 87 (29 amino acids) for L6, a sequence encoding a start codon and an N-terminal 9X histidine tag downstream of the *AOX1* promoter (Schmidt *et al.* 2007). Constructs used to produce protein for ATP hydrolysis and autophosphorylation assays contained the *M* gene coding sequence minus the first 63 base pairs with an N-terminal 6X histidine tag and a C-terminal 1X *Strep*-Tag® (WSHPQFEK) downstream of the *AOX1* promoter.

#### 2.7.1.2 RPS2

The *RPS2* and *RPS2 CC-NBS* coding sequences were obtained from Brian Staskawicz and Doug Dahlbeck at UC Berkeley. The constructs were as follows; full-length RPS2, RPS2 minus the first 75 base pairs, the CC and NBS domains only, the CC and NBS minus the first 75 base pairs and the NBS and LRR domains. Using PCR a sequence encoding a 9X histidine tag and a *Bst*BI cut site was added to the N-terminus and a sequence encoding a stop codon and *Not*I cut site to the C-terminus of each construct (Table 2-1). Each construct was then cut with *Bst*BI and *Not*I to prior to ligation into pPICZa.

#### 2.7.1.3 RPP1B

Sequenced pPICZa RPP1B constructs were obtained from Ksenia Krasileva. Four expression constructs were trialled; full-length RPP1B, RPP1B minus the first 180 base

pairs, the NBS and LRR domains only and the LRR domain alone. Each construct contained an N-terminal sequence encoding a 9X histidine tag.

### 2.7.2 Site-directed mutagenesis PCR and *E. coli* transformation

Mutations were made using mismatch primers and site-directed mutagenesis PCR with the following conditions: 98°C, 30 seconds; 30 cycles of 98°C for 10 seconds, primer dependent annealing temperature for 15 seconds; 72°C for three and a half minutes; 72°C for five minutes. Electrocompetent *E. coli* DH10B cells were prepared according to the methods of Sambrook *et al.* (1989). Mutated plasmid DNA (1-100 ng) (after *Dpn* I digestion) was mixed with 20 µL of electrocompetent *E. coli* DH10B cells and electroporation was performed with a Bio-Rad Laboratories Cell-Porator Voltage Booster according to the manufacturer's instructions. Cells were incubated in 200 µL of LB broth at 37°C for 45 minutes and grown on half salt LB plates with zeocin (40 µg/mL) for 12-16 hours at 37°C to select for transformants containing the expression vectors of interest. Plasmids were isolated from four independent colonies and the entire gene coding sequence was sequenced to confirm the integrity of the sequence and/or the presence of the desired mutation.

### 2.7.3 *P. pastoris* transformation

Electrocompetent *P. pastoris* X-33 cells were prepared on the day of transformation as per the methods adapted from those described in the Invitrogen 'EasySelect™ *Pichia* Expression Kit' manual. A 10 mL *P. pastoris* X-33 YPD starter culture was grown for 12-16 hours at 30°C and 200 rpm. This culture was used to inoculate 90 mL of 2X YPD and the subsequent culture was grown as above until it reached an $OD_{600\ nm}$ between 0.6 and 0.8. At this time the cells were harvested via 10-minute centrifugation at 4°C and 3000 rpm. Cells were washed twice by resuspension and subsequent centrifugation with 50 mL followed by 25 mL of sterile, ice-cold water and once with two millilitres of one molar sorbitol before being resuspended in 200 µL of one molar sorbitol and stored at 4°C. pPICZ expression vectors (5-10 µg) were linearised by digestion with *Pme* I according to the instructions from NEB, ethanol precipitated (Sambrook *et al.*, 1989), dried in a vacuum concentrator and resuspended in 10 µL of sterile water. Linear DNA was mixed with 80 µL of the freshly

prepared electrocompetent *P. pastoris*, transferred to an ice-cold 0.2 cm electroporation cuvette (Cell Projects; Kent, United Kingdom) and incubated on ice for five minutes. Electroporation was performed using the Bio-Rad Gene Pulser II as per the manufacturer's instructions (25 µF, 200 Ω, 1.5 kV). Transformed cells were immediately suspended in one millilitre of ice-cold 1M sorbitol in a sterile eppendorf and incubated for one hour at 30$^{o}$C. One millilitre of YPD was added and incubation was continued for a further 90 minutes with gentle shaking. Transformed colonies were grown on YPD plates with zeocin (100 µg/mL) at 30$^{0}$C for two days to select for transformants containing the expression vectors.

### 2.7.4 Test expression

*P. pastoris* colonies (six-to-eight cream, circular colonies) were grown in 10 mLs of BMGY containing 100 ug/µL of zeocin at 30$^{o}$C and 200 rpm for 72 hours. Cells were harvested by centrifugation for 10 minutes at 3000 rpm and 4$^{o}$C and the media was replaced with 10 mLs of BMMY to induce protein expression. Protein induction was conducted at 15$^{o}$C and 200 rpm for 72 hours with the addition of methanol to 0.5% (v/v) every 24 hours.

To extract intracellular proteins, cells were harvested as described above and resuspended in 200 µL of lysis buffer (50 mM Tris-base pH 8.0, 100 mM NaCl, 10% (w/v) glycerol, 0.25 mM Triton X-100 (TX-100), 10mM MgOAc, 10 mM imidazole). Two hundred microlitres of 425-600 µm of glass beads (pre-washed in lysis buffer) was added and the cells were lysed by a two-minute vortex at maximum speed. This was followed by centrifugation, as described above, to pellet cell debris. Equal amounts of total protein (30-40 µg), as determined by a total protein estimation, were separated by 10% SDS-PAGE. An immunoblot was performed to assess protein expression levels. Primary antibody: affinity purified rabbit anti-6X His epitope tag (Rockland; Gilbertsville, Pennsylvania, U.S.A.), dilution of 1:1000. Secondary antibody: peroxidase conjugated affinity purified goat anti-rabbit IgG (Rockland), dilution of 1:10 000. The colony from which the strongest chemiluminescent signal was obtained was used in the large-scale protein production of each mutant (see below).

### 2.7.5 Large-scale protein expression

The method described above for test expression was followed, except that 100 mL *P. pastoris* cultures were grown in 500 mL flasks, and cell pellets were washed once in cell wash buffer (50 mM Tris-base, pH 7.5, 130 mM NaCl, 10% (w/v) glycerol) before being snap frozen and stored at minus 80$^{o}$C. Note: a 100 mL *P. pastoris* culture typically resulted in a six gram cell pellet.

### 2.7.6 Nickel affinity chromatography

*P. pastoris* cells expressing M proteins were resuspended in three volumes of cell lysis buffer (50 mM Tris-base pH 8.0, 100 mM NaCl, 10% (w/v) glycerol, 0.25 mM, Triton X-100 (TX-100), 10mM MgOAc, 10 mM imidazole) supplemented with 20 mM β-mercaptoethanol, 100 mM PMSF, 100 mM benzamide, 100 mM benzamidine-HCL and 200 mM ε-amino-n-caproic acid. Lysis was performed by three sequential passes through a French Press (Aminco, Urbana, IL, U.S.A.) at 2000 to 6000 psi. Unlysed cells and cell debris were removed by centrifugation at 3000 rpm for 15 min at 4°C. The supernatant was ultracentrifuged at 45 000 rpm for 45 min at 4°C. The cleared lysate was titrated to a pH between 7.6 and 7.8 and loaded onto a five millilitre Ni-Sepharose 6 Fast Flow affinity column (GE Healthcare Life Sciences; Rydalmere, NSW, Australia) that had been pre-equilibrated with 5 column volumes (c/v) of equilibration buffer (20 mM Tris-base pH 8.0, 0.25 mM TX-100, 130 mM NaCl, 10 mM imidazole, 5 mM β-mercaptoethanol). The column was washed with 10 c/v of wash buffer (20 mM Tris-base pH 8.0, 0.25 mM TX-100, 130 mM NaCl, 55 mM imidazole, 5 mM β-mercaptoethanol). His-tagged proteins were eluted from the column with 5 c/v of elution buffer (20 mM Tris-base pH 8.0, 0.25 mM TX-100, 130 mM NaCl, 250 mM imidazole, 5 mM β-mercaptoethanol). To prevent the formation of soluble protein aggregates, as found by Williams *et al.* (2009) the total imidazole concentration was decreased to 75 mM by the addition of 70 mL of gel filtration buffer (20 mM Tris-base pH 7.5, 150 mM NaCl, 10% (w/v) glycerol, 10 mM MgOAc, 1 mM DTT) prior to protein concentration in an ultracentrifugal concentrator (30 kDa MWCO) (Millipore) Protein samples were concentrated to a final volume of one millilitre.

### 2.7.7 Nucleotide quantification assay

Nucleotide quantification was performed using the Sigma-Aldrich Adenosine 5'-triphosphate bioluminescent assay kit. It involves the luciferase-catalyzed conversion of ATP, luciferin, and $O_2$ to AMP, oxyluciferin, $CO_2$ and light, in which, using saturating levels of luciferin, luminescence is linearly proportional to ATP concentration. In triplicate, nickel affinity purified and concentrated M protein samples were denatured at 98°C for five minutes and centrifuged at 14 000 g for two minutes at 4°C to remove precipitated protein. Seventy five microlitres of pyruvate kinase buffer (125 mM TrisOAc, pH 7.4, 5 mM phosphoenolpyruvate, 2.5 mM $MgSO_4$) was added to 50 µL of each sample. For ADP measurements, 31 units of pyruvate kinase (PK) were added and samples were incubated at room temperature for 30 min. Samples were boiled for five minutes and centrifuged for two minutes at 14 000 g and 4°C to remove the PK. Ten microlitre samples were added to 100 µL of ATP assay mix, prepared as a 1:10 dilution using the buffer supplied. Luminescence output was measured in millivolts, using a 1250 luminometer (BioOrbit, Turku, Finland). A standard concentration of ATP ($2 \times 10^{-7}$ or $2 \times 10^{-8}$ M) was added to each reaction as an internal reference. ADP concentration was measured as the difference between the ATP detected in samples with and without PK. The average of triplicate measurements of ATP and ADP from an empty pPICZ vector purification (see appendix D) were subtracted from average ATP and ADP measurements of each protein. The ADP/ATP percent occupancy values represent the mean, plus or minus standard error as calculated from three or four independent *P. pastoris* cell pellets and protein purifications.

### 2.7.8 *Strep*-Tactin® chromatography

Protein eluted from nickel affinity chromatography was concentrated to ~20 mL, as above, before being loaded onto a 1 mL *Strep*-Tactin® column that had been pre-washed with strep binding buffer (100 mM Tris ($Cl^-$) pH 8.0, 150 mM NaCl). The column was washed with 20 c/v of strep binding buffer and protein eluted with 10 c/v of hydrolysis assay buffer B (50 mM Tris ($Cl^-$) pH 8.0, 50 mM NaCl, 10 mM $MgCl_2$, 10% (w/v) glycerol) into 4X the volume of hydrolysis assay buffer A (50 mM Tris ($Cl^-$) pH 7.5, 50 mM NaCl, 10 mM $MgCl_2$, 10% (w/v) glycerol) was concentrated to 500 µL

as described above. The *Strep*-Tactin® column was re-equilibrated with HABA buffer according to the manufacturer's instructions (Fisher Biotec).

### 2.7.9 ATP hydrolysis assay

ATP hydrolysis was measured at 25°C in the presence of [α-$^{32}$P]ATP and a cold (unlabelled) ATP concentration of 1 µM. All assays were performed were performed in LoBind protein tubes (Eppendorf, Australia) in triplicate on nickel affinity and *Strep*-Tactin® purified protein samples. Each 50 µL assay contained 200-400 ng of total protein (as estimated by Bio-Rad total protein assay (2.7.11)), hydrolysis assay buffer A (50 mM Tris ($Cl^-$) pH 7.5, 50 mM NaCl, 10 mM $MgCl_2$, 10% (w/v) glycerol) and 2 µL of cold ATP (1 µM), containing 0.4 µCi [α-$^{32}$P]ATP. Assays were started by the addition of ATP and at appropriate time intervals 2 µL aliquots were spotted onto polyethylenimine (PEI) cellulose thin layer chromatography (TLC) plates (Merck). [α-$^{32}$P]ATP and [α-$^{32}$P]ADP spots were resolved by developing the plates with a solution of 0.25 M LiCl and 2.5% (v/v) formic acid. Dry plates were exposed to a PhosphorImager plate (Molecular Dynamics; Amersham Biosciences) which was scanned in a Bio-Rad Molecular Imager. [α-$^{32}$P]ATP and [α-$^{32}$P]ADP spots were quantified using the Bio-Rad Quantity One software. To measure the rate of ATP hydrolysis the percent of ATP remaining and ADP formed was calculated and the percent formation of [α-$^{32}$P]ADP over time was plotted on a graph and the slope of a linear line of best fit was used as the rate of the ATPase reaction.

Myosin (approximately 8 µg; calcium activated from rabbit muscle; Sigma-Aldrich) was used as a positive control for ATP hydrolysis in all assays. These reactions were performed in an assay buffer containing 20 mM Tris-base pH 6.8, 10 mM $CaCl_2$, 600 mM KCl.

#### 2.7.9.1 Statistical analysis

The rates of ATP hydrolysis at 1 µM ATP were tested for outliers and the normality of the data tested by boxplot and Shapiro-Wilk test ($p$<0.05) respectively. Homogeneity of variance was assessed by Levene's Test for Homogeneity of Variances. As this assumption was violated ($p$ =0.000) a Welch Test of Equality of Means was used ($p$ =0.000). Games-Howell *post-hoc* analysis of one-way ANOVA ($p$ =0.05) was used to

assess whether the mean rates of ATP hydrolysis were statistically significantly different from one another. This analysis was performed with IBM SPSS Statistics for Windows.

### 2.7.10 Autophosphorylation assay

Reactions contained 0.4 µCi [γ-$^{32}$P]ATP or [α-$^{32}$P]ATP in 5 µM cold ATP, 250 ng of protein in hydrolysis assay buffer A (50 mM Tris ($Cl^-$) pH 7.5, 50 mM NaCl, 10 mM $MgCl_2$, 10% (w/v) glycerol) in a total volume of 50 µL. At appropriate time points 10 µL assay samples were mixed with 3X SDS-PAGE sample buffer containing 100mM DTT and immediately frozen. Proteins within the samples were separated by 8, 10 or 4-15% SDS-PAGE and the gels were then exposed to a PhosphorImager plate for an appropriate period of time. This plate was then scanned in a Bio-Rad Molecular Imager to identify radiolabelled areas in the gel. Gels were subsequently stained with SyproRuby®, as per the manufacturer's instructions, to visualise the protein bands. Samples (2 µL) were also taken for TLC analysis, as described in section 2.7.9.

### 2.7.11 Total protein assay

This assay to estimate total protein was carried out in accordance with the manufacturer's instructions (Protein Assay reagent, Bio-Rad Laboratories). BSA standards of 0, 50, 125, 250, 500 ng/µL were used to generate the standard curve.

### 2.7.12 Sypro®Ruby stain

In triplicate, 5-10 µL of concentrated purified protein sample was separated by SDS-PAGE and stained overnight in the dark with Sypro®Ruby (Bio-Rad Laboratories). Gels were destained for 45 minutes using a solution of 10% (v/v) ethanol and 7% (v/v) glacial acetic acid. Gels were visualised and images captured using the Bio-Rad VersaDoc system using the following settings: Sypro®Ruby 520LP UV TRANS with 1 gain and 1 bin and an appropriate exposure time. M protein bands were quantified using the Bio-Rad Quantity One software using a BSA standard curve for which the standards (100, 250, 500, 750 ng) were run on the same gel.

## 2.8 Methods for *In Planta* Protein Expression & Analysis

Plant growth conditions and methods for site-directed mutagenesis, *A. tumefaciens* transformation, *A. tumefaciens*-mediated transient protein expression, protein expression and immunoblot were carried out according to the methods described in Williams *et al.* (2011).

### 2.8.1 Expression constructs

Transient expression of Walker B M mutants was driven by the strong constitutive cauliflower mosaic virus 35S promoter (CaMV 35S) in the pEG100 expression vector (Earley *et al*. 2006). Constructs contained a mutated genomic *M* gene fused in-frame with sequence encoding a 3X C-terminal HA epitope tag (YPYDVPDYA). This construct was designated pEG100 gM:HA.

Expression of the $L6^{D350E}$ mutant, AvrM and avrM was driven by CaMV 35S in pTNSpec binary vectors. The $L6^{D350E}$ mutation was generated in the *L6* cDNA within a pTNSpec L6 vector supplied by Dr. Peter Dodds (CSIRO Plant Industry, Canberra) (see Table 2-1 for primers). pTNSpec binary vectors containing the *AvrM* and *avrM* genes were as described in Catanzariti *et al.* (2006).

The pE1776 vector containing the genomic *M* gene was obtained from Ann-Maree Catanzariti (UC Berkeley, California). Expression in this vector was driven by the very strong constitutive chimeric octopine and manopine synthase promoter. Transcription termination in all vectors was controlled with a sequence encoding the *nopaline synthase* terminator from *A. tumefaciens*.

### 2.8.2 Plant & plant growth conditions

*L. usitatissimum* (flax) cv. Hoshangabad was grown at 19°C with a 12 hour light/dark cycle in Plugger 222 soil (Debco; Tyabb, Vic., Australia) supplemented with Indoor, Pot and Planter slow release fertilizer (Osmocote; Baulkham Hills, NSW, Australia).

### 2.8.3 Site-directed mutagenesis PCR

Mutations were made using mismatch primers and site-directed mutagenesis PCR (Table 2-1) with the following conditions: 98°C, 30 seconds; 25 cycles, 98°C for 10

seconds, primer dependent annealing temperature for 15 seconds, 72°C for seven and a half minutes; 72°C for five minutes. *E. coli* transformation was carried out as described in 2.7.2 and coding regions were sequenced to confirm the presence of the desired mutations (Table 2-2).

### 2.8.4 *A. tumefaciens* transformation

Plasmid DNA (1-100 ng) was mixed with 20 µL of electrocompetent *A. tumefaciens* GV3101 cells and electroporation was performed with a Bio-Rad Laboratories Cell-Porator Voltage Booster according to the manufacturer's instructions. Cells were immediately incubated in 200 µL of LB broth at 37°C for 45 minutes before being plated onto LB plates containing rifampicin (25 µg/mL), gentamycin (50 µg/mL) and kanamycin (50 µg/mL) and incubated at 28 °C for 48 hours.

For the pTNSpec AvrM and avrM constructs, transformation was performed by tri-parental mating. *E. coli* pRK2013 (helper strain), recombinant *E. coli* carrying binary vector and *A. tumefaciens* GV3101 were each mixed with 400 µL of sterile milli-Q water. 50 µL of each of these mixtures were mixed together and 20 µL of this mixture was dropped onto a LB plate without antibiotic selection. After 48 hours of growth at 28°C a loopful of cells was streaked onto an LB plate with rifampicin (25 µg/mL), gentamycin (50 µg/mL) and tetracycline (5 µg/mL).

Multiple antibiotic resistant *A. tumefaciens* colonies were selected for liquid culture and infiltration into flax cotyledons.

### 2.8.5 *A. tumefaciens*-mediated transient expression (Agroinfiltration)

*A. tumefaciens* strains were cultured and harvested according to the methods described in Catanzariti *et al.* (2010). Cultures were resuspended in infiltration buffer (10 mM MES pH 5.6, 10 mM magnesium chloride, 200 µM acetosyringone) and infiltrations were performed with cultures at an optical density (OD) of 1.0. For co-infiltration experiments equal volumes *A. tumefaciens* expression strains were mixed together prior to infiltration. All infiltrations were performed on Hoshangabad cotyledons 11 days after planting. Cotyledons were harvested at appropriate time-

points, snap frozen in liquid nitrogen and stored at minus 80$^{o}$C for assessment of transient protein expression.

### 2.8.6 Protein extraction and immunoblot analysis

Two flax cotyledons were ground to a fine powder in liquid nitrogen and then in extraction buffer (0.24 M Tris ($Cl^-$) pH 6.8, 6% (w/v) SDS, 30% (w/v) glycerol, 16% (v/v) β-mercaptoethanol, 0.006% (w/v) bromophenol blue, 10 M urea) for 45 seconds. Cell debris was removed by centrifugation at 14 000 *g* for 10 minutes. Supernatant containing the HA-tagged protein was separated by 10% SDS-PAGE. Proteins were transferred into nitrocellulose membrane (GE Healthcare Life Sciences) and an immunoblot was performed using anti-HA purified monoclonal mouse antibody (clone 16B12) at a dilution of 1:1000 (Covance) and peroxidase conjugated affinity purified goat anti-mouse IgG (Rockland) at a dilution of 1:1000. HRP signal was visualised using SuperSignal West Pico chemiluminescence substrate, a KODAK X-omat developer and KODAK Medical X-ray film, General Purpose Blue. Coomassie blue stains of Ribulose-1,5-bis-phosphate carboxylase/oxygenase (RuBisCO) were used to indicate loading in SDS-PAGE.

**Table 2-1: Primers for site-directed mutagenesis and truncation**

| Primer | Sequence (5' to 3') |
|---|---|
| RPS2 F hisBstBI | TCATTCGAAAAAAAAATGGGACACCATCACCACCATCACCATCACCACGATT TCATCTCATCTCTT |
| RPS2 R stopNotI | GCAAGCGGCCGCCTAACTAGTATTTGGAACAAA |
| RPS2 F hisBstBI(-25) | TCATTCGAAAAAAAAATGGGACACCATCACCACCATCACCATCACCACGGAC ATAAGACTGATCTT |
| RPS2 NB-LRR F hisBstBI | TCATTCGAAAAAAAAATGGGACACCATCACCACCATCACCATCACCACGATG GCGGGTCAATTCAA |
| SDM M D363A F | CTTGTCGTTCTCGCTGATGTTGATGAGAAG |
| SDM M D363A R | CTTCTCATCAACATCAGCGAGAACGACAAG |
| SDM M D364A F | GTCGTTCTCGATGCTGTTGATGAGAAGTTT |
| SDM M D364A R | AAACTTCTCATCAACAGCATCGAGAACGAC |
| SDM M D364E F | GTCGTTCTCGATGAAGTTGATGAGAAGTTT |
| SDM M D364E R | AAACTTCTCATCAACTTCATCGAGAACGAC |
| SDM M D366A F | CTCGATGATGTTGCTGAGAAGTTTAAATTT |
| SDM M D366A R | AAATTTAAACTTCTCAGCAACATCATCGAG |
| SDM M D366E F | CTCGATGATGTTGAAGAGAAGTTTAAATTT |
| SDM M D366E R | AAATTTAAACTTCTCTTCAACATCATCGAG |
| SDM M E367A F | GATGATGTTGATGCTAAGTTTAAATTTGAA |
| SDM M E367A R | TTCAAATTTAAACTTAGCATCAACATCATC |
| SDM M D363A+D364A F | CTTGTCGTTCTCGCTGCTGTTGATGAGAAG |
| SDM M D363A+D364A R | CTTCTCATCAACAGCAGCGAGAACGACAAG |
| SDM M D366A+E367A F | CTCGATGATGTTGCTGCTAAGTTTAAATTT |
| SDM M D366A+E367A R | AAATTTAAACTTAGCAGCAACATCATCGAG |
| SDM L6 D350E F | CTTGTCGTTCTCGATGAAGTGGATGAGAAG |
| SDM L6 D350E R | CTTCTCATCCACTTCATCGAGAACGACAAG |

**Table 2-2: Sequencing primers**

| Primer | Sequence 5’ to 3’ |
|---|---|
| 5' AOX | GACTGGTTCCAATTGACAAGC |
| 3' AOX | GCAAATGGCATTCTGACATCC |
| M For 1 | CAGACTGGACCTTATCGAAAGGC |
| M For 2 | CGATGAGCGAACAACATTCGC |
| M For 3 | ATGCTATATGCTTGGGAGTCGG |
| M For 4 | GGACTCATAGAGCTTCGTCTCG |
| M For 5 | ATGCGAACTCCACGACCAAAC |
| M For 6 | GATCTGGATGTGATTGGATCCCC |
| M Rev 1 | GATTGCTCTGAGGAGGTTGACC |
| M Rev 2 | CCTTCTGCTCTTGCATTGCTC |
| M Rev 3 | CCGATGAAGAAGCAGGCTATATC |
| M Rev 4 | AAGAAAGTCTTTGGGAGACCCC |
| M Rev 5 | TCTCTATTCCTGGAAGCCACG |
| M Rev 6 | CGTGCAACCCTCTAAAATCAACC |
| L6 For 1 | GTGGATCCAAGTGACGTACGACATC |
| L6 For 3 | GGTGTCAAGTATGAGTTTAAGAGCG |
| L6 For 5 | CAAACAGTAGTAGTCCCCTCTATGG |
| L6 Rev 2 | GGCATCCTTCTCTTGTGTTTCACG |
| L6 Rev 4 | GCCAACAGTCAGAAAGGCGAACACG |
| L6 Rev 6 | GTAGTCTCCCCAACGACGTGCAACC |
| RPS2 For 1 | TCATCTCTTATCGTTGGCTGTGC |
| RPS2 For 2 | TGTCTGGGAAGAGATAGACTTGG |
| RPS2 For 3 | ATGCTAGTTGAGCCTAGCATGGG |
| RPS2 For 4 | CTGTAAGCCAAGATTGTCTGCGG |
| RPS2 Rev 1 | GACCCGCCATCTGTTTTGATAGC |
| RPS2 Rev 2 | TAGTTCATACCCTTCATCTCTGC |
| RPS2 Rev 3 | ACATATGGCATCCGTGGGATCG |
| RPS2 Rev 4 | CCTAGTTCTCAAGGTCTTCAGGC |

# 3 Heterologous Expression of RPP1B and RPS2

I would like to thank Ksenia Krasileva, Doug Dahlbeck and Brian Staskawicz from UC Berkeley for allowing me to utilise their *RPP1B* and *RPS2* constructs in this study. This work was conducted at Flinders University and also in UC Berkeley with Ksenia Krasileva who cloned the RPP1B sequences into the pPICZa vector and was involved in preliminary test expressions for the RPP1B proteins.

I also need to thanks Dr Simon Williams and the Protein Expression Facility at the University of Queensland for allowing me to share the results from the baculovirus/insect cell expression trial.

## 3.1 Introduction

Although more than 50 plant NB-LRR genes have been cloned, only a few have been expressed and purified as soluble protein from a heterologous expression system (Maekawa *et al.*, 2011a; Tameling *et al.*, 2002; Tameling *et al.*, 2006; Ueda *et al.*, 2006; Williams *et al.*, 2011). The production of proteins with whole domain truncations was initially more successful than the production of full-length R proteins. Truncated versions of the tomato R proteins I-2 and Mi-1, comprising the CC domain and a partial NBS domain, could be expressed in *E. coli*. Although these proteins did form inclusion bodies, a small proportion could be refolded and purified as soluble protein and studied in biochemical assays (Tameling *et al.*, 2002; Tameling *et al.*, 2006). A NB-LRR construct of the TIR-NB-LRR tobacco protein N could also be successfully expressed in *E. coli* and purified to levels suitable for use in biochemical assays. These truncated forms of I-2, Mi-1 and N were found to have intrinsic ATPase activity and provided the first evidence that the NBS is a functional nucleotide binding pocket capable of binding and hydrolysing nucleotides (Tameling *et al.*, 2002; Tameling *et al.*, 2006; Ueda *et al.*, 2006). While these studies did provide the first evidence of a biochemical property of plant NB-LRRs, they were limited by the refolding techniques that had to be employed to produce the heterologous proteins, and the relevance of the ATPase activity and nucleotide bound states of these truncated proteins to those properties of full-length NB-LRR proteins.

Most successful has been the expression and purification of R protein N-terminal domains, both CC and TIR. Recently the TIR domains of the L6, RPS5 and RRS1 proteins have been produced and crystallised, as has the CC domain of Mla10 (Bernoux *et al.*, 2011a; Maekawa *et al.*, 2011a; Ve *et al.*, 2011, Williams *et al.*, 2014). Whilst these structures provide some idea of potential signalling strategies, they do not provide any data about pre- or post-activation states of whole NB-LRRs.

Only one full-length R protein, Mla27 from Barley, has been expressed, purified and studied biochemically *in vitro* (Maekawa *et al.*, 2011). Flax M and L6 proteins have also been expressed, purified and their nucleotide binding capabilities studied *in vitro*, but these proteins are not entirely full-length, as production of soluble M and L6 proteins required the removal of the extreme N-terminal region of their TIR

domains. M is without its first 21 amino acids and L6 its first 29 (Schmidt *et al.*, 2007; Williams *et al.*, 2011). This hydrophobic region is conserved between some flax R proteins and is predicted to encode a signal anchor that localises M to the tonoplast membrane and L6 to the cytosolic face of the Golgi apparatus (Takemoto *et al.*, 2012). It is likely that such signal anchor sequences are recognised in heterologous non-plant eukaryotic expression systems like *P. pastoris*, and thus when they are removed, soluble cytosolic protein amenable to purification strategies is the result.

Biochemical data obtained from assays on the truncated proteins described above formed the basis for the molecular switch model of R protein activation. Later data from M, L6 and Mla27 provided further support for the model that was based on whole R proteins. Whilst there are similarities in the biochemical data obtained from the proteins described above, there are also differences. More recent data from Sornaraj (2013) indicates that the M protein contains intrinsic ATPase activity. This adds a modification to the molecular switch model, suggesting that this protein may not be held in inactive or active states, but rather cycles between these states and activation requires a shift in the equilibrium towards an ATP bound conformation. An outlier in the biochemical data obtained for R proteins is the tobacco mosaic resistance protein, N. It is purified bound to ATP (Ueda *et al.*, 2006), whilst the other NB-LRRs are purified ADP-bound (Maekawa *et al.*, 2011; Tameling *et al.*, 2002; Tameling *et al.*, 2006; Williams *et al.*, 2011).

In addition to biochemical differences, as more and more NB-LRR proteins are investigated *in planta*, contradictory functionally information is being uncovered. These differences are particularly evident in mutational studies in which mutations of equivalent residues in different R proteins have different functional consequences. For example, equivalent mutations to the proposed catalytic aspartate in the Walker B motif have been shown to autoactivate a CC-NB-LRR (Tameling *et al.*, 2006), but have no effect on the function of a TIR-NB-LRR (Williams *et al.*, 2011); this feature is explored further in the chapter 4. Additionally, some classes of R proteins, termed 'helper NB-LRRs', do not require a functional P-loop for their activity as a helper of signal transduction in ETI and their function in basal defence (Bonardi *et al.*, 2011). It now seems unlikely that all plant NB-LRR proteins share the same mechanisms of

activation and/or regulation. Nucleotide binding, ATP hydrolysis and other biochemical assays of many different R proteins as possible will help to refine, expand and/or redefine the current activation model.

### 3.1.1 RPP1B and RPS2 resistance proteins

The *RECOGNITION OF PERONOSPORA PARASITICA1-WsB* gene, here designated *RPP1B*, was identified in the *A. thaliana* ecotype Wassilewskija and encodes a protein with TIR-NBS-LRR domain architecture. It denotes disease resistance against *Hyaloperonospora arabidopsidis* (previously known as *Peronospora parasitica*) and recognises the cognate effector Arabidopsis thaliana Recognised1 (ATR1) (Botella *et al.*, 1998; Rehmany *et al.*, 2005). The *RPS2 A. thaliana* gene encodes a CC-NB-LRR protein that provides resistance against *P. syringae* strains carrying the *avrRpt2* avirulence gene (Axtell *et al.*, 2001). RPP1B has been shown to interact directly with its cognate effector ATR1 for detection, whilst RPS2 recognises the degradation of the RIN4 protein caused by its cognate effector AvrRpt2 in an example of indirect effector recognition (Axtell & Staskawicz, 2003; Krasileva *et al.*, 2010). RPS2 is also speculated to represent a plant NB-LRR that is held in a dormant state through interaction with partner proteins, rather than through an ADP-bound conformation and intramolecular interactions (Qi *et al.*, 2010).

The *RPP1B* gene contains four introns and the protein expression constructs used in this study were designed around the intron/exon boundaries as defined by cDNA analysis (Botella *et al.*, 1998). As *RPS2* does not contain any introns (Hammond-Kosak & Jones, 1997) the protein expression constructs were designed based on the borders of the CC, NBS and LRR domains. RPP1B contains a hydrophilic region of amino acids prior to the TIR domain (amino acids 1-60; Botella *et al.*, 1998) that also contains a putative nuclear localisation signal (NLS) (K. Krasileva, personal communication, 2009). In contrast, the extreme N terminus of RPS2 encodes a hydrophobic region (amino acids 7-22). The hydrophilic and NLS regions of RPP1B and the hydrophobic region of RPS2 were removed to see whether this assist in producing soluble cytosolic, near full length protein expression in *P. pastoris*.

### 3.1.2 Aims

In this study expression of the *A. thaliana* NB-LRR proteins RPS2 and RPP1B was trialled in the *P. pastoris* X-33 expression system that is used to produce soluble M and L6 proteins (used later in this study and by Schmidt, *et al.,* 2007; Sornaraj, 2013; Williams *et al.*, 2011). Additionally, in collaboration with Dr. Simon Williams at The University of Queensland, a trial of full-length RPS2 expression was conducted at the Protein Expression Facility at The University of Queensland in a baculovirus/insect cell system. Once expression of a construct was gained its purification would be optimised and its ATP and ADP occupancies and ATPase activity investigated.

This study represents the first time that the heterologous expression of NB-LRR proteins from a species other than flax, and of architecture other than TIR-NB-LRR has been undertaken in this laboratory. It would also represent the first study into the biochemistry of a protein that guards a host protein to indirectly detect its effector. In conjunction with data from M and L6 proteins, it was hoped that data obtained from these experiments could be used to support and refine, or possibly redefine, the current NB-LRR activation model.

## 3.2 Results

The *RPS2* sequences were amplified by PCR from vectors supplied by Doug Dahlbeck at UC Berkeley and the *RPP1B* sequences by Ksenia Krasileva at UC Berkeley. Primers used during PCR were designed to introduce a 9X histidine tag at the N-terminus of each construct. This would be used for the purposes of protein purification with nickel affinity chromatography and immunoblot analysis. The 5' primers used were also designed to encode a yeast translation initiation consensus sequence around the initiation ATG codon (AAAATGG; Table 2-1; Invitrogen 'EasySelect *Pichia* Expression Kit' manual). Each *RPP1B* and *RPS2* DNA construct was ligated into the pPICZa *P. pastoris* expression vector, transformed into *E. coli* and its integrity validated by sequencing. Plasmid DNA from one *E. coli* transformant of each construct was electroporated into *P. pastoris* X33.

Small-scale, 10 mL test expression cultures (see 2.7.4) were performed with six putative *P. pastoris* transformants of each construct taken from 100 µg/mL zeocin selection plates. RPS2 and RPP1B expression was examined with 10% SDS-PAGE and anti-His immunoblots of total protein extracted from lysed *P. pastoris* cultures. All conditions used during these procedures were identical to those used for M and L6 proteins (Schmidt *et al.*, 2007; Williams *et al.*, 2011). The procedures from *P. pastoris* transformation to the analysis of test expression were repeated two separate times for all RPP1B and RPS2 constructs. A colony known to heterologously express the M protein was included as a positive control in each test expression.

Despite the presence of M protein used as a positive control in the immunoblot, full-length RPS2 and near full-length RPS2 (RPS2(-25)) proteins were not expressed to a detectable level as represented in Figure 3-2A by the absence of any immunoreactive bands of the expected sizes (Table 3-1). Similar results were found for RPP1B, but shown for analysis of colonies selected from 6X zeocin selection plates in Figure 3-2B.

Expression levels of the NBS-LRR portion of M (MΔTIR) are higher than the near full-length M protein; higher protein purity of MΔTIR is also consistently achieved after nickel affinity purification (Williams, 2009). It was thought that if it was not possible to produce full-length versions of RPS2 and RPP1B, removing the N-terminal TIR or

CC domain of both might improve expression. However, unfortunately expression of neither of these truncated proteins was detected (Figure 3-2).

The hydrophobic N-terminal region of RPS2 was also removed to emulate the truncations of M and L6 that were required for the production of soluble protein (Schmidt *et al.*, 2007). Unfortunately this did not result in soluble expression of the RPS2(-25) protein (Figure 3-2A). Improved secretion of the synovial sarcoma X break point 2 (SSX2) protein from *P. pastoris* has previously been observed after the removal of a NLS (Huang *et al.*, 2010). Even though RPP1B proteins were being expressed intracellularly, the NLS was removed to see whether it would aid its intracellular expression; it did not (Figure 3-2B).

### 3.2.1 Investigating low or no protein expression

When low or no expression is observed from *P. pastoris* transformants it is recommended that the selection of transformed clones is performed with a significantly higher concentration of Zeocin than is normally recommended. This procedure allows for the selection of clones containing multiple integrated copies of the expression sequence (Higgins & Cregg, 1998; Invitrogen 'EasySelect *Pichia* Expression Kit' manual). Such a procedure has resulted in an increased yield of a number of recombinant proteins (Aw and Polizzi, 2013; Clare *et al.*, 1991a; Clare *et al.*, 1991b; Shen *et al.*, 2012). Increased expression of M has previously been observed using 6X (600 µg/mL) Zeocin selection, as compared to 1X selection (Williams, 2009).

Despite many colonies growing on YPD selection plates containing 600 µg/mL zeocin (6X zeocin), none of the RPP1B or RPS2 clones grown in 10 mL test expression cultures had any detectable levels of protein expression. Transformation with 6X zeocin selection was performed twice for all expression constructs. Each time test expression with 6X zeocin resistant clones did not show expression of any of the desired proteins, a *P. pastoris* clone containing the *M* gene used as a positive control, did express M protein (Figure 3-2B).

In case RPP1B or RPS2 proteins were being expressed by any of the clones, but the resultant protein levels were below the level of immunoblot detection, batch nickel

affinity chromatography was performed to enrich any expressed His-tagged RPS2 proteins. Total protein extracted from glass bead lysed *P. pastoris* test expression cultures were incubated with 1 mL of nickel affinity resin. The resin was briefly washed with wash buffer (20 mM Tris-base pH 8.0, 0.25 mM TX-100, 130 mM NaCl, 55 mM imidazole, 5 mM β-mercaptoethanol) before any bound proteins were eluted with elution buffer (20 mM Tris-base pH 8.0, 0.25 mM TX-100, 130 mM NaCl, 250 mM imidazole, 5 mM β-mercaptoethanol). SDS-PAGE and a subsequent anti-His immunoblot were performed with the eluted proteins (Figure 3-3). Whilst some immunoreactive bands were observed, they were not the sizes expected from expression of any of the proteins (Table 3-1). M protein expressed in a previous test expression was used as a positive control for the binding of histidine tagged protein to the affinity resin. M protein did bind and could be detected in this experiment (lane 11, Figure 3-3). Thus, disappointingly, this procedure did not result in the detection of any of the RPS2 proteins.

To determine whether the *RPS2* and *RPP1B* sequences had actually been integrated into the *P. pastoris* genome of zeocin resistant expression clones, PCR screening was performed. Such a procedure has never been required for any of the M and L6 proteins that have been expressed by *P. pastoris* in this laboratory. Standard laboratory procedure is to take six to eight zeocin resistant colonies through test expression and 90-100% of those colonies will express the protein of interest. However, as no RPP1B or RPS2 proteins were detected in the above experiments, screening for integration of the *RPP1B* and *RPS2* genes by PCR seemed necessary.

The PCR screen was performed on genomic DNA isolated from zeocin resistance *P. pastoris* colonies of each RPS2 construct with primers specific for regions of the *AOX1* promoter and transcription terminator. Since primers for *AOX1* were used, a product of 2.2 kbp was expected from those colonies that had integrated the expression constructs into their genome, as this band represents the native *AOX1* gene of *P. pastoris* X33 (Invitrogen 'EasySelect *Pichia* Expression Kit' manual). Of the clones obtained by 6X zeocin selection 11 were found to contain an *RPS2* construct of the expected size; RPS2-4, RPS2(-25)-4, CC-NB-1,-2,-3 and -4, CC-NB(-25)-1,-2,-3 and -4,

NB-LRR-1 (Figure 3-4). From colonies selected with 1X zeocin only one clone, NB-LRR-4 was amplified (data not shown).

Large-scale, 50 mL test expressions were subsequently performed with these clones that were found to have integrated an *RPS2* construct into their genome. An anti-His immunoblot was performed after SDS-PAGE to separate 30 µg of total protein extracted from the lysis of each clone. As in previous small-scale test expressions and despite M protein being detected, no bands of the expected sizes for the RPS2 proteins were observed (Figure 3-5).

Similar PCR analysis of RPP1B 6X zeocin clones showed that only one clone had positive integration of an RPP1B construct; RPP1B(-60) clone #4 (data not shown). Like the RPS2 clones, no RPP1B protein was detected after large-scale test expression of the RPP1B(-60)-4 clone (data not shown).

### 3.2.2 Trialling expression in the baculovirus/insect cell system

Concurrent with the above experiments, an expression trial of the full-length *RPS2* construct was conducted at the Protein Expression Facility at the University of Queensland. In this trial, commissioned by Dr. Simon Williams at the University of Queensland, the full-length *RPS2* gene was cloned into the pOPIN NHis vector (6X His N-terminal tag) and expression was attempted in a baculovirus/Sf9 insect cell system. While expression of the NHis-RPS2 construct was detected, the protein was largely insoluble (Figure 3-6).

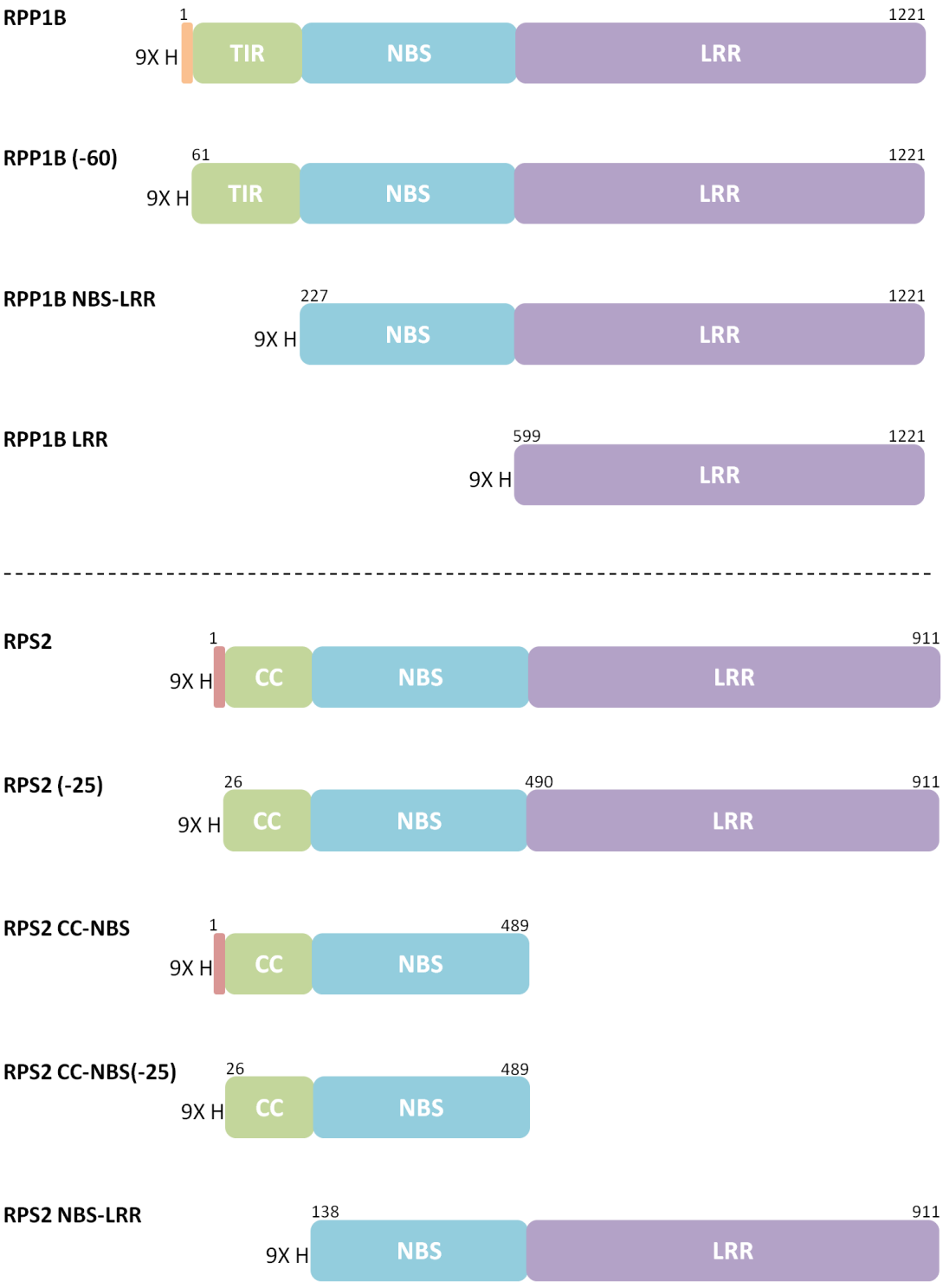



**Figure 3-1: Schematic representation of the RPP1B and RPS2 protein constructs**

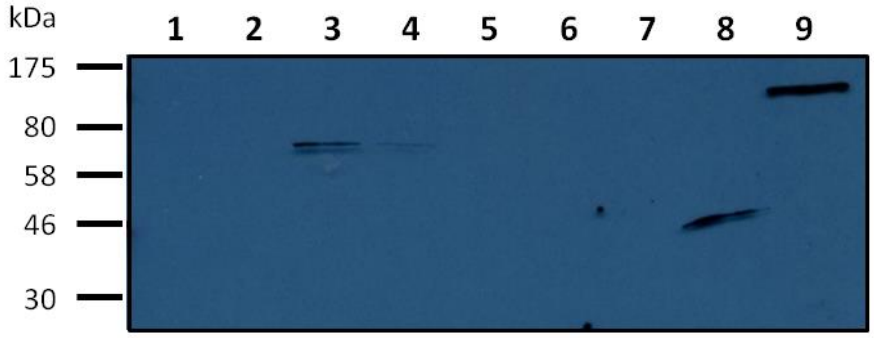
The red section in some of the RPS2 constructs represents the region containing the hydrophobic amino acids (residues 7 to 22). The orange section in the RPP1B construct represents the hydrophilic extension, which also contains a putative NLS. All protein constructs contained an N-terminal 9X histidine (9X H) tag to facilitate their purification by nickel affinity chromatography.

**Table 3-1: Predicted molecular weights of the RPP1B and RPS2 protein constructs**

Expected sizes were calculated with the 'Compute pI/Mw Tool' from the ExPASy server and included the N-terminal 9X histidine tag and a linker glycine residue.

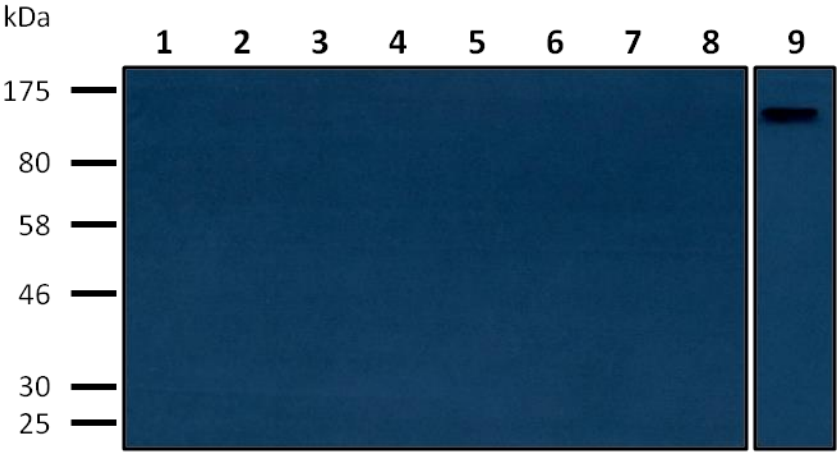
| Designation | Amino Acids | Expected Size (kDa) |
|---|---|---|
| **RPP1B** | 1-1221 | 140 |
| **RPP1B(-60)** | 61-1221 | 133 |
| **NBS-LRR** | 227-1221 | 115 |
| **LRR** | 599-1221 | 72 |
| **RPS2** | 1-911 | 106 |
| **RPS2(-25)** | 26-911 | 103 |
| **CC-NBS** | 1-489 | 57 |
| **CC-NBS(-25)** | 26-489 | 54 |
| **NBS-LRR** | 138-911 | 90 |

**A.**




**1X RPS2 test expression:** Protein ladder, NEB Prestained broad range; lane 1, NB-LRR-1; lane 2, NB-LRR-2; lane 3, NB-LRR-3; lane 4, NB-LRR-4; lane 5, RPS2-1; lane 6, RPS2(-25)-1; lane 7, CC-NB-1; lane 8, CC-NB(-25)-1; lane 9, M test expression positive control

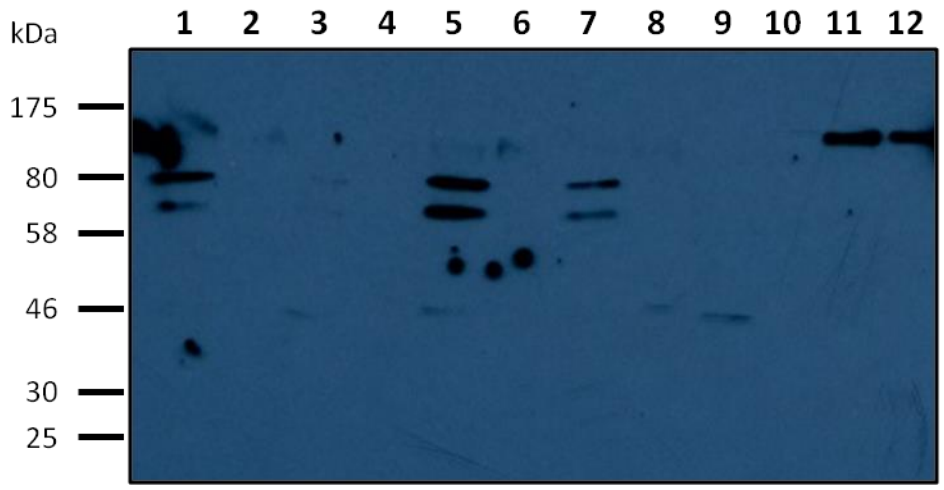
**B.**




**6X RPP1B test expression:** Protein ladder, NEB Prestained broad range; lane 1, RPP1B-1; lane 2, RPP1B-2; lane 3, RPP1B(-60)-1; lane 4, RPP1B(-60)-2; lane 5, NB-LRR-1; lane 6, NB-LRR-2; lane 7, LRR-1; lane 8, LRR-2; lane 9, M test expression positive control

**Figure 3-2: None of the RPS2 or RPP1B proteins were expressed by *P. pastoris***

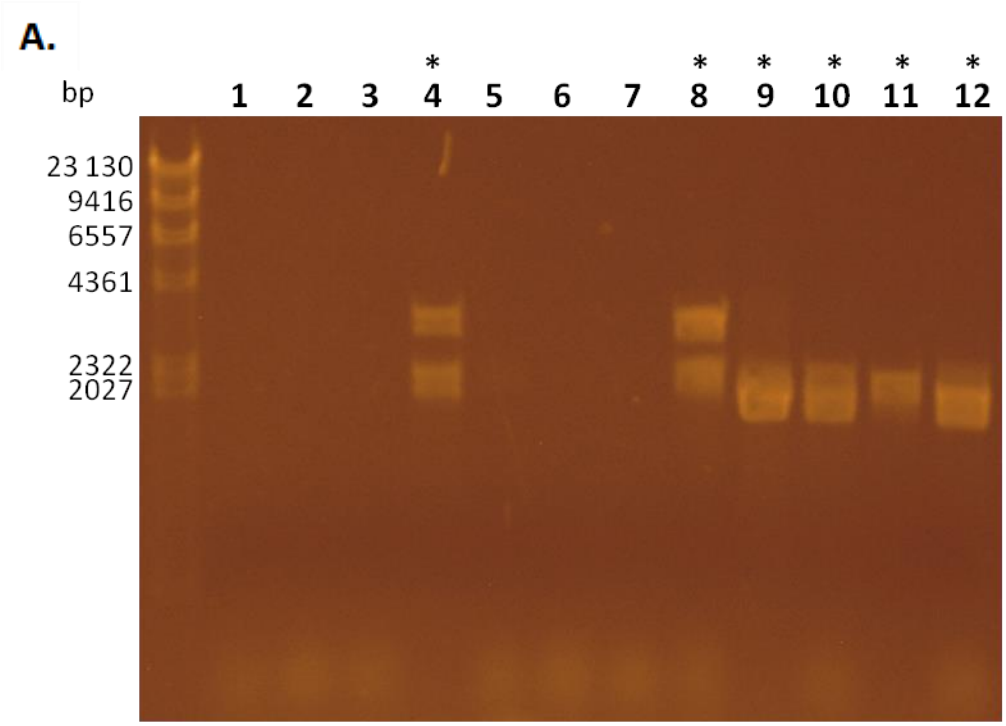
Transformed *P. pastoris* colonies were selected on YPD plates with either 100 µg/µL (1X) or 600 µg/µL (6X) of zeocin. Eight colonies for each construct were put through the test expression procedure and at the conclusion of protein induction the cultures were lysed with glass beads and the crude lysate separated by SDS-PAGE. Total protein concentration was estimated using a total protein assay and 30 µg of extract from each colony was loaded into every lane. The absence of anti-His immunoreactive bands of the expected molecular weights showed that none of the RPS2 or RPP1B proteins was expressed successfully.



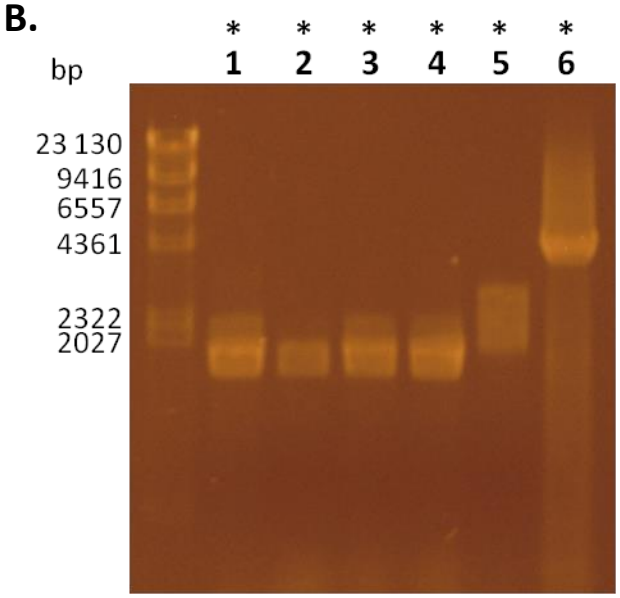

Protein ladder, NEB Prestained broad range; lane 1, RPS2-1 bound; lane 2, RPS2-1 unbound; lane 3, RPS2(-25)-1 bound; lane 4, RPS2(-25)-1 unbound; lane 5, CC-NB-1 bound; lane 6, CC-NB-1 unbound; lane 7, CC-NB(-25)-1 bound; lane 8, CC-NB(-25)-1 unbound; lane 9, NB-LRR-1 bound; lane 10, NB-LRR-1 unbound; positive control M protein from test expression, lane 11, bound; lane 12 M unbound

**Figure 3-3: Enrichment using nickel affinity chromatography did not result in the detection of any histidine-tagged RPS2 proteins**

*P. pastoris* cultures subjected to test expression were lysed with glass beads and the crude lysate was incubated with 1 mL of nickel affinity resin. The resin was briefly washed with wash buffer (20 mM Tris-base pH 8.0, 0.25 mM TX-100, 130 mM NaCl, 55 mM imidazole, 5 mM β-mercaptoethanol) before any bound proteins were eluted with a buffer containing 20 mM Tris-base pH 8.0, 0.25 mM TX-100, 130 mM NaCl, 250 mM imidazole, 5 mM β-mercaptoethanol. A sample of the eluted protein was separated using 10% SDS-PAGE and an anti-His immunoblot was performed. Whilst M protein from a test expression culture was detected in the bound and unbound fractions no immunoreactive bands corresponding to the expected sizes of any of the RPS2 proteins were observed.



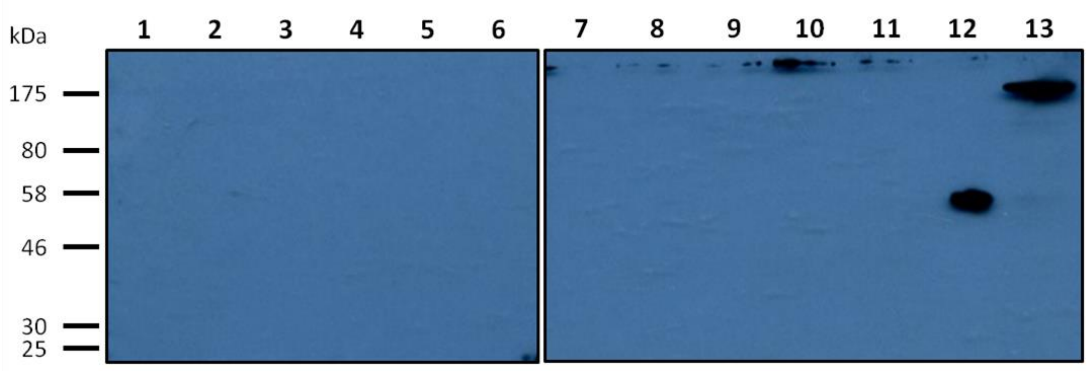

DNA ladder, NEB λ DNA-*HIND*III digest; lane 1, RPS2-1; lane 2, RPS2-2; lane 3, RPS2-3; lane 4, RPS2-4; lane 5, RPS2(-25)-1; lane 6, RPS2(-25)-2; lane 7, RPS2(-25)-3; lane 8, RPS2(-25)-4; lane 9, CC-NB-1; lane 10, CC-NB-2; lane 11, CC-NB-3; lane 12 CC-NB-4




DNA ladder, NEB λ DNA-*HIND*III digest; lane 1, CC-NB(-25)-1; lane 2, CC-NB(-25)-2; lane 3, CC-NB(-25)-3; lane 4, CC-NB(-25)-4; lane 5, NB-LRR-1; lane 6, pPICZ M plasmid control

**Figure 3-4: Not all colonies that grew on 6X zeocin selection plates contained the *RPS2* constructs**

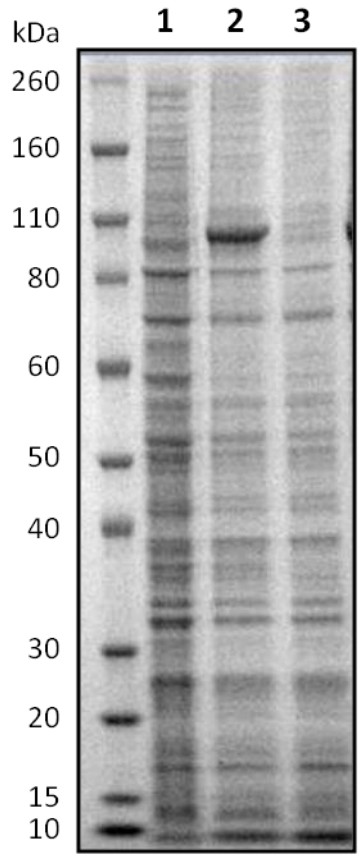
PCR products amplified from genomic DNA isolated from *P. pastoris* colonies that grew on YPD plates with 600 µg/mL zeocin with AOX1 primers were run on 1% agarose gels at 80 V for 40 minutes. The gels were stained with ethidium bromide to visualise the products. Bands observed between the 2027 and 2322 bp markers are around the expected size (~2.2 kbp) of the *AOX1* gene. The expected sizes of the RPS2 constructs were ~2700 bp for *RPS2* and *RPS2(-25)*, ~1500 bp for *CC-NB* and *CC-NB(-25)* and ~2300 bp for *NB-LRR*. The colonies found to contain an expression construct are identified with an *. These were the RPS2-4, RPS2(-25)-4, CC-NB-1,-2,-3 and -4, CC-NB(-25)-1,-2,-3 and -4 and NB-LRR-1 clones.




Protein ladder, NEB Prestained broad range; lane 1, RPS2-4; lane 2, RPS2(-25)-4; lane 3, CC-NB-1; lane 4, CC-NB-2; lane 5, CC-NB-3; lane 6, CC-NB-4; lane 7, CC-NB(-25)-1; lane 8, CC-NB(-25)-2; lane 9, CC-NB(-25)-3; lane 10, CC-NB(-25)-4; lane 11, NB-LRR-1; lane 12, NB-LRR-2; lane 13, M test expression positive control

**Figure 3-5: Large-scale test expression of RPS2 *P. pastoris* colonies with confirmed construct integration**

*P. pastoris* clones that contained *RPS2* expression constructs, as determined by PCR, were grown in 50 mL cultures for a large-scale test expression. Post-induction a sample of each culture was lysed and 30 µg of total protein was analysed by SDS-PAGE and an anti-His immunoblot. The absence of anti-His immunoreactive bands at the expected sizes confirmed that not one of the RPS2 proteins was successfully expressed.




Protein ladder, Novex® Sharp Pre-stained ladder; lane 1, Sf9 negative control; lane 2, pOPIN RPS2 total protein after lysis; lane 3, pOPIN NHisRPS2 soluble protein after lysis.

**Figure 3-6: Full-length RPS2 protein was expressed in the baculovirus/insect cell system but was not soluble**

The Protein Expression Facility at the University of Queensland carried out a solubility test with Sf9 cells that had been transformed with a pOPIN vector encoding full-length RPS2 with an N-terminal 6X histidine tag. The expected size of this protein was ~106 kDa. SDS-PAGE was performed with fractions of insoluble (total) and soluble protein obtained after lysis of the transfected Sf9 cells. In the gel stained for total protein a band of the expected size (106 kDa) can be seen in the insoluble or total protein fraction (lane 2), but no corresponding band is present in the soluble fraction (lane 3).

## 3.3 Discussion

Unfortunately no immuno-detectable RPP1B or RPS2 proteins were expressed by *P. pastoris* using the conditions that consistently produce soluble and biochemically active M and L6 proteins. Efforts to produce and detect recombinant protein were not positively influenced by domain or regions truncations, increasing gene dosage in the recombinant *P. pastoris* clones, or by enriching His-tagged proteins with low expression rates via a crude purification strategy. Whilst reoccurring bands around 47, 75 and 83 kDa were detected in many of the immunoblots conducted after test expressions, they were not the predicted size of any of the RPS2 or RPP1B proteins or truncation of thereof. As they were consistently detected in multiple protein extracts of all constructs tested they most likely represent non-specific anti-His immunoreactive proteins present in *P. pastoris*.

Expression of NB-LRRs other than M and L6 in *P. pastoris* has been attempted by other research groups and has also been unsuccessful. In Frank Takken's laboratory, expression of full-length and truncated variants of the CC-NB-LRR protein Rx was trialled. M was used as a positive control for the system and whilst it was expressed, none of the Rx proteins were (Lukasik-Shreepaathy, 2011).

Removing the hydrophobic N-terminal regions that contain predicted signal anchors was necessary to achieve soluble M and L6 protein in *P. pastoris* (Schmidt *et al.*, 2007). In flax, L6 has been localised to the Golgi body and it has a predicted signal anchor and a transmembrane helical segment from residues seven to 24 which direct and enable this localisation (Takemoto *et al.*, 2012). Even though it is known that M is located at the tonoplast membrane, its N-terminal hydrophobic region does not contain a software predicted signal anchor or other membrane attachment sequences (Takemoto *et al.*, 2012).

Such trafficking sequences are not found at the N-terminus of RPS2, although it does also contain a hydrophobic region. Despite not containing a known signal anchor or transmembrane helical region, RPS2 has been found to behave like an integral membrane protein and experimentally to have membrane attachment with an incomplete plasma membrane localisation (Axtell & Staskawicz, 2003; Takemoto *et*

*al.*, 2012). Putative palmitoylation motifs have been identified within this hydrophobic region of RPS2 and may contribute to its subcellular membrane localisation (Takemoto *et al.*, 2012). Contrary to M and L6, truncation of this region of RPS2 did not result in the production of any soluble protein in *P. pastoris*. However, as insoluble full-length RPS2 protein was expressed in the baculovirus/insect cell system the other expression constructs, particularly the truncation of the N-terminal hydrophobic region should also be tested in this system. As insect cells are capable of modifying proteins by adding palmitic acid to palmitoylation motifs it is possible that this mechanism of membrane attachment did contribute to the insoluble production of RPS2 (Resh, 1999).

Although RPP1B does not contain a predicted N-terminal signal peptide, or an anchor sequence, it does have three predicted N-terminal myristoylation and palmitoylation motifs (Takemoto *et al.*, 2012). As RPP1B has been localised to the plasma membrane it is possible that these sites of possible myrisitc and palmitic acid attachment are providing an alternative means of membrane anchorage (Takemoto *et al.*, 2012). Even if RPP1B was directed to the plasma membrane of *P. pastoris* removal of these motifs with the truncation of the TIR domain (NBS-LRR construct) did not results in the production of any detectable protein in *P. pastoris*. Given more time and funds an investigation of RPP1B protein expression in the baculovirus/insect cell system should also be undertaken.

Within *A. thaliana*, in the absence of its cognate effector, RPS2 is found to associate with RIN4 at the plasma membrane (Axtell & Staskawicz, 2003; Mackey *et al.*, 2003). RIN4 contains a carboxy-terminal palmitoylation motif that directs this plasma membrane attachment. This placement allows RIN4 to mediate the interaction between RPS2 and its effector AvrRpt2 (Day *et al.*, 2005; Kim *et al.*, 2005; Takemoto & Jones, 2005). This interaction likely holds RPS2 in a conformation that acts as a form of negative regulation (Axtell & Staskawicz, 2003; Mackey *et al.*, 2003) and may also help to maintain its stability. It is therefore highly likely that RPS2 will require such an association for it to be stably produced in any heterologous system. However, localisation at a membrane makes heterologous expression more difficult. Unfortunately, whilst mutation of one of the cysteine residues believed to be

responsible for the membrane localisation of RIN4 disrupts this localisation, it also disrupts the RIN4-RPS2 interaction (Day *et al.*, 2005). Changing the subcellular localisation of both proteins whilst maintaining their interaction may not be possible and this may hamper attempts to improve solubility through co-expression with RIN4. Nevertheless co-expression should be attempted (perhaps through alternative mutation of the membrane localisation region) in both the *P. pastoris* and baculovirus/insect cell systems, as purification of membrane bound proteins is possible.

### 3.3.1 Further optimisation of expression in *P. pastoris*

That RPP1B and RPS2 clones which did amplify the appropriate size product during PCR screening did not give expression begs the question why? The most likely explanation for lack of RPS2 expression is insolubility due to the absence of its regulatory partner RIN4. However, why RPP1B failed to yield detectable levels of protein remains a mystery. Some success in improving expression levels in *P. pastoris* has been found by optimising the codons within the gene of interest to those preferred by *P. pastoris* (Hu *et al.*, 2013a, Li *et al.*, 2012). Whether this improvement in expression is an indirect consequence of a favourable decrease in the A+T content, rather than a direct result of codon optimisation remains controversial (Boettner *et al.*, 2007). Nevertheless, altering the target gene sequence to more favourably resemble native *P. pastoris* genes has increased the expression of a number of proteins, and it will be interesting to see whether expression of any of the RPP1B or RPS2 proteins could be gained by codon optimisation.

AT-rich regions have been known to resemble polyadenylation consensus sites and therefore act as premature polyadenylation sites, resulting in the premature termination of transcription. Increasing the G+C content of genes to 40-50% and changing any such consensus sequences have resulted in a favourable increase in the expression levels of a number of proteins (Gurkan & Ellar, 2003; Huang *et al.*, 2008; Li *et al.*, 2012; Scorer *et al.*, 1993). The *RPP1B* and *RPS2* genes do not have very different A-T contents than the *M* gene (RPP1B 60.2%, RPS2 55.1%, M 61.2%) and do not contain the sequence TTTTTATA, which resembles a sequence found in HIV-1 gp120, ATTATTTTAT AAA that is known to prematurely terminate mRNA when

expressed in *P. pastoris* (Scorer *et al.*, 1993). Whilst it is therefore less likely that premature transcription termination was the reason that no expression was observed in *P. pastoris*, it is still possible that these genes do contain other sequences that are causing the premature termination of transcription. Yeast cryptic splicing, processing and polyadenylation sequences are not highly conserved and therefore many sequences can fulfil the requirements of such a signal and prematurely terminate or alter a transcript (Graber *et al.*, 1999). Real-time PCR and/or Northern blot analysis could be undertaken to determine if a premature termination event has taken place or cryptic processing is occurring.

Heterologous gene instability has been reported in *P. pastoris* and has been predominantly observed with multi-copy clones after the induction of expression with methanol (Aw & Polizzi, 2013). PCR screening was not conducted with the putative multi-copy clones after methanol induction and it is thus plausible that integrated heterologous genes had become unstable during the test expression procedure, thus resulting in a lack of observable protein expression.

### 3.3.2 Further optimisation of expression in the baculovirus/insect cell system

Expression in eukaryote expression systems appears to be the best method for the production of functional R proteins. Mla27 was successfully expressed in insect cells (Maekawa *et al.*, 2011) and M and L6 in *P. pastoris* (Schmidt *et al.*, 2007). Whilst immuno-detectable expression of RPP1B and RPS2 in *P. pastoris* was not possible, insoluble full-length RPS2 protein was detected in the baculovirus/Sf9 insect cell system. Given that heterologous RPS2 expression was obtained in this system, further modification of the expression constructs could be introduced in order to aid solubility.

Fusion of genes that encode stable proteins to expression constructs with poor solubility has been employed in a number of cases and has resulted in improved solubility (Young *et al.*, 2012). Such fusion proteins could include the *E. coli* maltose binding protein and thioredoxin A proteins (Young *et al.*, 2012). An engineered small ubiquitin-related modifier (SUMO) protein fusion tag has also been used to

successfully improved protein expression in insect cells (Liu *et al.*, 2008). The use of such fusion proteins should be investigated for future RPS2 and/or RPP1B production.

Chaperone proteins have also been successfully utilised to stabilise proteins that are prone to folding poorly and thus become insoluble. Protein disulphide-isomerase is one such chaperone that has been co-expressed in the *P. pastoris* system to increase yields of recombinant proteins (Aw & Polizzi, 2013). Co-expression is usually employed when a protein is being secreted to alleviate stress on the secretory pathway that, when overwhelmed, can trigger protein degradation (Aw & Polizzi, 2013). However, there is no reason why co-expression with a chaperone should not be attempted with the intracellular expression of RPS2 or RPP1B in either heterologous system.

### 3.3.3 Conclusions

Although it was not possible to detect expression of RPP1B or RPS2 proteins in this study further optimisation may yet see soluble expression. Whilst the reasons for the absence of detectable protein maybe different for the two R proteins, further exploration of RPS2 expression in the baculovirus/insect cell system is warranted. Given more time and funding expression of RPP1B in the baculovirus/insect cell system would also have been attempted and this should be carried out in the future.

Further understanding of the activation and regulation mechanisms of NB-LRR proteins will only be possible if more are studied *in vitro*; however this study, and others before it, have proven that this is not an easy accomplishment. The only way to get around this is to trial more R proteins and more constructs in more heterologous expression systems.

Given the limited time available in my candidature I decided to leave this expression study at this point and proceed with the work that will be presented in the subsequent results chapters of this thesis.

# 4 The Role of the Walker B Motif in NB-LRR Protein Function

## 4.1 Introduction

The Walker B, or Kinase-2, motif is a highly conserved amino acid signature that is a common characteristic of most P-loop NTPases, including STAND proteins. This region is typically found at the C-terminus of a hydrophobic β-strand that is positioned to allow it access to a bound nucleotide. In plant R proteins, the Walker B is located in the predicted NB sub-domain of the NB-ARC region (Albrecht & Takken, 2006; van Ooijen *et al.*, 2008). Conserved acidic aspartate (D) and/or glutamate (E) residues within this motif play roles in nucleotide coordination and NTP hydrolysis (Albrecht *et al.*, 2003; Leipe *et al.*, 2004; Story & Steitz, 1992; Tameling *et al.*, 2006). In the majority of STANDs the Walker B motif has the consensus sequence hhhhD-D/E (where h is a hydrophobic residue); however, there can be considerable variability in this sequence. The only invariant features are a negatively charged residue following a stretch of hydrophobic amino acids (Leipe *et al.*, 2004; Koonin, 1993).

### 4.1.1 The canonical Walker B motif

In the classic Walker B (hhhhD-D/E), the proximal aspartate or glutamate coordinates the active site magnesium cation ($Mg^{2+}$) that is linked to the β- and γ-phosphates of the bound NTP molecule. It consequently acts to chelate the bound nucleotide within a protein's nucleotide binding pocket. Mutation of this residue in NTPases can affect nucleotide binding and generally has a negative effect on nucleotide hydrolysis rates (Marquenet & Richet, 2007; Tameling *et al.*, 2006; Traut, 1994).

The second highly conserved acidic residue provides the catalytic carboxylate side chain for the hydrolysis of a nucleoside triphosphate (NTP) (Albrecht *et al.*, 2003). In many proteins this residue participates in the hydrolytic event via a process called general base catalysis. In these reactions, the basic residue abstracts a proton from the hydrolytic water molecule, thereby activating it for nucleophilic attack of the β-γ phosphoanhydride bond (Figure 4-1). This is the critical rate limiting step of many STAND-catalysed ATP hydrolysis events (Leipe *et al.*, 2004). Consequently, mutation of this catalytic base reduces or eliminates ATP hydrolysis, but generally does not affect ATP binding in NLR proteins (Marquenet & Richet, 2007; Tameling *et al.*, 2006).

### 4.1.2 Non-canonical Walker B motifs

There are varying arrangements of conserved functional amino acids within Walker B regions. Proteins belonging to the NACHT sub-class of STAND proteins are identified by the presence of a small residue (glycine, alanine or serine) directly downstream of the $Mg^{2+}$ coordinating residue, in the place of the archetypal catalytic residue. However, the majority of NACHT proteins have two additional acidic residues three positions downstream of the first aspartate; hhhhD[GAS]hDE (Leipe *et al.*, 2004). This unconventional amino acid arrangement can, in some cases, functionally compensate for the missing up-stream catalytic residue (Ye *et al.*, 2007; Zurek *et al.*, 2011b). This non-canonical Walker B region has been re-defined as the extended Walker B box (Proell *et al.*, 2008).

In the STAND AP-ATPase sub-class, which includes Apaf-1, CED4 and R proteins, the Walker B motif consensus includes only two conserved acidic residues. This is true for most plant CC-NB-LRRs (consensus LLVLDDVW); however, similar to many NACHTs, some NB-LRRs have an extended Walker B region (Figure 4-2). All TIR-NB-LRR proteins contain an additional acidic amino acid downstream of the canonical acidic pair (consensus LLVLDDVD) (Meyers *et al.*, 1999). Furthermore, some have an additional conserved glutamate four residues downstream of the first aspartate (LLVLDDVDE). Flax rust proteins from the M, L and N loci contain a Walker B motif with this additional acidic pair.

### 4.1.3 Biochemical roles of Walker B residues in CC-NB-LRRs

The acidic residues within the Walker B motif of the CC-NB-LRR protein I-2 perform the biochemical roles of classic Walker B residues (Tameling *et al.*, 2006). Substituting the proposed catalytic aspartate with a glutamate, I-$2^{D283E}$, created an autoactive protein that displayed reduced ATPase activity in comparison to wild-type I-2. It was proposed that this was not due to a decreased affinity for ATP; instead that the mutation prevented the correct bond forming with the catalytic water molecule. This loss of autoinhibition suggests that ATP hydrolysis is not required for the signalling activity of I-2, but rather acts as a reset or off switch. An aspartate to glutamate mutation of the proposed catalytic residue has also been shown to cause autoactivity

in RPS5 (Ade *et al.*, 2007), as has an aspartate to alanine mutation in RPM1 (Eitas, 2010). Neither protein has been studied biochemically.

An $Mg^{2+}$ ion is required for I-2 to bind ADP and accordingly, mutation of the proposed magnesium coordinating aspartate of I-2 (I-$2^{D282C}$) results in inactivity (Tameling *et al.*, 2002; Tameling *et al.*, 2006). Studies on Rx and RPS2 have also highlighted the functional importance of the archetypal pair of conserved aspartates in CC-NBS-LRRs. In both proteins, dual mutation of their Walker B aspartate pair (Rx$^{D244A+D245A}$ and RPS2$^{D262T+D263A}$) generated inactive proteins (Bendahmane *et al.*, 2002; Tao *et al.*, 2000). Further research is required for it to be determined whether or not the biochemical functions of each residue are conserved between all CC-NBS-LRR proteins.

### 4.1.4 Walker B residues in TIR-NB-LRRs

Consistent with TIR-NB-LRR proteins containing non-canonical Walker B regions, different phenotypes have been generated by mutations of proposed catalytic residues analogous to those within CC-NB-LRR proteins. Previous work on the M protein has shown that mutation of the proposed catalytic aspartate to glutamate (M$^{D364E}$) does not affect HR activation, but does increase ADP occupancy (Williams *et al.*, 2011), although some question is placed on the phenotypic data of this mutant in the work presented in this chapter. In L6, a flax TIR-NB-LRR that is 86% similar to the M protein, the analogous mutation (L6$^{D350E}$) creates a non-functional protein (Peter Dodds, personal communication; also see 4.2.1). It is unknown what effect this mutation has on the nucleotide occupancies of L6. Given that the analogous mutations in CC-NB-LRRs generated autoactivity (Tameling *et al.*, 2006) these results indicate that the additional downstream acidic residues may play important functional roles in TIR-NB-LRRs. The work presented here attempts to address this possibility.

### 4.1.5 The roles of ATP hydrolysis and the Walker B motif in the NB-LRR activation model

Based on the reduced ATPase activity of the autoactive I-$2^{D283E}$ mutant, and further evidence of ATPase activity in related STAND proteins, ATP hydrolysis is currently

believed to act as the off switch for NB-LRR protein activation. Mutations of residues critical for ATP hydrolysis should result in autoactivity because they trap the proteins in an 'on' state. To achieve ATP hydrolysis, general base catalysis is performed by the second conserved Walker B acidic residue (Iyer *et al.*, 2004; Marquenet & Richet, 2007; Tameling *et al.*, 2006). The intrinsic ATPase activities of the M and I-2 proteins suggests that at least some NB-LRR proteins continuously cycle between ADP and ATP bound states. With limited biochemical evidence, the role of ATP hydrolysis in NB-LRR activation and activity remains controversial and needs to be better understood.

### 4.1.6 Aims

Given that functional differences have been observed after analogous Walker B mutations, it is unlikely that the functions performed by the individual acidic residues are conserved between NB-LRRs. To elucidate whether the acidic residues within its extended Walker B are functionally important, rational mutagenesis of the M protein was performed.

To determine whether or not the additional acidic residues, (LLVLD$^{363}$D$^{364}$VD$^{366}$E$^{367}$), can functionally compensate for the upstream aspartates, the first and second pairs of conserved acidic residues were individually and jointly substituted with alanine (M$^{D363A}$, M$^{D364A}$, M$^{D363A+D364A}$, M$^{D366A}$, M$^{E367A}$ & M$^{D366A+E367A}$). Additionally, to mimic the D to E mutation of the second residue in the archetypal acidic pair, a glutamate substitution of the non-canonical downstream aspartate (M$^{D366E}$) was also made.

The function of each mutant was tested *in planta* using *A. tumefaciens*-mediated transient expression. Mutants were expressed alone and co-expressed with AvrM and avrM. Those mutants with a phenotype different to that of wild-type M were heterologously expressed using the *P. pastoris* expression system developed by Schmidt *et al.* (2007), and purified according to the methods reported by Williams *et al.* (2011). The quantity of any ATP or ADP bound within the mutant proteins was determined with an *in vitro* bioluminescence-based ATP quantification assay. The addition of pyruvate kinase, to convert ADP into ATP, permitted ADP occupancies to be calculated (Williams *et al.*, 2011).

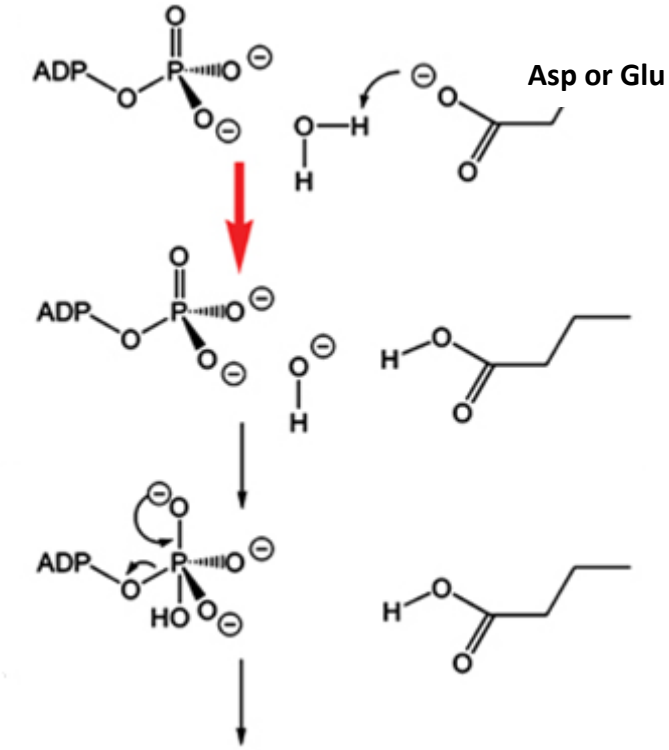



**Figure 4-1: ATP hydrolysis achieved by the mechanism of general base catalysis (Zaitseva *et al.*, 2005)**

In general base catalysis the hydrolytic water molecule is polarised when a basic residue in the Walker B motif abstracts a proton and becomes protonated. The activated water molecule can then perform nucleophilic attack on the β-γ phosphodiester bond of the ATP molecule bound within the nucleotide binding pocket. This is the rate limiting step of the ATP hydrolysis reaction that is catalysed by some STAND proteins (Marquenet & Richet, 2007; Zaitseva *et al.*, 2005).

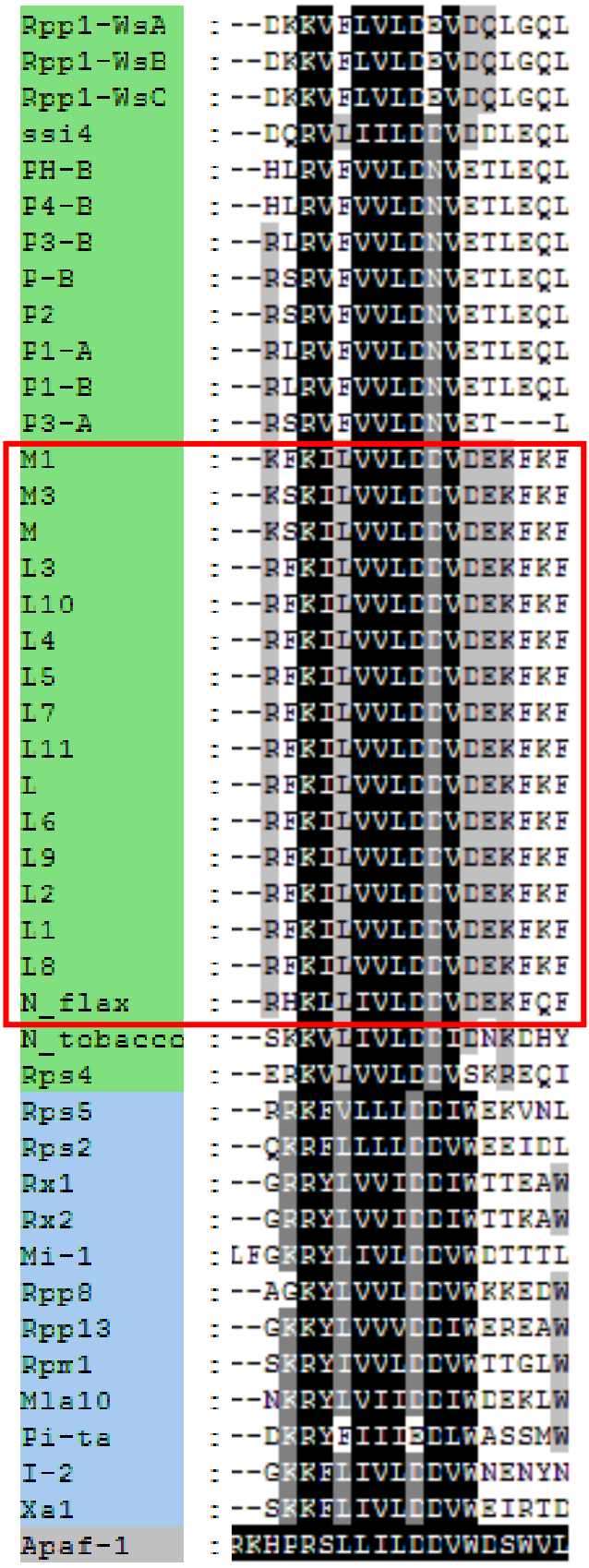

Rpp1-WsA
Rpp1-WsB
Rpp1-WsC
ssi4
PH-B
P4-B
P3-B
P-B
P2
P1-A
P1-B
P3-A
M1
M3
M
L3
L10
L4
L5
L7
L11
L
L6
L9
L2
L1
L8
N_flax
N_tobacco
Rps4
Rps5
Rps2
Rx1
Rx2
Mi-1
Rpp8
Rpp13
Rpm1
Mla10
Pi-ta
I-2
Xa1
Apaf-1

**Figure 4-2: Multiple sequence alignment of the Walker B motifs of 42 plant NB-LRR proteins and the human NOD-like receptor Apaf-1**

To highlight the differences in Walker B motif amino acid conservation within the two groups, the plant TIR- and CC-NB-LRR classes are highlighted in green and blue respectively. TIR-NB-LRR proteins all contain at least one additional acidic residue downstream of the archetypal Walker B acidic pair. Flax rust NB-LRRs, found within the red box, contain two additional downstream acidic residues. Amino acid conservation is highlighted as follows: 100% conservation: black background with white text; 80% conservation: dark grey background with white text; 60% conservation: light grey background with black text.

FASTA protein sequences were obtained from NCBI and the multiple sequence alignment was performed using CLUSTAL-W, BLOSUM and GeneDoc software.

## 4.2 Results

The Walker B motif is located within the NB sub-domain of the central nucleotide binding site or NB-ARC domain (Figure 4-3A). Resides within this motif are predicted to play roles in both nucleotide binding and ATP hydrolysis. In flax rust resistance proteins from the L, M and N loci this motif is made up of two acidic aspartate residues separated by a valine from two additional acidic residues, an aspartate and glutamate. A total of eight mutations have been made to these acidic residues in the TIR-NB-LRR flax M protein to determine if they play functionally important roles in the activation and function of M. These mutants are: $M^{D363A}$, $M^{D364A}$, $M^{D364E}$, $M^{D363A+D364A}$, $M^{D366A}$, $M^{D366E}$, $M^{E367A}$ and $M^{D366A+E367A}$. Their positions relative to the perceived structure of the NB pocket of M based on the structure of the human apoptosis promoting protein Apaf-1 are shown in Figure 4-3B.

Wild-type M and the $M^{D364E}$ mutant have previously been reported by Williams *et al.* (2011). For this reason the nucleotide quantification and immunoblot analysis of their *in planta* transient expression were not repeated in this study. Wherever those results appear in this thesis they are the work of Simon J. Williams and Pradeep Sornaraj, as published in Williams *et al.* (2011) and reported in Williams (2009) and Sornaraj (2013).

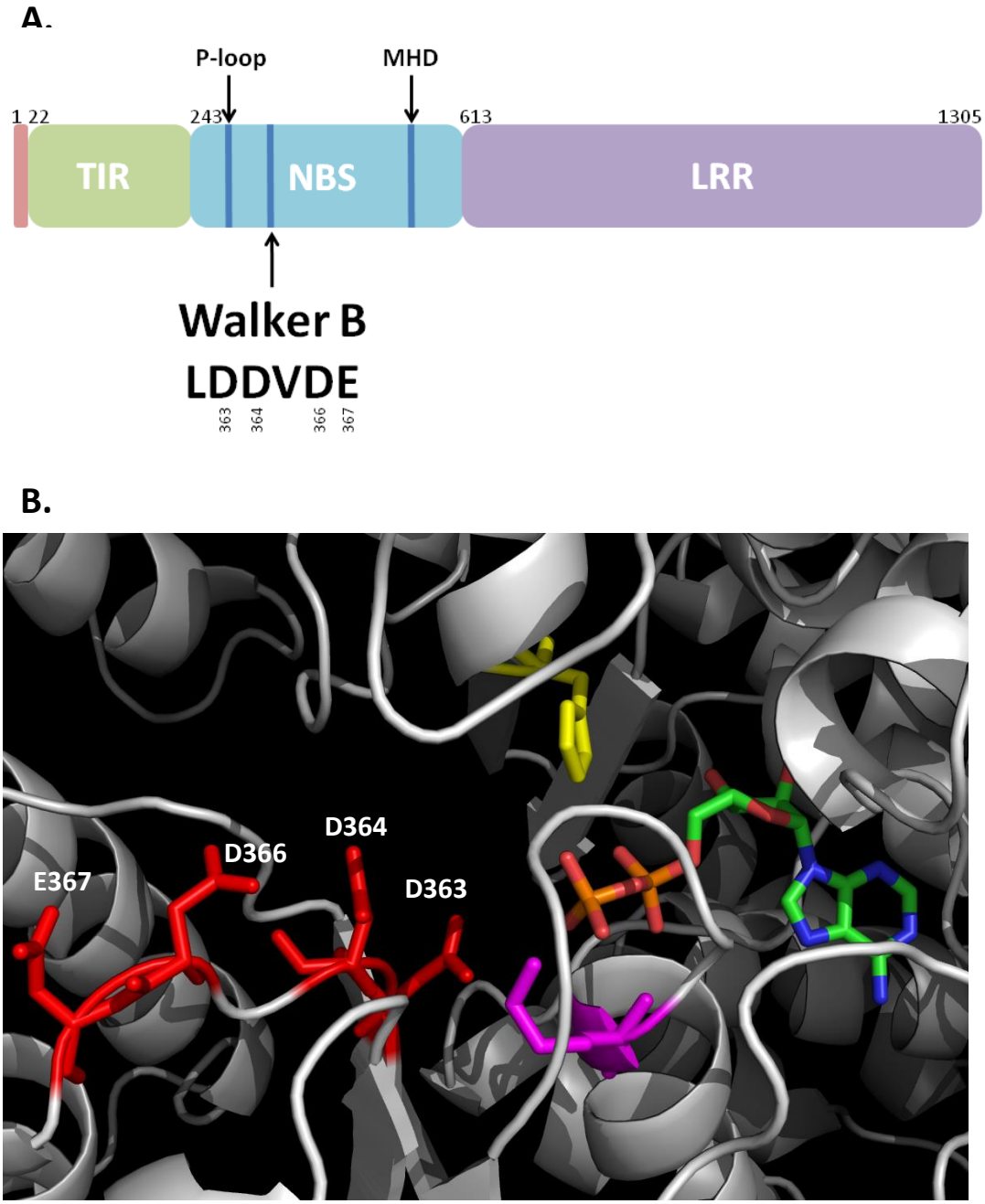



**Figure 4-3: The Walker B motif of the flax rust TIR-NB-LRR M protein**

**A)** The Walker B motif within M protein. **B)** A model of M's Walker B amino acids superimposed on the crystal structure of human Apaf-1 bound to ADP (PDB ID# 1z6t). The side chains of the four conserved acidic residues of the Walker B are highlighted in red. From the Apaf-1 structure, the invariant lysine of the P-loop is shown in magenta and the conserved histidine of the LHD motif in yellow. Although this model gives may have some broader architectural value, as Apaf-1 does not contain the D366 and E367 residues, this model is meant only as a guide as to where they may sit in relation to a bound nucleotide.

**Table 4-1: Predicted roles and mutations of the conserved Walker B acidic residues**

| Residue | Proposed Function | Single Mutations | Double Mutations |
|---|---|---|---|
| **D363** | $Mg^{2+}$ coordination | D363A | D363A+D364A |
| **D364** | Catalytic base | D364A<br>D364E | |
| **D366** | Unknown | D366A<br>D366E | D366A+E367A |
| **E367** | Unknown | E367A | |

### 4.2.1 Equivalent mutations of the proposed catalytic residues of the flax TIR-NB-LRRs M and L6 have different functional consequences

As reported by Williams *et al.* (2011) an aspartate to glutamate substitution of the proposed catalytic residue, D364, does not autoactivate or inactive the M protein. However, the infiltration timelines presented here show that this $M^{D364E}$ mutation does delay the onset of the HR and reduces its overall intensity (Figure 4-4).

Co-expression of the analogous mutation in genomic L6, $L6^{D350E}$, with the cognate effector AvrL (both driven by *Cauliflower mosaic virus* 35S promoter (CaMV 35S) in pTNSpec) did not develop a HR over the timeframe in which the co-expression of wild-type L6 and AvrL does (Bernoux *et al.*, 2011b; Howles *et al.*, 2005). This result corroborated personal communications from Peter Dodds in which $L6^{D350E}$ was found to be an inactive mutant. The anti-HA immunoblots of the single and co-infiltrations showed that the $L6^{D350E}$ mutant was expressed in the infiltrated cotyledons (Figure 4-4).

### 4.2.2 Overexpression of M protein generates spontaneous cell death

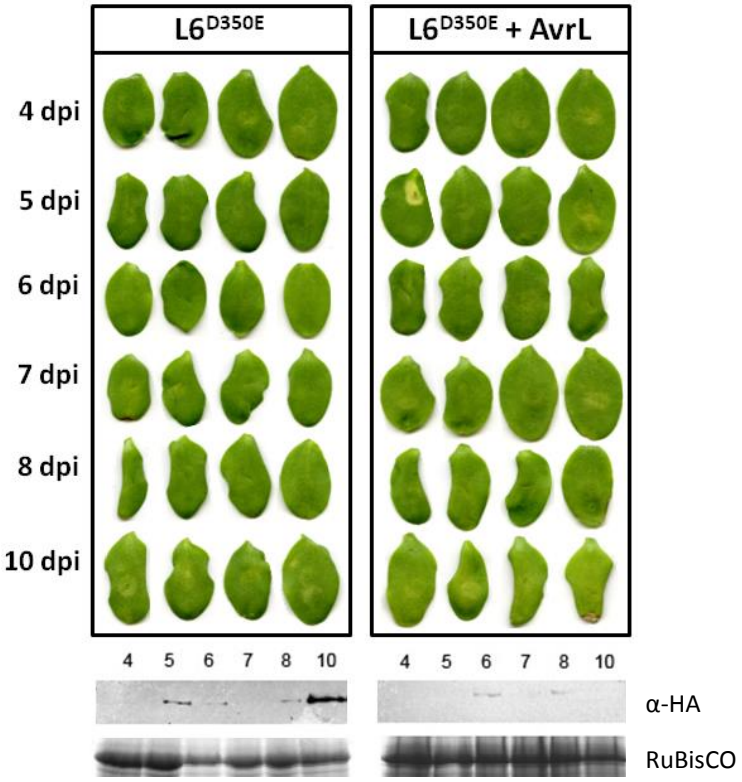
As previously observed by Williams *et al.* (2011), a robust HR was generated by the co-expression of CaMV 35S driven genomic *M* (pEG100 gM:HA) and AvrM (pTNSpec AvrM) (Figure 4-5B). Though, when pEG100 gM:HA was overexpressed alone some chlorosis and necrosis, symptomatic of a spontaneous cell death, was observed 8-12 dpi. Some chlorosis was also observed 12 dpi when pEG100 gM:HA was overexpressed with pTNSpec avrM (cDNA). Equal volumes of infiltration cultures were mixed together for co-infiltration, so delayed HR development was expected.

The intensity of the HR generated by M, independent of the flax rust cognate effector AvrM, was increased when it was expressed from a modified pE1776 vector with a chimeric octopine and manopine synthase promoter that has previously been reported to generate as much as 156X more expression than CaMV 35S (Figure 4-5A; Day *et al.*, 2006; Ni *et al.*, 1995). The ensuing early onset of the HR and its intensity did not permit evaluation of M protein expression levels.

Given the autoactive HR of M driven by the p1776 promoter, and the inability of SDS-PAGE and immunoblot analysis to detect the protein in extracts of infiltrated leaves,

the use of this promoter would not facilitate straightforward visualisation of functional changes caused by Walker B motif mutations in M. All subsequent transient expression studies were therefore done in the pEG100 expression vector where *M* is driven by the CaMV 35S promoter.

It should be noted that neither infiltration of *A. tumefaciens* GV3101 alone, nor GV3101 containing pTNSpec AvrM or pTNSpec avrM triggered any HR over the 12 dpi (Figure 4-6).


L6D350E
L6D350E + AvrL
4 dpi
5 dpi
6 dpi
7 dpi
8 dpi
10 dpi
4 5 6 7 8 10
4 5 6 7 8 10
α-HA
RuBisCO

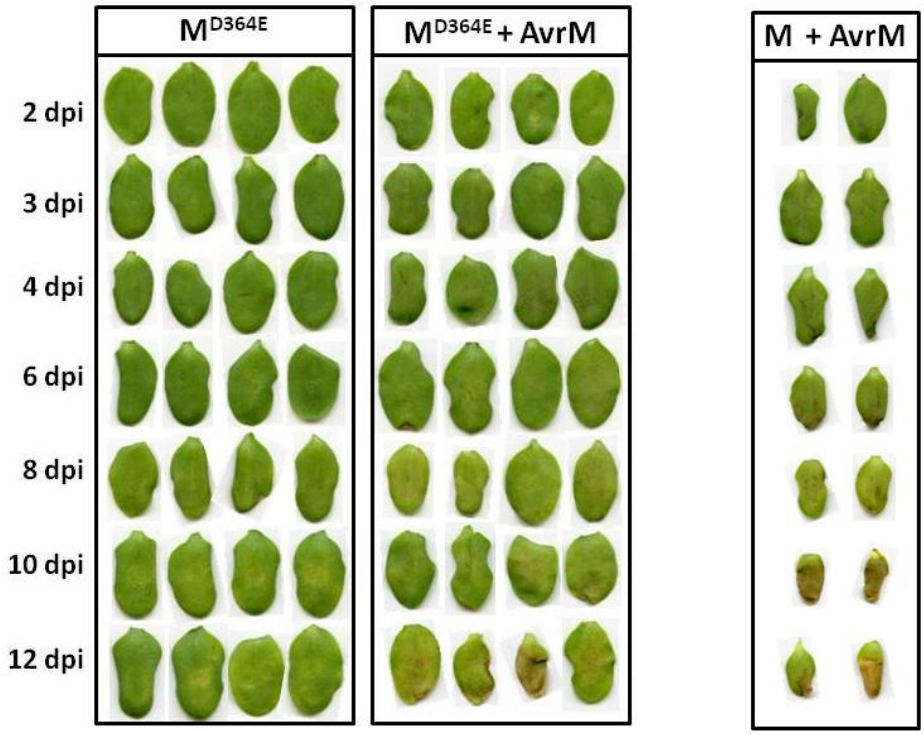
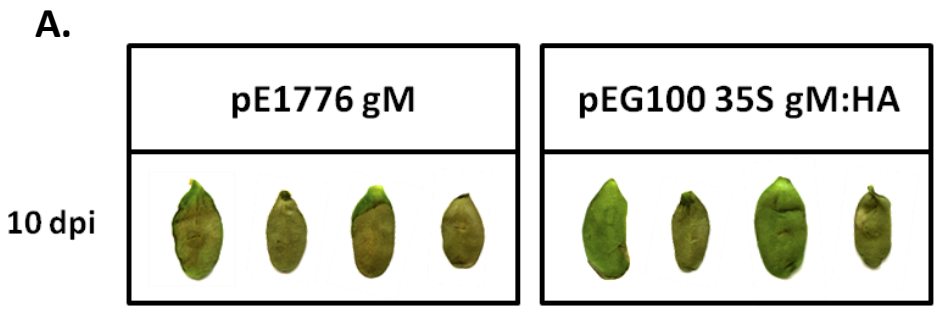

MD364E
MD364E + AvrM
M + AvrM
2 dpi
3 dpi
4 dpi
6 dpi
8 dpi
10 dpi
12 dpi

**Figure 4-4: Equivalent mutations of the proposed catalytic residues within the Walker B regions of the TIR-NB-LRR proteins M and L6 have different functional consequences**

An aspartate to glutamate substitution of the proposed catalytic residue in M, D364, does not change the function of M. The same substitution of the analogous residue in $L6^{D350E}$ generates an inactive protein that cannot signal a HR in the presence of L6's cognate effector, AvrL. Infiltrations were performed in 11 day-old *L. usitatissimum* cv. Hoshangabad cotyledons with recombinant *A. tumefaciens* GV3101 strains resuspended to an OD of 1 in infiltration media. Equal volumes of *A. tumefaciens* suspensions were mixed together for co-infiltrations with AvrM and AvrL. The expression of all constructs was driven by the strong constitutive CaMV promoter; pEG100 $gM^{D346E}$:HA, pTNSpec $gL6^{D350E}$:HA, pTNSpec AvrL and pTNSpec AvrM. Cotyledons were harvested and snap frozen in liquid nitrogen 2, 3, 4, 6, 8, 10 and 12 dpi for the analysis of protein expression. Anti-HA immunoblots show the levels of mutant protein expression from infiltration alone and with AvrM or AvrL. Coomassie blue staining of the large subunit of RuBisCO indicates loading in the SDS-PAGE.


A.
pE1776 gM
pEG100 35S gM:HA
10 dpi

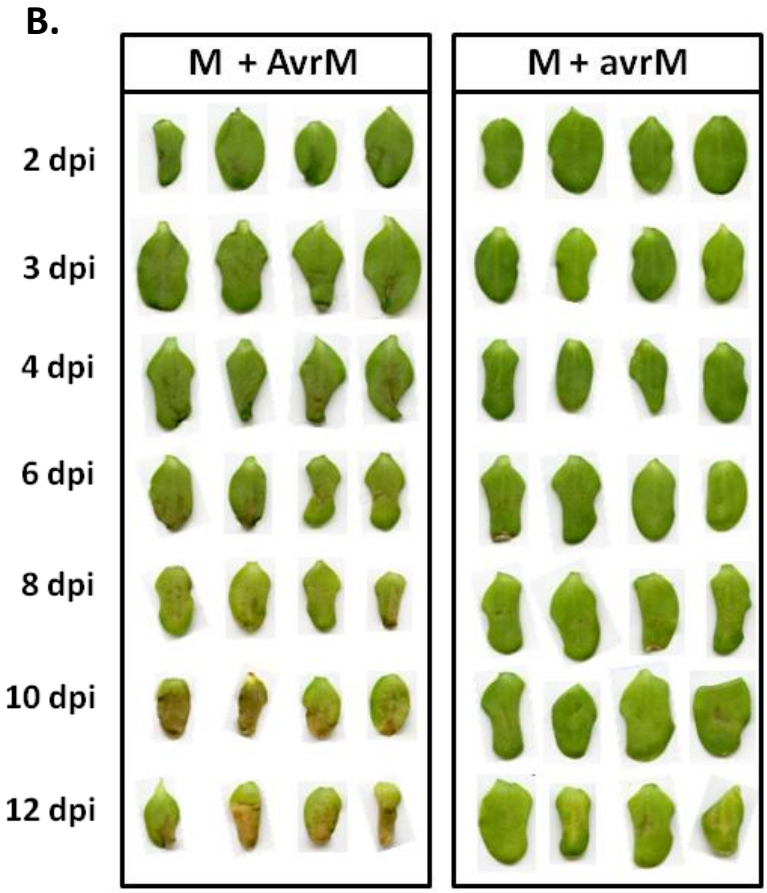

B.
M + AvrM
M + avrM
2 dpi
3 dpi
4 dpi
6 dpi
8 dpi
10 dpi
12 dpi

**Figure 4-5: Overexpression of M generates spontaneous cell death in Hoshangabad cotyledons**

Infiltrations were performed in 11 day-old *L. usitatissimum* cv. Hoshangabad cotyledons with recombinant *A. tumefaciens* GV3101 strains resuspended to an OD of 1 in infiltration media. Equal volumes of *A. tumefaciens* suspensions were mixed together for co-infiltrations with AvrM and avrM. Images were taken 10 dpi. **A)** Independent expression of genomic *M* was driven by the very strong constitutive chimeric octopine and manopine synthase promoter in a modified pE1776 vector (pE1776 gM) and the strong constitutive CaMV 35S promoter (pEG100 35S gM:HA). Spontaneous cell death was generated by the overexpression of M with both promoters, less with CaMV 35S. **B)** Co-expression of genomic *M* driven the CaMV 35S promoter (pEG100 35S gM:HA) with AvrM and avrM in pTNSpec expression vectors. Some chlorosis was observed from the co-expression of M driven by its native promoter with AvrM. A robust HR was observed after the same period when M was expressed with the strong constitutive CaMV 35S promoter.

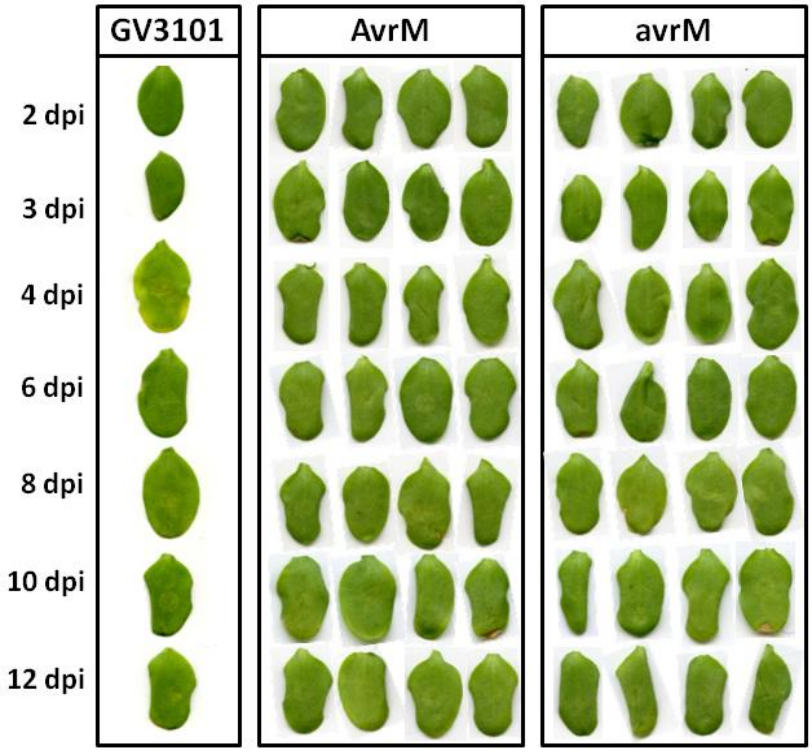



**Figure 4-6: Transient expression of *A. tumefaciens* GV3101 and CaMV 35S driven AvrM and avrM give no visible HR or chlorosis in Hoshangabad cotyledons**

Infiltrations were performed in 11 day-old *L. usitatissimum* cv. Hoshangabad cotyledons with recombinant *A. tumefaciens* GV3101 strains resuspended to an OD of 1 in infiltration media. No HR was observed during the 12 dpi of the untransformed control *A. tumefaciens* GV3101 or the pTNSpec AvrM and pTNSpec avrM CaMV 35S overexpression vectors.

### 4.2.3 Functional analysis of Walker B mutants

Each Walker B mutation was made in a genomic version of the *M* gene with a C-terminal 3X HA tag. Expression was driven by the CaMV 35S promoter in the pEG100 binary vector (pEG100 gM:HA). *A. tumefaciens*-mediated transient expressions were performed in 11 day-old *L. usitatissimum* cv. Hoshangabad cotyledons. Mutants were infiltrated alone (Figure 4-7 & Figure 4-8) and co-infiltrated with CaMV 35S driven AvrM (Figure 4-9 & Figure 4-10) and avrM (Figure 4-11 & Figure 4-12) in pTNSpec vectors. Equal volumes of cultures, diluted to an $OD_{600nm}$ of 1.0, were mixed together for co-infiltrations. Cotyledons were visually observed for symptoms of the HR until 12 dpi. Mutant expression was evaluated with anti-HA immunoblots of proteins extracted from cotyledons 2, 3, 4, 6, 8, 10 and 12 dpi of mutants alone, and with AvrM and avrM. An HA-tagged M protein heterologously expressed in *P. pastoris* was used as a positive control in each immunoblot (data not shown). The large subunit of RuBisCO, the most prominent polypeptide in extracts of total leaf protein separated by SDS-PAGE and stained with Coomassie blue, was used to indicate protein loading.

Only the $M^{D363A+D364A}$ mutation generated a protein that did not activate a HR and was therefore deemed inactive. The $M^{D364A}$, $M^{D364E}$, $M^{D366E}$, $M^{E367A}$ and $M^{D366A+E367A}$ mutants all activated the HR when co-infiltrated with AvrM. The $M^{D364A}$ and $M^{E367A}$ mutants developed less intense HRs over the 12 day period than wild-type M in AvrM co-infiltrations. As has previously been shown by Williams *et al.* (2011), the $M^{D364E}$ mutation retained its AvrM induced activity, but generated a less intense and slightly delayed HR than wild-type M over the 12 day period.

The $M^{D363A}$ and $M^{D366A}$ mutants both displayed some autoactivity, $M^{D366A}$ more than $M^{D363A}$. An increased HR intensity was observed from co-infiltration of the $M^{D366A}$ mutant with AvrM. Both autoactive phenotypes were severely reduced in avrM co-infiltrations. A full summary of the mutant phenotypes and HR onsets can be found in Table 4-2.

As shown in the immuonblots from the single infiltration of each, all mutants were stably expressed in the flax cotyledons. No HA-tagged protein was detected after the co-infiltration of any active mutant with AvrM. Due to the reduced inoculum strength,

mutant protein detection from co-infiltrations with avrM was delayed 1-2 days compared to when they were expression alone. The weak autoactive mutant $M^{D363A}$ was detected after co-infiltration with AvrM, but the $M^{D366A}$ mutant was not (Figure 4-9 & Figure 4-10).

**Table 4-2: Walker B mutant phenotypes as determined via *A. tumefaciens*-mediated transient expression in Hoshangabad cotyledons**

| Protein | Motif | Function | HR Onset (dpi) | | |
|---|---|---|---|---|---|
| | | | Alone | + AvrM | + avrM |
| M | DDVDE | Autoactive | 8 | 2-3 | 12 |
| $M^{D363A}$ | ADVDE | Autoactive (very weak) | 10-12 | 8 | - |
| $M^{D364A}$ | DAVDE | Active (weak) | - | 3 | - |
| $M^{D363A+D364A}$ | AAVDE | Inactive | - | - | - |
| $M^{D364E}$ | DEVDE | Active (weak) | - | 3-4 | - |
| $M^{D366A}$ | DDVAE | Autoactive (weak) | 6-8 | 2-3 | - |
| $M^{D366E}$ | DDVEE | Active | - | 4 | - |
| $M^{E367A}$ | DDVDA | Active (weak) | - | 3 | - |
| $M^{D366A+E367A}$ | DDVAA | Active | - | 3 | - |
| $L6^{D350E}$ | DEVDE | Inactive | - | - | - |

**- signifies no HR was visually observed**

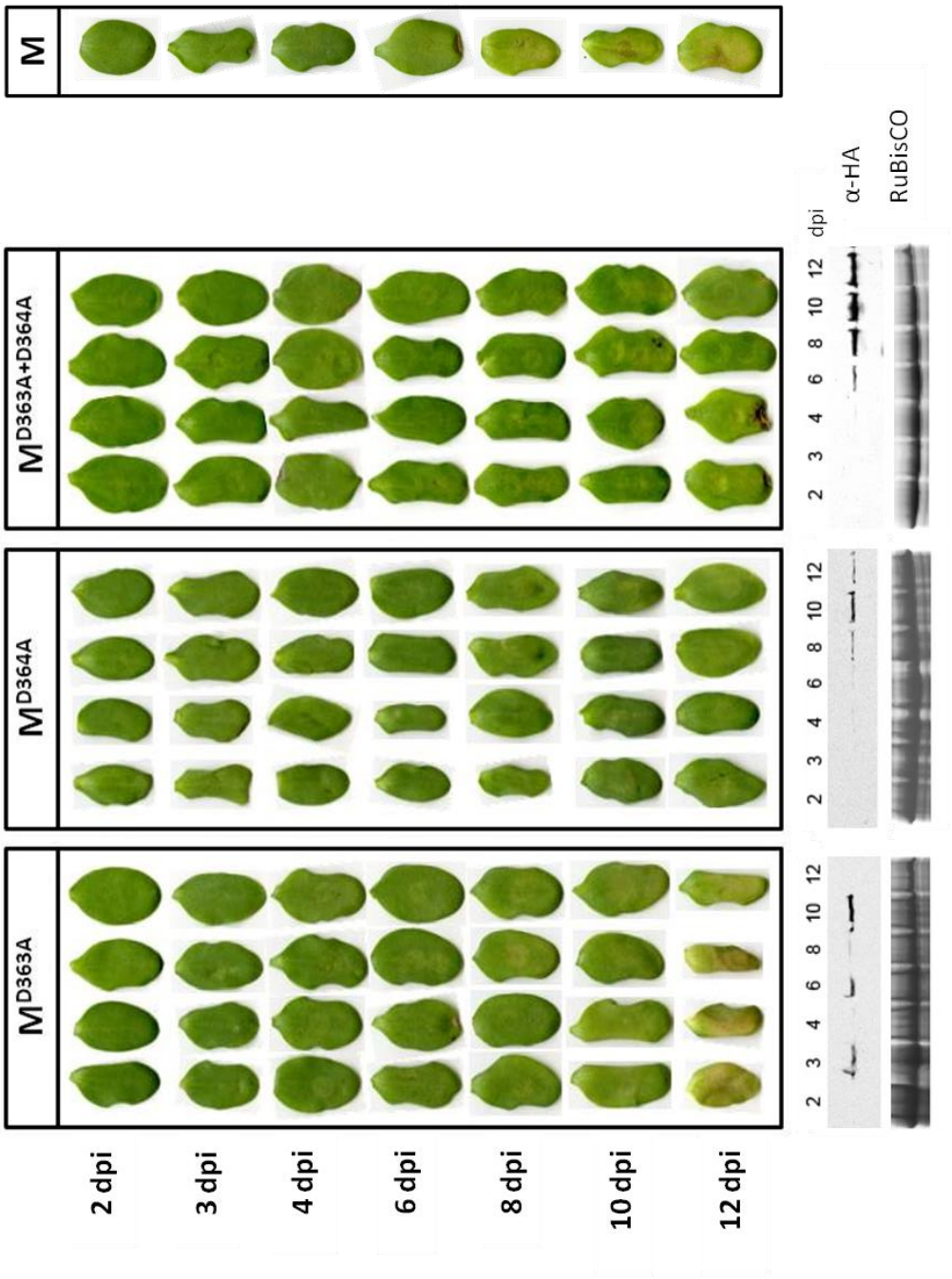



**Figure 4-7: Transient expression in *L. usitatissimum* cv. Hoshangabad of mutations made to the first pair of acidic residues in the Walker B motif of M protein; pEG100 gM$^{D363A}$:HA, pEG100 gM$^{D364A}$:HA and pEG100 gM$^{D363A+D364A}$:HA**

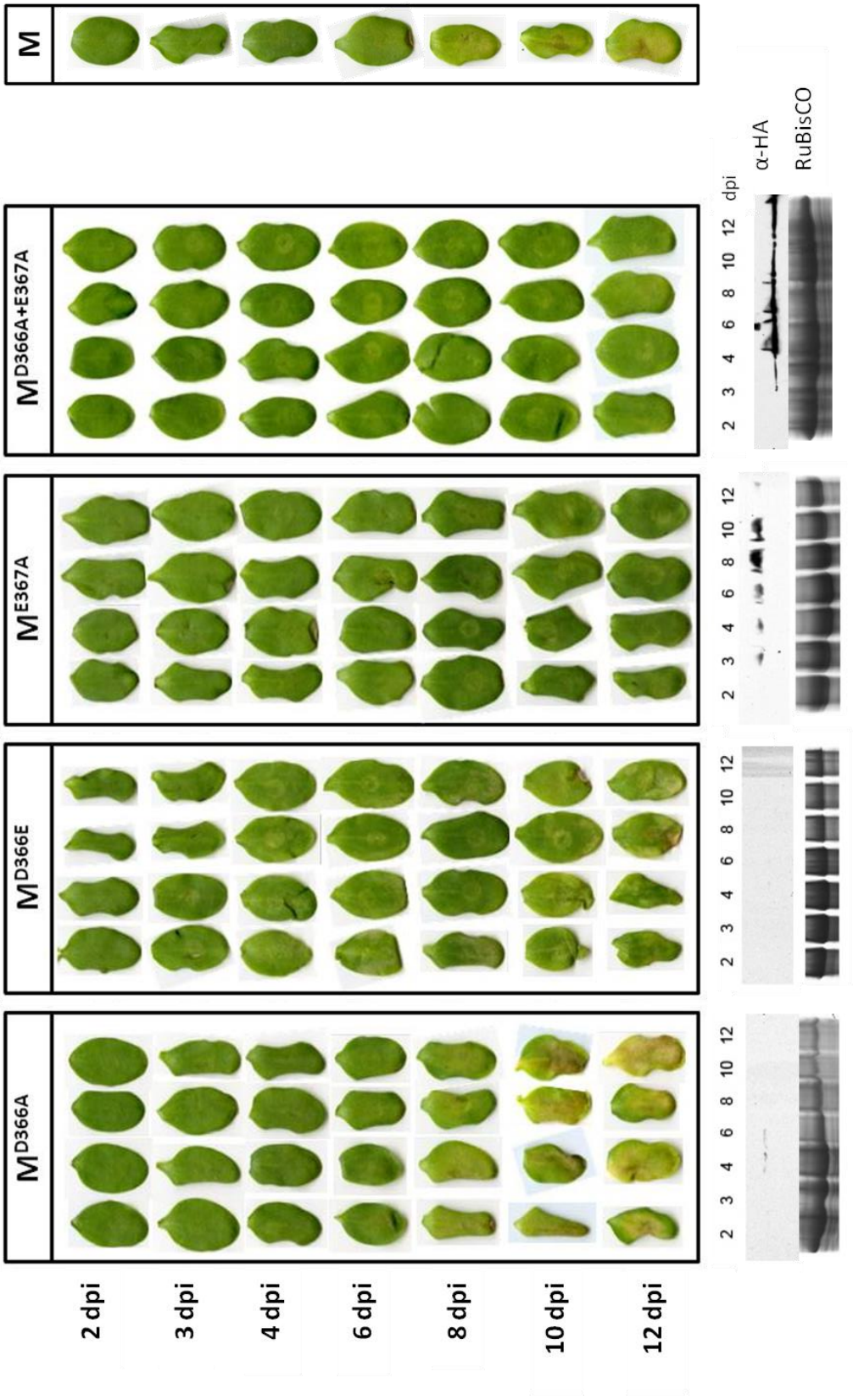



**Figure 4-8: Transient expression in *L. usitatissimum* cv. Hoshangabad of mutations made to the second pair of acidic residues in the Walker B motif of M protein; pEG100 gM$^{D366A}$:HA, pEG100 g:M$^{E367A}$:HA and pEG100 g:M$^{D366A+E367A}$:HA**

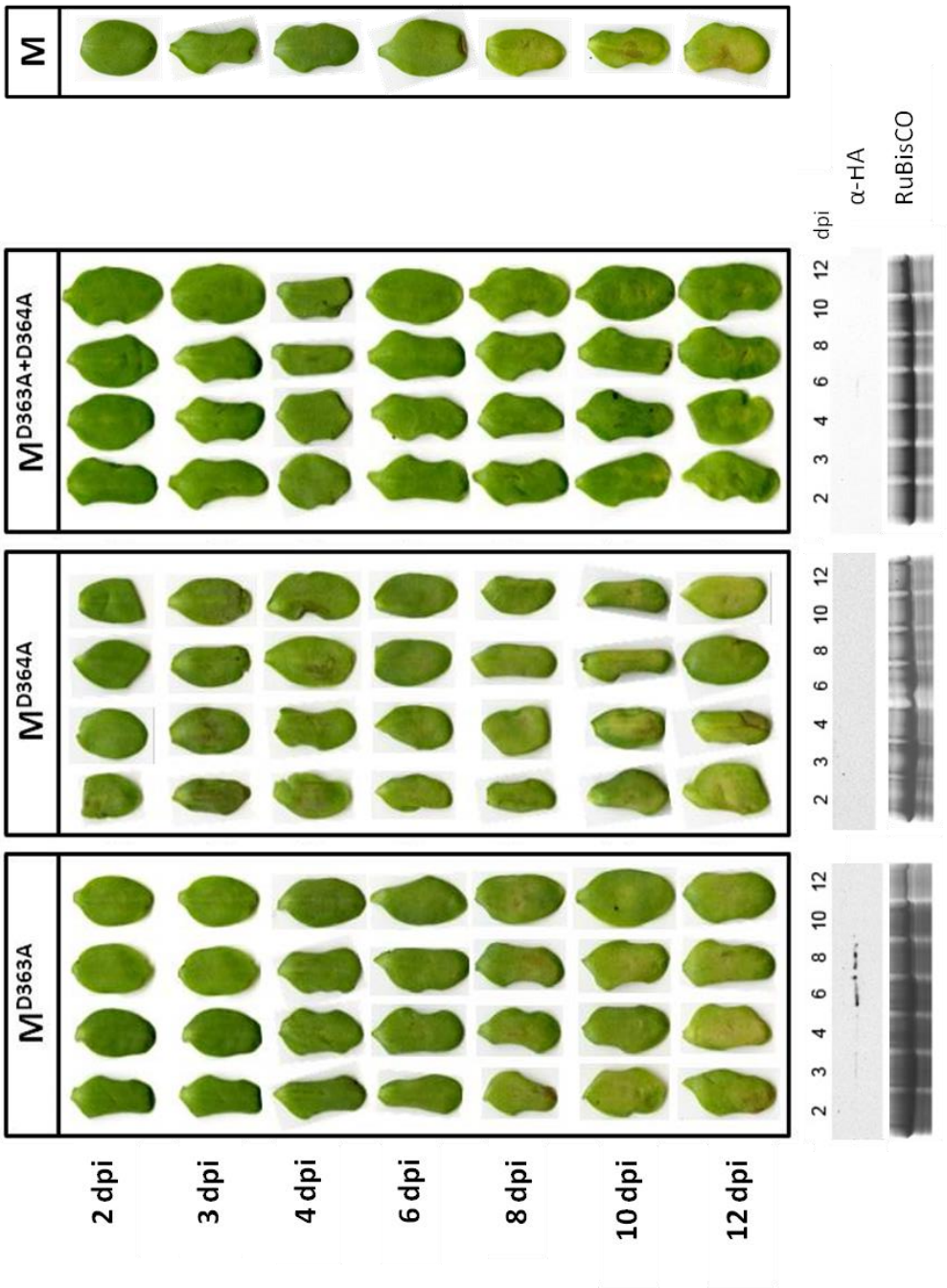



**Figure 4-9: Transient co-expression in *L. usitatissimum* cv. Hoshangabad of AvrM with mutations made to the first pair of acidic residues in the Walker B motif of M protein; pEG100 gM$^{D363A}$:HA, pEG100 gM$^{D364A}$:HA and pEG100 gM$^{D363A+D364A}$:HA + pTNSpec AvrM**

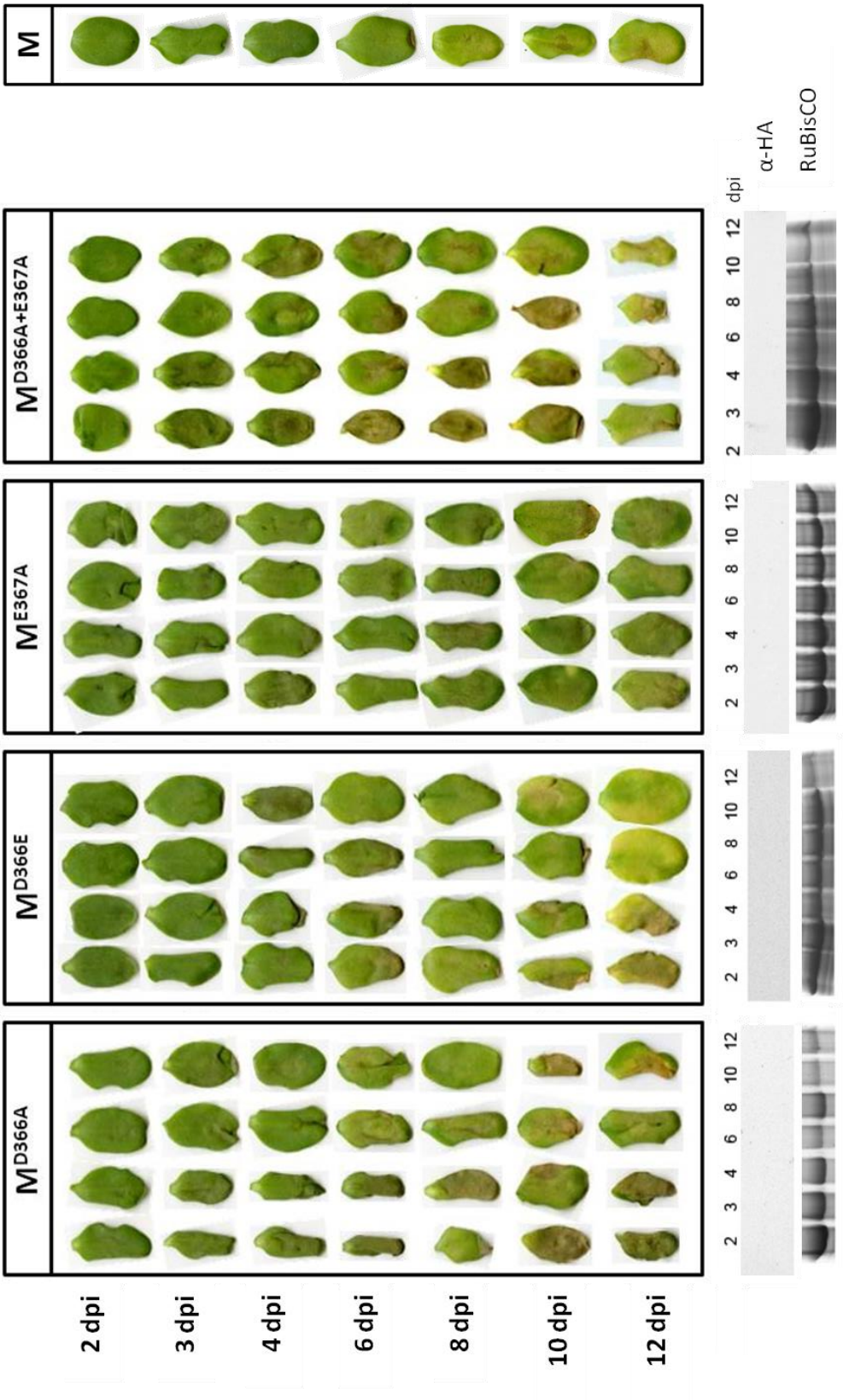



**Figure 4-10: Transient co-expression in *L. usitatissimum* cv. Hoshangabad of AvrM with mutations made to the second pair of acidic residues in the Walker B motif of protein; pEG100 gM$^{D366A}$:HA, pEG100 g:M$^{E367A}$:HA and pEG100 g:M$^{D366A+E367A}$:HA + pTNSpec AvrM**

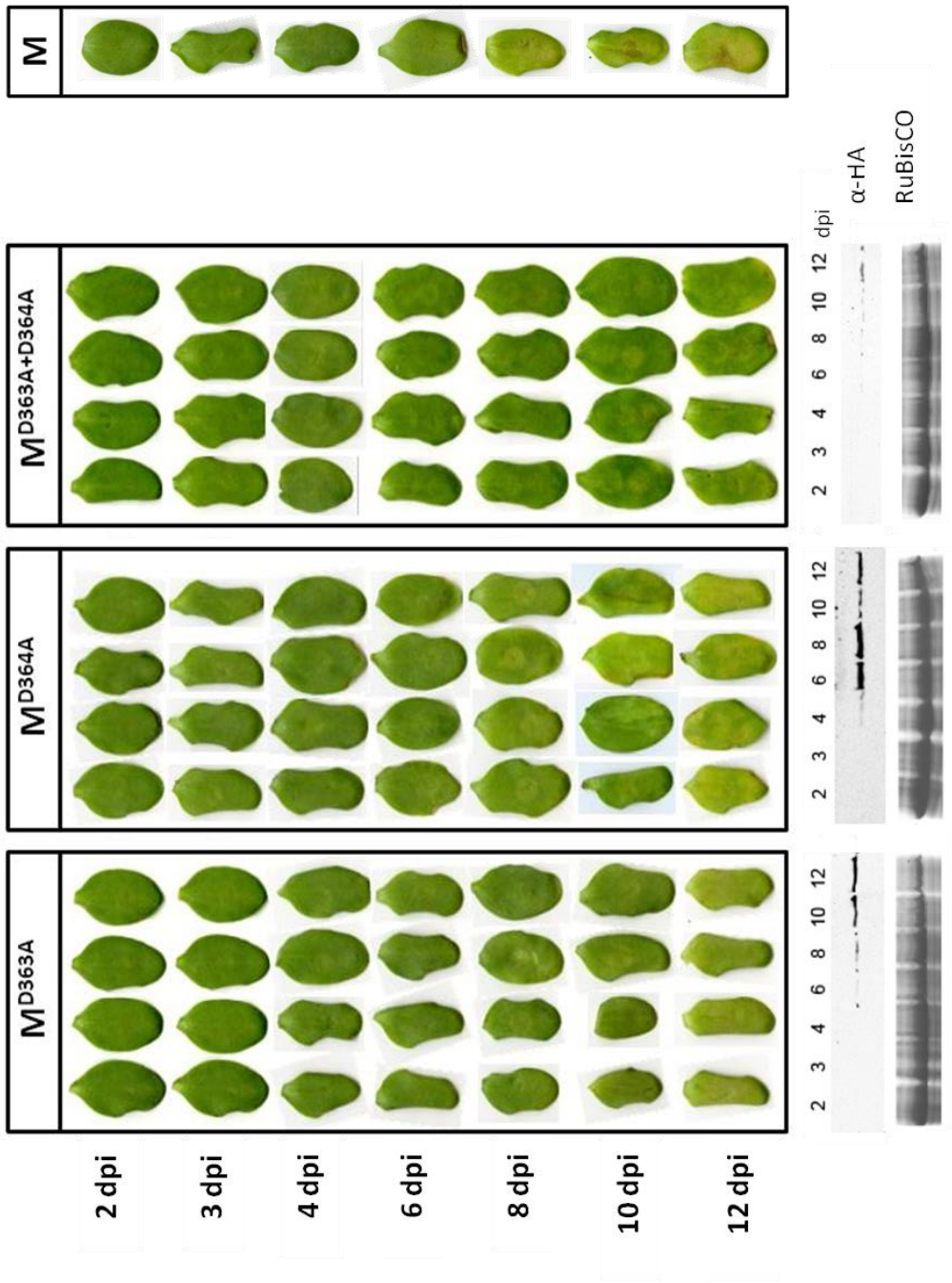



**Figure 4-11: Transient co-expression in *L. usitatissimum* cv. Hoshangabad of avrM with mutations made to the first pair of acidic residues in the Walker B motif of M protein; pEG100 gM$^{D363A}$:HA, pEG100 gM$^{D364A}$:HA and pEG100 gM$^{D363A+D364A}$:HA + pTNSpec avrM**

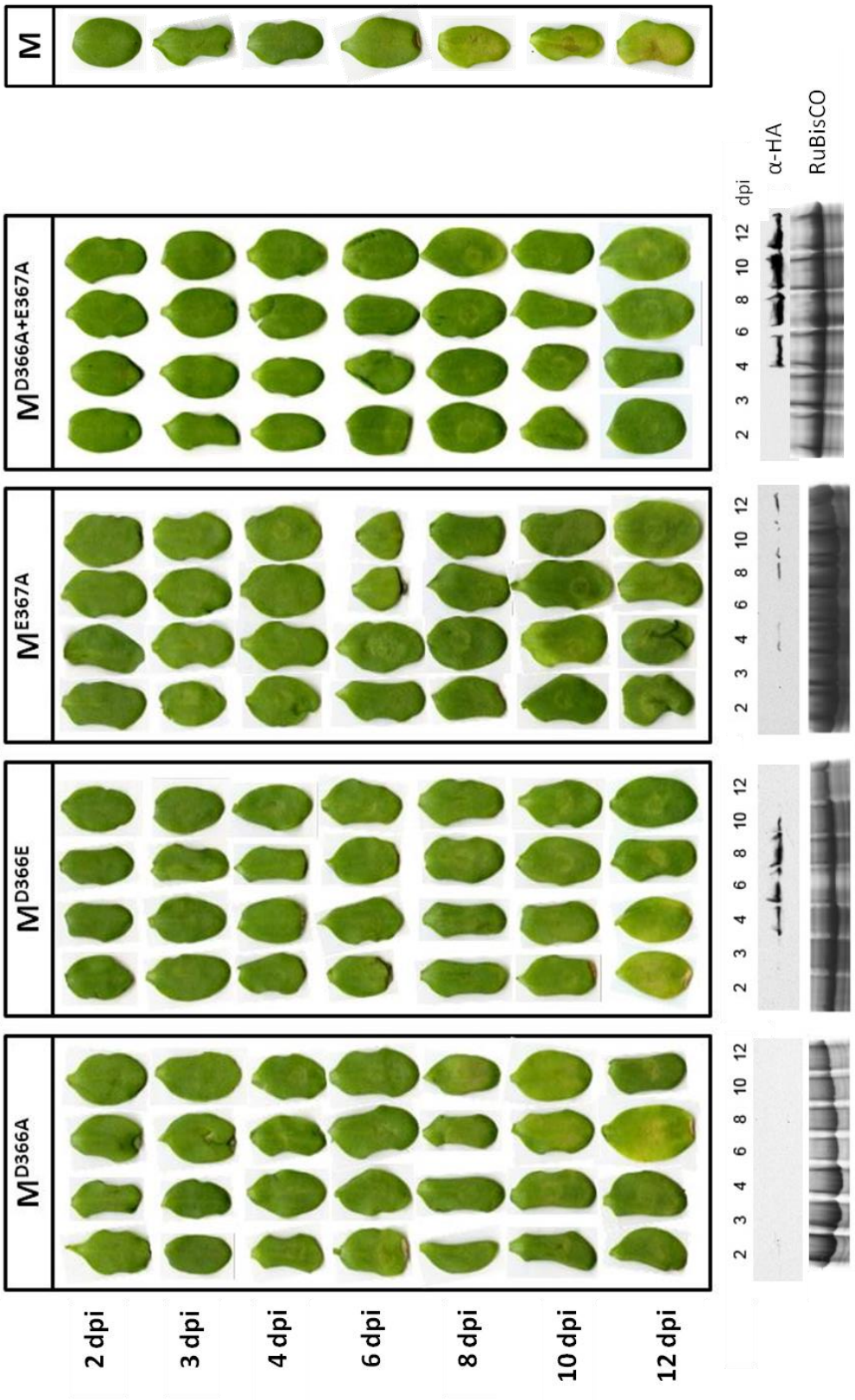



**Figure 4-12: Transient co-expression in *L. usitatissimum* cv. Hoshangabad of avrM with mutations made to the second pair of acidic residues in the Walker B motif of the M protein; pEG100 gM$^{D366A}$:HA, pEG100 g:M$^{E367A}$:HA and pEG100 g:M$^{D366A+E367A}$:HA + pTNSpec avrM**

**Figure 4-7 to Figure 4-12: *A. tumefaciens*-mediated transient expression of Walker B mutants in *L. usitatissimum* cv. Hoshangabad cotyledons**

Expression constructs containing mutant genomic *M* genes with 3X HA C-terminal tags were transiently overexpressed in 11 day-old *L. usitatissimum* cv. Hoshangabad cotyledons with the CaMV 35S promoter (pEG100 gM:HA). Infiltrations were performed with recombinant *A. tumefaciens* GV3101 strains resuspended to an OD of 1 in infiltration media. Equal volumes of *A. tumefaciens* suspensions were mixed together for co-infiltrations with pTNSpec AvrM (Figures 3-9 & 3-10) and pTNSpec avrM (Figures 3-11 & 3-12). Cotyledons were harvested and snap frozen in liquid nitrogen 2, 3, 4, 6, 8, 10 and 12 dpi for the analysis of protein expression. Anti-HA immunoblots show the levels of mutant protein expression from infiltration alone, with AvrM or avrM. Coomassie blue staining of the large subunit of RuBisCO indicates protein loading in the SDS-PAGE.

### 4.2.4 Identification and quantification of bound nucleotide

The $M^{D363A}$, $M^{D364A}$, $M^{D363A+D364A}$ and $M^{D363A+D364A}$ mutant proteins were expressed in and purified from the yeast *P. pastoris* in order to determine the identity and quantity of any bound adenine nucleotides (nucleotide occupancy). To generate each expression construct, site-directed mutagenesis was performed on pPICZ vectors containing the coding region of the *M* gene, minus the sequence encoding the 21-amino acid predicted N-terminal signal region. Each construct contained a sequence encoding an N-terminal 9X histidine tag to facilitate protein purification via nickel affinity chromatography.

Recombinant protein expression, purification and occupancy calculations were carried out as per the methods described in Williams *et al.* (2011). Bound ATP and ADP amounts were measured using an ATP bioluminescence-based assay. Nucleotides were released from the protein via boiling and the subsequent addition of pyruvate kinase facilitated the quantification of ADP levels. Nucleotide occupancies were calculated using protein concentrations estimated from SyproRuby® stained 10% denaturing polyacrylamide gels of the purified mutants and BSA standards. ATP and ADP levels obtained after the extraction and nickel affinity purification performed on a pPICZ empty vector *P. pastoris* pellet were treated as background ATP and ADP concentrations. These values were subtracted from the ATP and ADP concentrations obtained for each mutant (see appendix D). Each mutant was purified a minimum of three times from three independent *P. pastoris* cultures and nucleotide occupancy values are reported as the mean ± standard error (Figure 4-13 & appendix D).

Williams *et al.* (2011) determined that wild-type M is bound with 2 ± 0.1% ATP and 35 ± 3% ADP. The autoactive $M^{D555V}$ mutant displays a preference for bound ATP, but it also contains some bound ADP (18 ± 4% ATP, 5 ± 2% ADP). The inactive mutant, $M^{K286L}$, was purified with negligible amounts of ATP and ADP (0.5 ± 0.1% ATP, 3 ± 1% ADP). $M^{D364E}$, an aspartate to glutamate substitution of the proposed catalytic residue, was reported to have no affect on M function, but did increase the affinity for ADP (1 ± 0.1% ATP, 105 ± 5% ADP) (Williams *et al.*, 2011). These occupancy levels

were used as a comparison for expected nucleotide occupancies of active, autoactive and inactive mutants.

$M^{D363A}$, which displayed weak autoactivity, was purified bound to 0.4 ± 0.2% ATP and 27 ± 9% ADP. The active $M^{D364A}$ mutant contained 4 ± 0.6% of bound ATP and 23 ± 3% of ADP. The non-functional double mutant $M^{D363A+D364A}$ contained very low amounts of both ATP and ADP, 0.4 ± 0.1% and 3 ± 2%, respectively. This latter result is consistent with that reported by Williams *et al.* (2011) for the non-functional P-loop mutant, $M^{K286L}$, which is also purified without any significant amounts of bound ATP or ADP. The autoactive $M^{D366A}$ mutant contained 4 ± 0.4% of bound ATP and 12 ± 2% ADP. Unlike the strongly autoactive mutant, $M^{D555V}$ (Williams *et al.*, 2011), neither autoactive Walker B mutant contained any bound ATP.

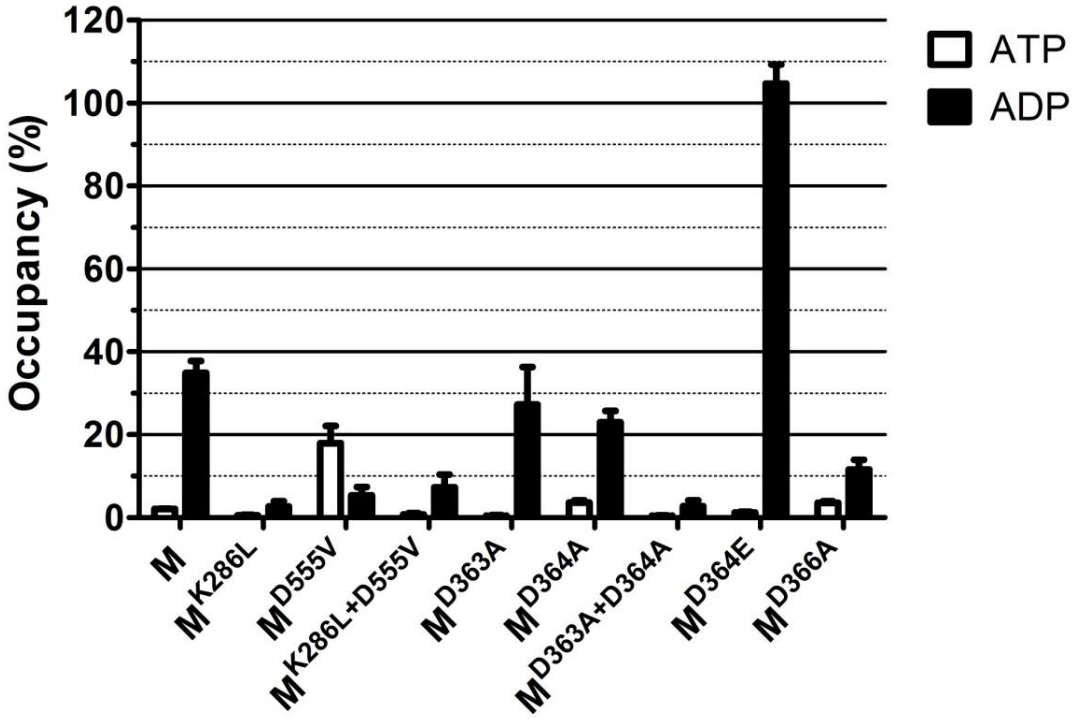



**Figure 4-13: ATP and ADP occupancies of M Walker B mutants compared to wild-type M and other inactive and autoactive M mutants**

To determine nucleotide occupancy levels, protein concentration was estimated using a SyproRuby® stain of nickel affinity purified protein together with a BSA standard curve. Occupancy values represent the mean ± standard error as calculated from four independent protein purifications, except for $M^{D363A+D364A}$ and $M^{D366A}$ which represent the average of three. No significant levels of ATP were bound within any of the mutant proteins. All mutants, except $M^{D363A+D364A}$, which contained negligible amounts of ATP and ADP, contained some bound ADP.

NB: M, $M^{K286L}$, $M^{D555V}$ and $M^{K286L+D555V}$ and $M^{D364E}$ occupancy values are as reported in Williams *et al.* (2011).

### 4.2.5 ATP hydrolysis activity of the $M^{D363A}$, $M^{D364E}$ and $M^{D366A}$ mutants

To measure ATPase activity the two mutants that displayed some autoactivity, $M^{D363A}$ and $M^{D366A}$, and the $M^{D364E}$ mutant that was purified fully occupied with ADP, were purified using a two-step affinity chromatography process. Protein purified using only a one-step nickel affinity chromatography method has previously been shown to have a high background level of ATP hydrolysis from which specific activity attributed to the M protein could not be distinguished (deCourcy-Ireland, 2007; Williams, 2009). Sornaraj (2013) showed that this two-step purification process improves the purity of the purified protein and reduces background ATP hydrolysis to negligible levels.

Each pPICZ expression construct contained a sequence encoding an N-terminal 6X histidine tag followed by the coding region of the *M* gene, minus the 21-amino acid predicted N-terminal signal region, and a sequence encoding a C-terminal 1X *Strep*-Tag®. Each protein was purified first with nickel affinity chromatography and then with *Strep*-Tactin® chromatography. The concentration of each purified mutant protein was determined using a total protein assay (2.7.11) and purity was visualised with 10% SDS-PAGE and Coomassie blue staining. ATP hydrolysis assays were carried out as per the methods described in deCourcy-Ireland (2007) and Sornaraj (2013) and Myosin was used as a positive control for ATPase activity in each assay performed.

Low yields of total protein were achieved for all mutants after the two-step purification process (Figure 4-14). This limited the number of ATP hydrolysis assays that could be performed with each batch of purified protein and therefore full kinetic analysis of each mutant could not be performed. Consequently, rates of hydrolysis at 1 µM ATP were determined for each Walker B mutant and compared to the rates for M, $M^{K286L}$ and $M^{D555V}$ at 1 µM, as determined by Sornaraj (2013). As recommended when doing kinetic measurements of activity, product formation (ADP production) over the course of the assays at this ATP concentration was approximately 10%.

The hydrolysis rates of each mutant were measured at 25°C in the presence of 0.4 µCi of [α-$^{32}$P]ATP and cold (unlabelled) ATP at a concentration of 1 µM. Each assay was performed in triplicate and contained 200-400 ng of total protein depending on the yield of the respective protein. Assays were started with the addition of ATP and

at appropriate time points 2 µL aliquots of the reaction were spotted onto PEI cellulose TLC plates. Assays were conducted for 30 or 60 minutes. The [α-$^{32}$P]ATP and [α-$^{32}$P]ADP spots on the TLC plates were resolved with a solution of 0.25 M LiCl and 2.5% (v/v) formic acid. Dry plates were exposed to a PhosphorImager plate and scanned in a Bio-Rad Molecular Imager to visualise the spots of [α-$^{32}$P]ATP and [α-$^{32}$P]ADP. The formation of [α-$^{32}$P]ADP per minute per µg of protein was used as the measure of the rate of ATP hydrolysis. Myosin was used as a positive control for ADP generation in each assay (Figure 4-15).

Games-Howell *post-hoc* analysis of one-way ANOVA revealed that the rates of ATP hydrolysis (µmol/min/µg ± standard error of the mean) of each Walker B mutant ($M^{D363A}$ 4.1e-007 ± 5.9e-008 µmol/min/µg, $p$ =0.209; $M^{D364E}$ 2.9e-007 ± 3.2e-008 µmol/min/µg, $p$ =0.736; $M^{D366A}$ 2.8e-007 ± 1.0e-008 µmol/min/µg, $p$ =0.565) were not significantly different from the rate of hydrolysis of M (2.4e-007 ± 2.1e-008). The Walker B mutants also did not have significantly different rates of ATP hydrolysis from each other ($p$ >0.05). The rates of M, $M^{K286L}$ (1.5e-007 ± 4.3e-009 µmol/min/µg) and each Walker B mutant were significantly lower than the hydrolysis rate of $M^{D555V}$ (8.3e-007 ± 1.9e-008 µmol/min/µg) (M $p$ =0.000, $M^{K286L}$ $p$ =0.000, $M^{D363A}$ $p$ =0.003, $M^{D364E}$ $p$ =0.000, $M^{D366A}$ $p$ =0.000).

Analysis also showed that the rates of hydrolysis of M ($p$ =0.035) and $M^{D555V}$ ($p$ =0.000) were significantly greater than the rate of $M^{K286L}$. The rates of hydrolysis of $M^{D363A}$ ($p$ =0.042) and $M^{D366A}$ ($p$ =0.009) were also significantly higher than that of $M^{K286L}$. However, the rate of hydrolysis of $M^{D364E}$ ($p$ =0.000) was not significantly different from the rate of hydrolysis achieved by $M^{K286L}$ (Figure 4-16).

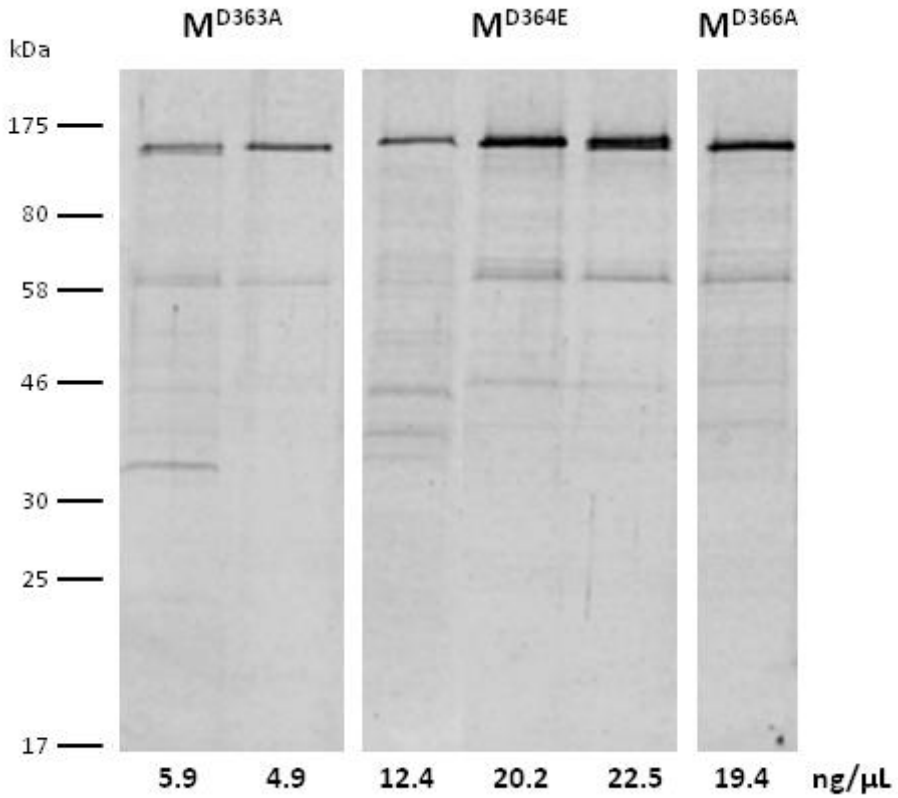



**Figure 4-14: SDS-PAGE of $M^{D363A}$, $M^{D364E}$ and $M^{D366A}$ after purification by nickel affinity and then *Strep*-Tactin® chromatography**

A two-step purification procedure was used to purify the proteins for use in ATP hydrolysis assays. Each protein was purified with nickel affinity chromatography then concentrated before being subsequently purified with *Strep*-Tactin® and concentrated again. An assay for total protein was used to assess the yields of each purified protein and a SyproRuby® stain of the proteins after SDS-PAGE showed overall protein purity.

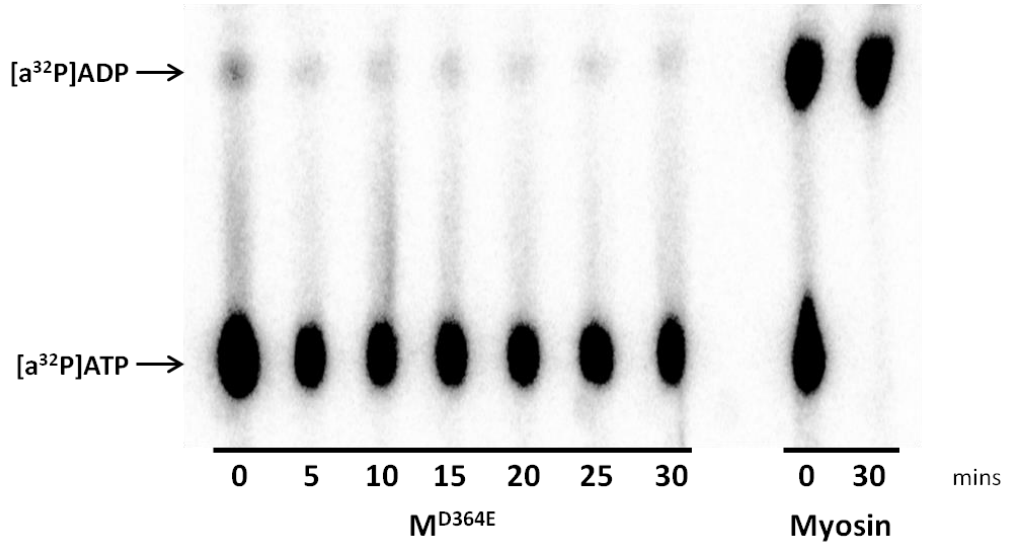



**Figure 4-15: A phosphoimage of a resolved TLC plate showing the substrates and products in a 30 minute ATPase assay of $M^{D364E}$ protein with 1 µM ATP**

ATPase assays were performed in triplicate for each protein using an ATP concentration of 1 µM. This phosphoimage is an example of the time course of the ATPase reaction for $M^{D364E}$ and Myosin. [$\alpha^{32}$P]ATP and [$\alpha^{32}$P]ADP spots were quantified and the formation of [$\alpha^{32}$P]ADP ADP per minute per µg of protein was used to calculate the hydrolysis rate of each protein.

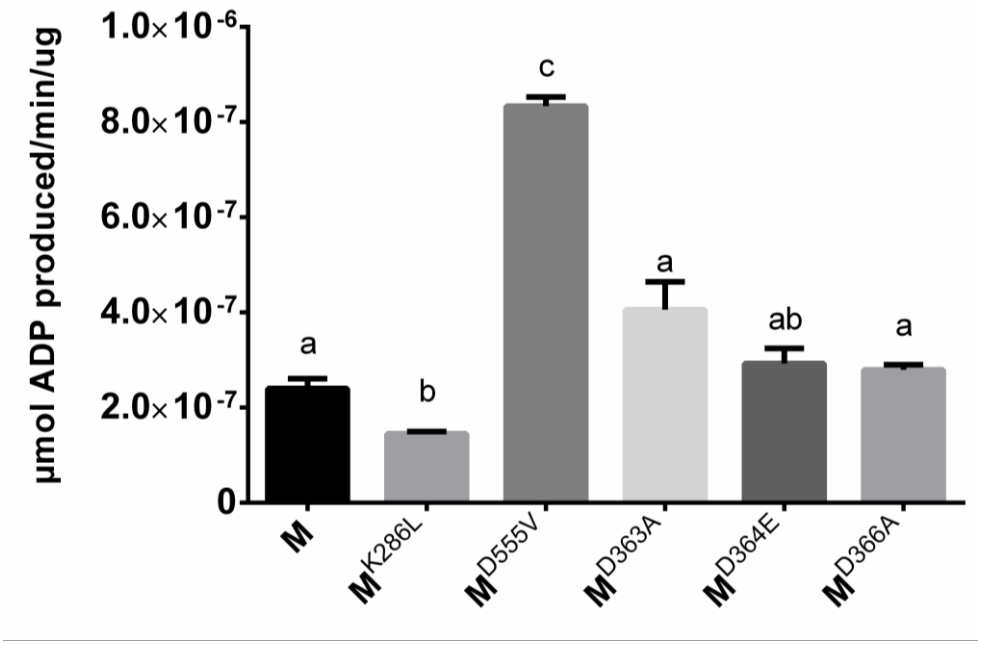



**Figure 4-16: Rates of ADP formation in ATP hydrolysis assays containing 1 µM ATP**

The mean ATP hydrolysis rates for each protein were calculated using data obtained from 1 µM ATP hydrolysis assays performed in triplicate using protein from multiple purifications of each mutant (note $M^{D366A}$ was only successfully purified once). Error bars represent the standard error of the mean. Statistical significance was calculated by one-way ANOVA with Games-Howell *post-hoc* test ($p$ =0.05). The ATP hydrolysis rates of $M^{D363A}$, $M^{D364E}$ and $M^{D366A}$ are not significantly different from M. $M^{K286L}$ has a significantly lower rate of hydrolysis than each of the other proteins, except for $M^{D364E}$. The rate of $M^{D555V}$ is significantly higher than all other proteins tested.

## 4.3 Discussion

M, like many other TIR-NBS-LRR and STAND proteins, contains an extended Walker B motif. In addition to the archetypal acidic pair, its Walker B motif contains two additional downstream acidic residues. As the biochemical roles of the Walker B amino acids within M are unknown and mutation of the proposed catalytic residue does not result in autoactivity, it is possible that they are not equivalent to those in other NB-LRR proteins.

Given that M has been shown to hydrolyse ATP (Sornaraj, 2013) many questions were raised about the mechanism of this activity. For instance, which residue(s) within the Walker B motif of the M protein is responsible for catalysing ATP hydrolysis? Are the first two aspartate residues coordinating $Mg^{2+}$ binding and catalysing ATP hydrolysis, as in other P-loop NTPases? Is there flexibility or sharing of roles within this region? Do the additional acidic residues also have functional roles?

In an attempt to answer these questions, the acidic Walker B residues of M were mutated and their phenotype assessed. Those with a function different to that of wild-type M were expressed in *P. pastoris* and the nickel affinity purified protein assayed for any bound ATP or ADP.

Despite having a high level of amino acid identity, 86% over entire protein and 82% over the NB-ARC domain, M and L6 appear to function differently. While mutation of the proposed catalytic residue in M (D364E) reduces the strength of the HR, the equivalent mutation in L6 (D350E) inactivates the protein (Figure 4-4). These results suggest that the second conserved aspartate residue may have different functions in M and L6, or that the architure of the M and L6 NB pockets are subtly different, or the overall role of ATP hydrolysis is different in these proteins.

### 4.3.1 Overexpression of M generates spontaneous cell death

Transient overexpression of M with CaMV 35S generated very low levels of spontaneous cell death in the infiltrated flax cotyledons that was visible 8-12 dpi. This phenotype was stronger when the expression of M was driven by the chimeric octopine and manopine synthase promoter, which as mentioned earlier, has been

reported to generate as much as 156X more expression than CaMV 35S (Day *et al.*, 2006; Ni *et al.*, 1995).

Pathogen-independent activation has previously been observed for the *A. thaliana* proteins RPS2, RPS4 and RPP13 and the potato protein Rx during their constitutive overexpression in the heterologous tobacco expression system (Bendahmane et al. 2002; Leister & Katagiri, 2000; Oldroyd & Staskawicz, 2000; Rentel et al., 2008; Zhang *et al.*, 2004). It was therefore not unexpected that autoactivity was the result of the overexpression of M. That an autoactive phenotype was observed suggests that negative regulation of M signalling was overcome (implying that a protein partner may be involved in negative regulation), or that a basal level of activity was raised above a threshold level for ETI signalling. A similar result is observed when the L6 protein is overexpressed with the CaMV 35S promoter, but only with genomic L6. The L6 cDNA does not elicit an effector-independent response (Howles *et al.*, 2005). However, pathogen-independent necrosis is not always observed after NB-LRR *in planta* overexpression. When expressed with CaMV 35S in the heterologous host tobacco, RPP1-WsB is not autoactive, suggesting that not all NB-LRRs are regulated in the same manner (Krasileva *et al.*, 2010).

The early onset of cell death from the solitary expression of M with the chimeric octopine and manopine synthase promoter meant that visualisation of any gain-of-activity phenotypes by the expression of mutant M proteins from this construct was not possible. Cotyledon senescence occurred before the HR was fully developed and this system was therefore deemed unsuitable for functional analysis. Whilst it did cause some effector-independent cell death, CaMV 35 driven expression generated a level of M protein that could easily be detected in immunoblot analysis. It also allowed both gain- and loss-of-function phenotypes to be easily detected visually.

### 4.3.2 M requires the canonical acidic pair to activate the HR

Joint mutation of the first two conserved aspartate residues in M, $M^{D363A+D364A}$, produced a non-functional protein with significantly reduced adenine nucleotide binding capacity; 0.4 ± 0.1% ATP and 3 ± 2% ADP compared to 2 ± 0.1% ATP 35 ± 3% ADP in wild-type M. Loss-of-function *in planta* was not due to a lack of protein

expression as the anti-HA immunoblot showed a protein band of the expected size for full-length M protein (Figure 4-7).

Loss-of-function as a result of mutation of analogous residues in the CC-NB-LRRs Rx and RPS2 has previously been demonstrated (Bendahmane *et al.*, 2002; Tao *et al.*, 2000). Similar inactivity was also observed after the double mutation of the two conserved aspartate residues in the STAND protein MalT from *E. coli* (Marquenet & Richet, 2007). This result for M is the first time that this first conserved aspartate pair has been shown to be essential for the function of a TIR-NB-LRR protein. Further investigation of their individual roles in protein activation and function were investigated by separate mutations.

### 4.3.3 Mutation of the proposed magnesium coordinating residue

Mutation of the proposed $Mg^{2+}$coordinating aspartate to alanine, $M^{D363A}$, generated a protein with weak autoactivity that retained bound ADP (27 ± 9%). Interestingly, both inactive and autoactive phenotypes have been reported after various substitutions of analogous aspartates in different STANDs. In general, when autoactive proteins have been generated by mutation of the $Mg^{2+}$ coordinating aspartate the ability to bind nucleotides, in particular ATP, has not been impaired, but their ATPase activity has (Marquenet & Richet, 2007, Danot *et al.*, 2009). An asparagine substitution of this residue in the STAND protein MalT trapped the protein in the active ATP-bound state, thereby reducing ATPase activity and generating autoactivity (Marquenet & Richet, 2007).

In I-2, an inactive protein was created when aspartate was substituted with cysteine (Tameling *et al.*, 2006). As the activity of I-2 is dependent on $Mg^{2+}$, this result was attributed to the inability of the mutant to coordinate the catalytic ion (Tameling *et al.*, 2002). In MalT, an alanine substitution also generated an inactive protein that was unlikely to be able to bind ATP (Marquenet & Richet, 2007).

As an autoactive phenotype, albeit it a weak one, was generated by the $M^{D363A}$ mutation its biochemical properties were further investigated. No ATP was bound within the purified mutant protein, as would be expected if this mutation trapped the protein in the proposed active ATP-bound conformation, but this mutant did contain

27 ± 9% ADP after nickel affinity purification. It is unknown whether ATP was not detected because the ATP-bound state is unstable, or because this mutation prevents ATP coordination. The autoactive $M^{D555V}$ mutant is purified bound with greater levels of ATP than ADP making the former possibility less likely. If the active state of M is ATP-bound the latter is also unlikely given the autoactive phenotype of this mutant.

Investigation of the ability of $M^{D363A}$ protein to hydrolyse ATP was carried out in ATPase assays with 1 µM ATP. Statistical analysis of rate data obtained from these assays and comparable assays of M, $M^{K286L}$ and $M^{D555V}$ performed by Sornaraj (2013) showed that this mutant does not have a significantly different hydrolysis rate to M. It does have a statistically significantly higher rate than $M^{K286L}$ and a lower rate than $M^{D555V}$. Unfortunately full kinetic analysis was not possible due to the very low yields of protein after the two-step chromatography procedure. Further analysis is required before the biochemical role of this residue can be explored and confirmed.

### 4.3.4 Mutations of the proposed catalytic residue

If ATP hydrolysis is required for the deactivation of an activated NB-LRR protein like M, it is expected that an amino acid substitution that does not support ATPase activity may result in its constitutive activation because it sustains the active ATP bound state. Autoactive phenotypes have been demonstrated with glutamate substitutions of the catalytic bases of the RPS5 and I-2, both CC-NB-LRR proteins and from alanine substitutions in the *E. coli* transcription activator MalT (Ade *et al.*, 2007; Marquenet & Richet, 2007; Tameling *et al.*, 2006). The I-2 mutant displayed a reduced ATPase activity indicating that the second conserved acidic residue is likely its catalytic base. This was also observed for the MalT mutant, with the additional observation that it was purified with bound ATP, in contrast to wild-type MalT, which was bound to ADP (Marquenet & Richet, 2007; Tameling *et al.*, 2006).

Interestingly, in many other ATP/ADP binding proteins, including other STANDs, glutamate is the catalytic residue. Like aspartate, glutamate is a negatively charged, carboxylate moiety containing amino acid and can therefore polarise the water molecule that is required for ATP hydrolysis. In RecA, a DNA dependent ATPase, a glutamate located outside of the Walker B motif maintains the catalytic role of the

second or third aspartate or glutamate in a classical Walker B. Its substitution with aspartate causes the protein to exhibit a 100-fold lower $k_{cat}$ or hydrolysis rate and an enhanced enzyme activity. This substitution shortens the amino acid side chain, increasing the distance between the activating carboxylate moiety and the activating water molecule and thus allows for a reduction in the hydrolysis rate. Like observed in I-2 and MalT, this RecA mutant is predicted to exist longer in the active ATP-bound state (Campbell & Davis, 1999).

Both glutamate and alanine residues at the proposed site of the catalytic residue in M appear to be functionally interchangeable with aspartate. Instead of causing autoactivity, as observed in RPS5 and I-2, the $M^{D364E}$ and $M^{D364A}$ mutations retained the ability to activate the HR. While these amino acids do support the overall function of M, both mutations slightly delayed HR onset and/or reduced its strength when co-expressed with AvrM (Figure 4-4 & Figure 4-9).

As reported by Williams *et al.* (2011) the $M^{D364E}$ mutant was purified with a higher amount of bound ADP compared to M, suggesting that it has an increased affinity for ADP (bound ADP 105 ± 5% compared with 35 ± 3.0% bound to M). The rate of ATP hydrolysis of $M^{D364E}$ was not statistically different to the rate of $M^{K286L}$ or M (Figure 4-16). It is therefore likely that this mutant can perform ATP hydrolysis, but that this activity is impaired and hence results in a slower and less robust HR than that observed for M. Unfortunately, poor protein yields limited the number of ATP hydrolysis assays that could be performed with this mutant. Clearly, more protein, tested at more ATP concentrations, will be required before a definitive statement can be made regarding its ATPase activity.

Given that neither the D364E nor D364A substitution caused inactivity or autoactivity, it is unlikely that D364 is the catalytic base. However, given that the HR does seem to be partially compromised, and if ATP hydrolysis is crucial for the activity of M, it is still possible. It is equally possible that one of the additional acidic residues in the Walker B motif is the catalytic base. A semi-cooperative amino acid arrangement within the Walker B could result in compromised polarisation of the catalytic water, which would reduce the protein's ability to activate signalling, but not abolish it. As the HR activated by this mutation appears weaker and more

chlorotic than the M-activated HR this latter scenario is likely. As additional acidic residues are present in other TIR-NB-LRR Walker B domains similar flexibility may also be observed in them.

The possibility that these phenotypes are due to the fact that the hydrolysis of ATP is not required for M to activate the HR, but is rather required as a deactivation step that allows M to cycle and amplify the HR cannot be ruled out. Nor can the possibility that the active state is ATP-bound and that ATP binding is impaired or destabilised by these mutations. It is also possible that more than one biochemical reaction is being performed in the NBS and activation of M involves more than just a simple β-γ ATP hydrolysis reaction.

### 4.3.5 Are the additional acidic amino acids distal to the archetypal Walker B acidic pair important for function?

Mutation of aspartate 366 to alanine generated an autoactive phenotype, while its substitution with a glutamate had no affect on function. As described above, autoactivity generally results when a protein is trapped in an ATP-bound, pre-hydrolytic state that mimics the active state. However, no significant amount of ATP was found bound with the purified $M^{D366A}$ protein; 4 ± 0.4% of bound ATP and 12 ± 2% ADP. As mentioned above, the ATP-bound state may not be stable (it may dissociate from the protein during purification) and therefore will not often be observed.

As assessed in ATP hydrolysis assays at 1 μM ATP, this mutant protein does not hydrolyse ATP at a statistically different rate from M, but it does have a lower rate of hydrolysis than $M^{D555V}$. Given the autoactive phenotype of $M^{D366A}$ it remains possible that this residue plays the role of the catalytic base in M. However, as this protein was only successfully purified once by the dual affinity chromatography strategy further investigation of its ATPase activity is required before this residue can definitively be assigned a function in ATP hydrolysis, and a reason can be found for its autoactive phenotype.

Interestingly, this autoactive mutant responded to co-infiltration AvrM by generating a more intense and faster HR. Such a response has previously been observed with the

strongly autoactive $M^{D555V}$ mutant. Very preliminary evidence also showed $M^{D555V}$ may have an increased ATP hydrolysis rate in the presence of AvrM, suggesting that ATP hydrolysis is important for HR activation in M. It will therefore be important, once purification of this mutant is optimised, to see whether AvrM similarily influences its ATPase activity.

While substitution of residue E367 with alanine did not completely abolish the HR, it did weaken it. It therefore appears that the role played by this residue is not essential for function or can be performed by another of the conserved acidic residues. Interestingly, when residues D366 and E367 were together substituted with alanine, $M^{D336A+E367A}$, the resultant HR when co-expressed with AvrM appeared visually stronger compared with the co-expression of M and AvrM. This suggests that the first acidic pair of Walker B aspartates may function more efficiently without the extended Walker B region. What this suggests about the evolution of these additional residues and the roles that they play is unknown.

### 4.3.6 Using evidence from other NB-LRRs to make assumptions about critical residues is not advised

Plant NB-LRR proteins share significant homology, especially in their NB-ARC regions, and for this reason it was expected that they would share common mechanisms of activation and regulation. However, conflicting functional evidence suggests that this may not be the case. As highlighted by the many different functional outcomes of analogous mutations of predicted functional residues, there appear to be many functional arrangements of amino acids within the Walker B motif. Consequently, the biochemical mechanisms of ATP hydrolysis and the role/s of this reaction in R protein activation and/or regulation most likely vary. Functional and biochemical evidence from one NB-LRR should therefore not necessarily be used to draw conclusions about the activation and regulation of another, even the proteins are related.

### 4.3.7 Does ATP hydrolysis play a role in the activation and/or signalling of M?

The role of ATP hydrolysis in M's activation and signalling remains unclear and requires further investigation. While we do know that M has an intrinsic ATPase

activity, it is also unclear which residue is responsible for catalysing this reaction. Unlike analogous mutations in other NB-LRRs, mutation of the second aspartate did not generate autoactivity, providing further evidence that the mechanism of ATP hydrolysis is different in M to other NB-LRRs. It is also likely to be different again in L6 as the equivalent mutation caused inactivity. It is possible that the third conserved aspartate (D366) is the catalytic base in M given its substitution with alanine resulted in autoactivity. Could there be more than one catalytic amino acid and could different biochemical reactions be performed by each?

Further biochemical investigations are required before functions can be assigned to these residues and it is understood why some NB-LRRs have additional Walker B acidic residues. To do this, improved expression and purification strategies are required to increase flax R protein yields. The hydrolysis experiments presented in this study were hampered by poor yields and this made definitive conclusions difficult to draw.

With the data provided, two scenarios are likely models for M's activation:

1. ATP hydrolysis is required for the activation and activity of M. Mutations that increase hydrolysis rate or the catalytic turnover ($k_{cat}$) cause autoactivity. Mutations that reduce the rate of hydrolysis or $k_{cat}$ decrease the HR intensity, or cause it to be delayed;

2. As in the current model of NB-LRR protein activation, ATP hydrolysis is not required for the activity of M. The active state is the ATP-bound form and this conformation is stabilised and/or triggered upon interaction with AvrM.

Scenario one, that M requires ATP hydrolysis to activate the HR, is supported by evidence from ATPase assays which show that autoactivity was not the result of impaired ATPase activity. This indicates that ATP hydrolysis is unlikely to only be a mechanism of deactivation for M, as it is suggested to be for I-2 and Mi-1. Evidence that the strongly autoactive mutant, $M^{D555V}$, has a higher ATP turnover than wild-type M also supports this possibility (Sornaraj, 2013).

Interestingly, the mechanisms behind R protein autoactivity and effector triggered activity may not be the same. $M^{D555V}$ and $M^{D366A}$ appear to both respond with a more intense HR when co-infiltrated with AvrM compared to when they are infiltrated alone. Could this suggest that different signalling pathways are initiated by autoactive and effector-activated proteins? Whether AvrM triggers an increase in their ATP hydrolysis rate, or influence the nucleotide switch to favour the active ATP-bound state, remains to be determined. Certainly, as reported by Sornaraj (2013) the $M^{D555V}$ protein has a higher maximum ATP hydrolysis rate than wild-type M, and possibly has its ATPase rate increased by AvrM. Efforts to detect an *in vitro* interaction between purified M and AvrM protein were made in this study using co-immunoprecipitation, but were also inconclusive (data not shown). I remain frustrated (like many others) by the lack of effect of the rust effectors on the biochemical properties of the flax R proteins.

### 4.3.8 Conclusions

Extended Walker B motifs have been found to be functional in the STAND proteins including NOD1, NOD2 and NALP12 (Proell *et al.*, 2008); however, from this investigation it seems unlikely that the M protein's extended Walker B region is indispensible for its activity. While it is clear that the first acidic pair of aspartate residues is required for M to activate the HR, the roles played in the molecular mechanism of ATP hydrolysis remain uncertain, but it does seem likely that the hydrolysis of ATP is needed for the M protein to activate the HR. While the activity of M is partially compromised by mutation of the proposed catalytic residue, D364, it is not eliminated. It therefore seems likely that there may be some redundancy in the roles that the individual aspartates play in stimulating the hydrolysis reaction. Alternatively the biochemical reaction performed by the NBS is more complex than first predicted and we do not have the appropriate assay to clearly determine the outcomes of these mutations.

Difficulties in producing recombinant protein of a high enough concentration and purity still dampen attempts to conclusively investigate the molecular mechanism of the ATP hydrolysis reaction. These issues will need to be addressed before significant progress can be made in this area of research. Only then will it be possible to further

our understanding of the role of the ATP hydrolysis reaction in the NB-LRR activation pathway.

As discussed in Chapter 1, STAND proteins that do catalyse the hydrolysis of ATP do so poorly and the hydrolysis rates of the M proteins tested to date are no exception to this. Is hydrolysis of the β-γ phosphate bond of ATP the only function of NB-LRR proteins? It is plausible that more complex reactions are performed by the NBS domain of the M protein and this possibility is explored in the next chapter.

# 5 Other Biochemical Functions of the M Protein

## 5.1 Introduction

There are many classes of proteins that use ATP as a substrate for their biochemical function. Phosphatases use water to remove a phosphonate group from a donor, which can be ATP. Kinases transfer a phosphonate group from a donor (usually ATP) to an acceptor. ATPases catalyse the decomposition of ATP into ADP and a free phosphate ion and the dephosphorylation event releases energy that is used to drive other chemical reactions. R proteins and other proteins of the STAND classification contain many motifs, like the P-loop and Walker B, which classify them as ATPases, but they have also been predicted and/or shown to possess other biochemical functions.

### 5.1.1 ATPase activity

Proteins typically hydrolyse ATP for two main reasons, to provide energy for mechanochemical events, or to provide a phosphate for phosphorylation. The former is believed to be the reason that plant NB-LRRs, and other STAND proteins, possess intrinsic ATPase activity. Energy derived from the hydrolysis of the β-γ phosphoanhydride bond is believed to be utilised as a mechanism of deactivation by affecting conformational rearrangement within the NB-LRR protein and returning the protein to the inactive ADP bound state (Danot *et al.*, 2009; Takken & Tameling, 2009). This activity therefore forms a crucial part of the current molecular switch activation model for NB-LRRs.

To date only four NB-LRRs have been shown to possess ATPase activity: I-2, Mi-1, (Tameling *et al.*, 2002) tobacco N (Ueda *et al.*, 2006) and M (Sornaraj, 2013). The ATPase activities of I-2, Mi-1 and N were all measured with domain truncated proteins and the regulation of these proteins may therefore not reflect that of full-length NB-LRRs. The M protein studied was not full-length either, but only lacked the first 21 N-terminal amino acids and its ATPase activity is thus more likely to accurately represent that of a full-length resting NB-LRR.

### 5.1.2 Nucleotide phosphatase activity

ATPase activity is not the only biochemical activity that has been attributed to a plant R protein. Fenyk *et al.* (2012) recently described a nucleotide phosphatase activity in the NB sub-domain of an orphan rice protein, the NB-ARC regions of the *A. thaliana* CC-NB-LRR RPM1 protein and the Pollen Signalling Protein from maize, which has a similar domain structure to plant NB-LRR proteins. The biochemical action of these proteins is to remove the terminal phosphate from a nucleotide to produce the nucleoside product. This can occur sequentially, i.e. ATP to ADP, ADP to AMP, AMP to Adenosine, but their preference is for the nucleotide monophosphate (AMP). As this action requires functional P-loop and Walker B motifs the cleavage mechanism is thought to be nucleophilic attack on the terminal phosphodiester bond by metal ion-activated water.

Nucleotide phosphatase activity could fit in with the current activation model, but suggests that some R proteins may be able to undergo multiple rounds of nucleotide hydrolysis, rather than the single round that is currently proposed (Williams *et al.*, 2011). The purpose of such activity could be to sustain activity, to multiply the activation signal and it also hints that phosphate release could be crucial to the activation and/or signalling of R proteins.

### 5.1.3 Phosphorylation-mediated signalling

In this study by Fenyk *et al.* (2012) no mention is made regarding the fate of the phosphate groups that would be produced by these sequential cleavage events. A much earlier bioinformatics study has suggested that some R proteins may possess a region within their NBS domain that is structurally homologous to the CheY family of response regulator receiver domains of His-Asp phosphotransfer signalling pathways (Rigden *et al*., 2000). Whilst this prediction was made before nucleotide binding had been demonstrated by any R protein, this proposal could hint at the involvement of R proteins in more sophisticated phosphorylation-mediated mechanisms of self-regulation, activation and signalling than has previously been thought.

These pathways, also known as two-component signalling or phosphorelay systems, involve a sensor histidine kinase and a response regulator that contains the receiver

domain. They can also employ hybrid histidine kinases that contain both histidine kinase and receiver domains and also proteins with a phosphotransfer domain (Schaller *et al.*, 2011; West & Stock, 2001). The pathway is activated when, in response to a signal, a sensor histidine kinase catalyses ATP-dependent autophosphorylation of a conserved histidine residue in a transmitter domain. The response regulator then catalyses transfer of the phosphoryl group to a conserved aspartate in its receiver domain (West & Stock, 2001). Phosphorylation of the conserved response receiver aspartate domain is translated into an outcome through an associated output domain. These output domains include DNA-binding motifs that regulate transcription of target genes, serine/threonine kinases, phosphatases and protein-protein interaction domains (Schaller *et al.*, 2011). Rigden *et al.* (2000) suggest that the phosphorylated aspartate residue in the CheY receiver domain is analogous to the first conserved aspartate of the Walker B motif of NB-LRRs.

Interestingly, a large-scale phosoproteomics study carried out by Nakagami *et al.,* (2010) identified phosphorylated residues in functionally significant NB-LRR motifs. Perhaps most interesting in the context of this study was the discovery of two phosphorylated residues in the NB-ARC domain. The authors suggest that the conservation of NB-LRR phosphopeptides within different plant species possibly points to a conservation of phosphorylation-mediated regulation of NB-LRRs across diverse species (Nakagami *et al.*, 2010). While phospho-histidine or phospho-aspartates were not identified in this study, they were also not targeted, and thus maybe as yet unidentified within NB-LRRs. Further research is clearly warranted to explore if they exist.

### 5.1.4 Aims

We know that M possesses the ability to hydrolyse ATP and that it can do so in the absence of AvrM, but we do not know the purpose of this activity; is it important for activation or deactivation, both or neither? While an exact role for ATP hydrolysis in M protein activation and activity remains to be defined, other biochemical activities are worthy of investigation.

To date, the fate of the cleaved phosphoryl group of the hydrolysed bound ATP molecule with an NB-LRR R protein has not been explored and this study focused on investigating an expanded function of the M protein. The results presented in this chapter represent the first attempt to uncover if only the γ-phosphate is cleaved and if the reaction is a simple hydrolysis step, or if further catalysis of ADP occurs, or if a phosphorelay system is in place, either or both of which may be important in downstream signalling of the resistance response. Whether or not M undergoes autophosphorylation with the phosphoryl product of the ATP hydrolysis reaction was also explored.

## 5.2 Results

With the aim of observing phosphorylation, M and mutant M proteins were purified using the tandem protocol of nickel affinity and *Strep*Tactin® chromatography that is used when measuring ATPase activity (2.7.5 & 2.7.6). Each phosphorylation assay contained 0.4 µCi [γ-$^{32}$P]ATP or [α-$^{32}$P]ATP in 5 µM ATP and 250 ng of protein in hydrolysis assay buffer A (50 mM Tris ($Cl^-$) pH 7.5, 50 mM NaCl, 10 mM $MgCl_2$, 10% (w/v) glycerol) in a total reaction volume of 50 µL and was carried out at 25$^o$C. M protein used in the assays had been stored at -80$^o$C since its purification and their concentrations were determined with a total protein assay (2.7.11).

At appropriate time points during each assay, aliquots were taken for analysis by TLC and SDS-PAGE. Assays were typically performed for one or four hours and samples were taken after 30 minutes or at one hour intervals. Samples to be separated by SDS-PAGE were mixed with 3X sample buffer containing 100 mM DTT and immediately frozen. Samples spotted onto TLC plates were resolved with a solution of 0.25 M LiCl and 2.5% (v/v) formic acid, making sure not to run the buffer to the end of the TLC plate. Myosin (~ 8 µg; calcium activated from rabbit muscle) was used as a positive control for γ-phosphate cleavage in all assays (2.7.10).

Proteins within the samples were separated by 8, 10 or 4-15% SDS-PAGE. Once run, the gels were exposed to a PhosphorImager plate for an appropriate period of time (typically 24 hours), after which the plate was scanned in a Bio-Rad Molecular Imager to identify any radiolabelled areas in the gel. Gels were subsequently stained with SyproRuby®, as per the manufacturer's instructions, to visualise protein bands. The resolved TLC plates were also exposed to the PhosphorImager plate to visualise the products formed throughout the course of the assay.

### 5.2.1 A potential phosphorylation phenomenon is time and M protein dependent

Separate assays containing M protein and Myosin with [γ-$^{32}$P]ATP were performed to determine if any $^{32}$P signal could be seen associated with protein after 10% SDS-PAGE analysis. A scan of the PhosphorImager plate after it was exposed to the resolved TLC plate containing samples from each reaction showed that free $^{32}$P was produced in

the reactions containing the positive control for γ-phosphate cleavage, Mysoin and M protein. Each reaction also still contained some [γ-$^{32}$P]ATP (Figure 5-1). A mystery spot was also observed on the TLC plate that had run further from the origin than the $^{32}$P spot.

As seen in Figure 5-2, over the course of the one hour assay the $^{32}$P signal increased in the reaction containing M protein, but not in the reaction containing Myosin. A low $^{32}$P signal was observed in the buffer only sample taken at 60 minutes and the TLC plate showed that some free $^{32}$P was formed in this reaction (Figure 5-1). This signal was less than that observed from the M protein reaction.

Generation of this $^{32}$P signal was M protein dependent, but as illustrated by SDS-PAGE in Figure 5-2, the $^{32}$P signal was not associated with the near-full length M protein. Rather than co-migrating with the ~ 150 kDa M proteins band the signal accumulated at the interface between the stacking and resolving acrylamide layers.

Saturated spots were obtained on the phosphoimages due to the lengthy time of exposure of TLC plates to the PhosphorImager and it was therefore not possible to accurately determine a rate of ATP hydrolysis from the M and M$^{K286L}$ reactions. A comparison of the rates of hydrolysis achieved with [γ-$^{32}$P]ATP and those achieved with [α-$^{32}$P]ATP (Sornaraj, 2013) was therefore not made.

### 5.2.2 The potential phosphorylation phenomenon is reliant on a functional nucleotide binding pocket

Reactions were then performed to compare the $^{32}$P signal produced by the M protein with any produced by the inactive mutant M$^{K286L}$. As M$^{K286L}$ retains some ability to bind nucleotides and hydrolyse ATP (Sornaraj, 2013; Williams *et al.*, 2011) it was not unexpected to observed some $^{32}$P in the M$^{K286L}$ samples when they were separated by TLC and 10% SDS-PAGE (Figure 5-3 and Figure 5-4). The intensity of the observed in-gel $^{32}$P labelling phenomenon was visibly reduced in the reactions containing the M$^{K286L}$ protein (Figure 5-4). The same mystery spot as observed in the first M reaction was also observed on the TLC plate of this reaction.

It is worth noting that although the $^{32}$P signal on the TLC plate of these reactions (Figure 5-3) showed that the intensity of the free $^{32}$P and the mystery spots were

higher in the samples from the assay with $M^{K286L}$ than in the samples from the assay containing M protein. However more $M^{K286L}$ than M protein was used in this assay. This can be visualised in Figure 5-4 by the more intense staining of the $M^{K286L}$ band in the samples taken at time zero from the reactions containing M and $M^{K286L}$.

As the $^{32}P$ signals in the acrylamide gels had been concentrated around the interface between the 4% stacking gel and the 10% resolving gel the M and $M^{K286L}$ samples were re-run with 8% SDS-PAGE. The stacking gel was also kept on the gel during its exposure to the PhosphorImager plate. It can be observed in Figure 5-5 that some of the $^{32}P$ signal remained in the wells of the 4% stacking gel and some signal also remained at the interface of the 4 and 8% acrylamide layers. In the M samples taken at 2.5 and 4 hours it appears as though some signal is concentrated in distinct bands that have begun to move into the 10% resolving gel. A similar pattern may also be present in the $M^{K286L}$ samples from 2.5 and 4 hours, but the signal strength was not amenable for its visualisation and quantification.

### 5.2.3 Near-full length M protein disappears during the assay

As evidenced by the slow disappearance of M and $M^{K286L}$ bands in the SyproRuby stained acrylamide gels, M protein appears to break down during the assay coincident with the increase in the $^{32}P$ signal (Figure 5-2 and Figure 5-4). An assay was subsequently performed without [γ-$^{32}P$]ATP to determine whether this is due to the conditions in the assay, or if it is an indication of M protein shifting to a higher molecular weight size that could be co-migrating with the $^{32}P$ signal. As shown in the lanes where M was not incubated with ATP (Figure 5-6), M protein was lost over the course of the assay in the absence of ATP. Unfortunately, quantitation of this loss was not possible in these reactions and it was impossible to determine if this loss was associated with a shift in mobility.

### 5.2.4 M likely only cleaves the β-γ phosphoanhydride bond of ATP

If M has nucleotide phosphatase activity it should be able to sequentially cleave phosphates from tri-, di- and mono-nucleotides. It consequently should be able to release the α-phosphate from ATP and potentially, the possible phosphorylation signal should be observed in reactions performed with [α-$^{32}P$]ATP as it had been with

[γ-$^{32}$P]ATP. To determine whether this was the case, M and M$^{K286L}$ proteins were incubated in separate reactions with [γ-$^{32}$P]ATP and [α-$^{32}$P]ATP. A free $^{32}$P signal was observed in the reactions that were performed with [γ-$^{32}$P]ATP (Figure 5-6), but not in reactions with an [α-$^{32}$P]ATP substrate (Figure 5-7).

Interestingly, a low $^{32}$P signal was observed in the acrylamide gel from the reaction that contained [γ-$^{32}$P]ATP without any protein (Figure 5-2). A similar signal was not detectable in the gel containing samples of the [α-$^{32}$P]ATP only reaction (Figure 5-7).

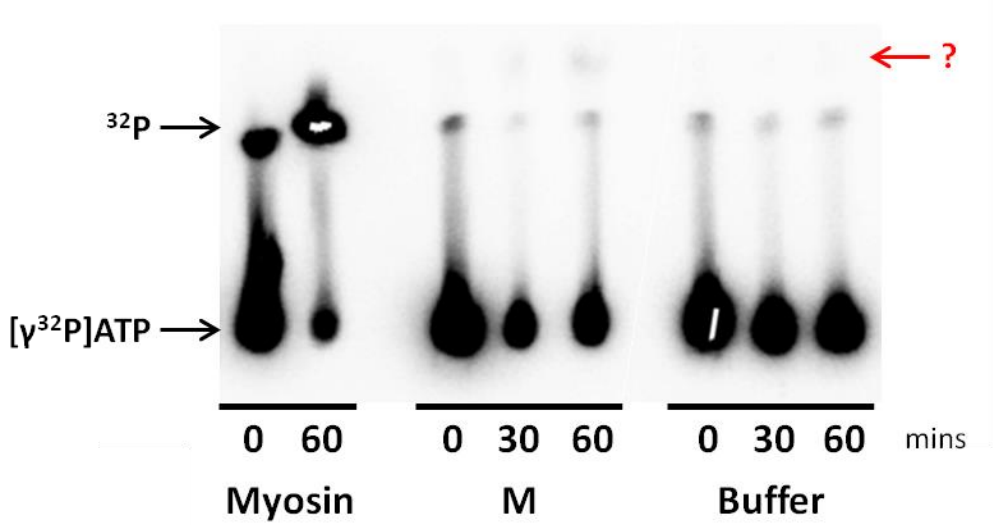
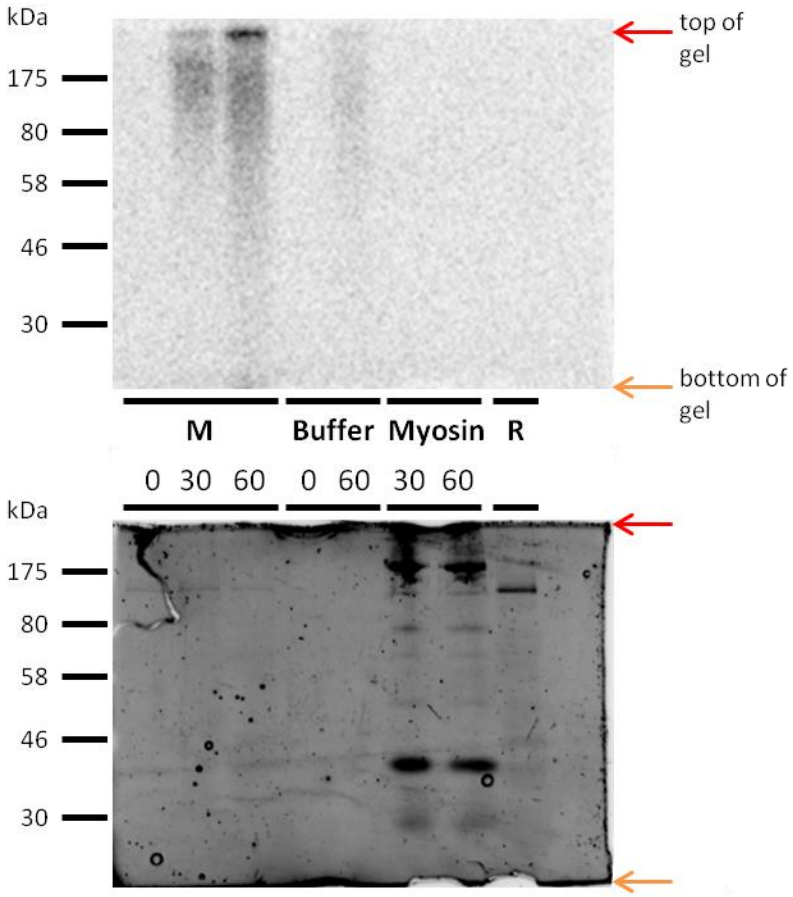



**Figure 5-1: The phospho-image of a resolved TLC plate showing the reaction time course from M, Myosin and buffer only [γ-$^{32}$P]ATP assays**

Each assay contained 0.4 µCi [γ-$^{32}$P]ATP in 5 µM ATP in hydrolysis assay buffer A (50 mM Tris ($Cl^-$) pH 7.5, 50 mM NaCl, 10 mM $MgCl_2$, 10% (w/v) glycerol) and 250 ng of M protein purified by nickel affinity and *Strep*Tactin® chromatography, 8 µg of Myosin or buffer only in a total volume of 50 µL. The assays were performed at 25°C for one hour and two microlitre samples of each reaction were spotted onto the TLC plate after 0, 30 and 60 minutes (0 and 60 for Myosin). Once dry, the plates were resolved with a solution of 0.25 M LiCl and 2.5% (v/v) formic acid. The TLC plate was then exposed to a PhosphorImager plate for an appropriate period of time before being scanned in a Bio-Rad Molecular Imager to detect radiolabelled areas.

All reactions generated some free $^{32}$P, although not all of the [γ-$^{32}$P]ATP was converted into free $^{32}$P during the course of the assay. A mystery spot (see red arrow) that had run further from the origin than the $^{32}$P spot was also observed on the TLC plate in samples from the assay containing M protein, and to a lesser extent in the buffer only samples. The spot was not observed in the samples from the reaction containing the positive control for β-γ bond hydrolysis, Myosin.




R = M protein input for size reference; 0, 30, 60 = minutes

**Figure 5-2: The phospho-image (top) of a 10% acrylamide gel (bottom) containing samples taken from [γ-$^{32}$P]ATP assays of M, Myosin and buffer only**

During the course of each assay 10 µL samples were taken after 0, 30 and 60 minutes (0 and 60 only for buffer only and Myosin) and immediately mixed with 3X sample buffer containing 100 mM DTT before being frozen. Proteins within these samples were then separated via 10% SDS-PAGE. M protein as used in the assay was run in the gel as a size reference and as an unlabelled control (lane R). Once run, the gel was exposed to a PhosphorImager plate for an appropriate period of time, after which the plate was scanned in a Bio-Rad Molecular Imager to identify any radiolabelled areas in the gel (top image). Gels were subsequently stained with the total protein SyproRuby® stain to visualise protein bands (bottom image).

Over the course of the one hour assay $^{32}$P signal increased in the reaction containing M protein, but not in the reaction with Myosin. A low $^{32}$P signal was also observed in the buffer only sample taken after 60 minutes. The $^{32}$P signal was not associated with M protein (band of ~150 kDa); rather it accumulated at the top of the resolving gel, the interface between the 4 and 10% acrylamide layers. The M protein band (~150 kDa) disappeared over the course of the reaction, whilst the Myosin bands heavy chain ~200 kDa) did not.

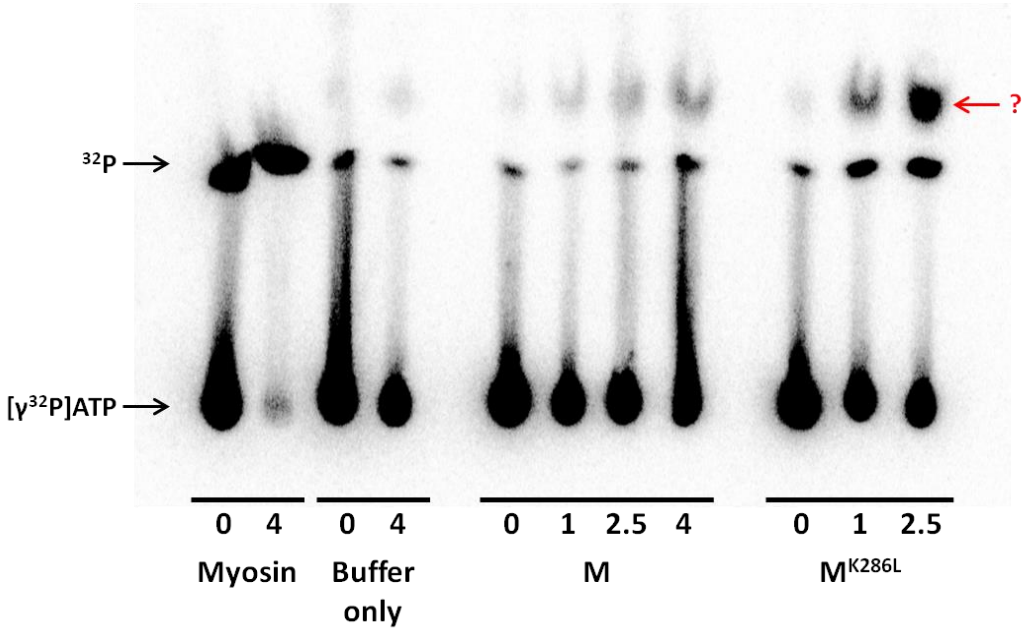
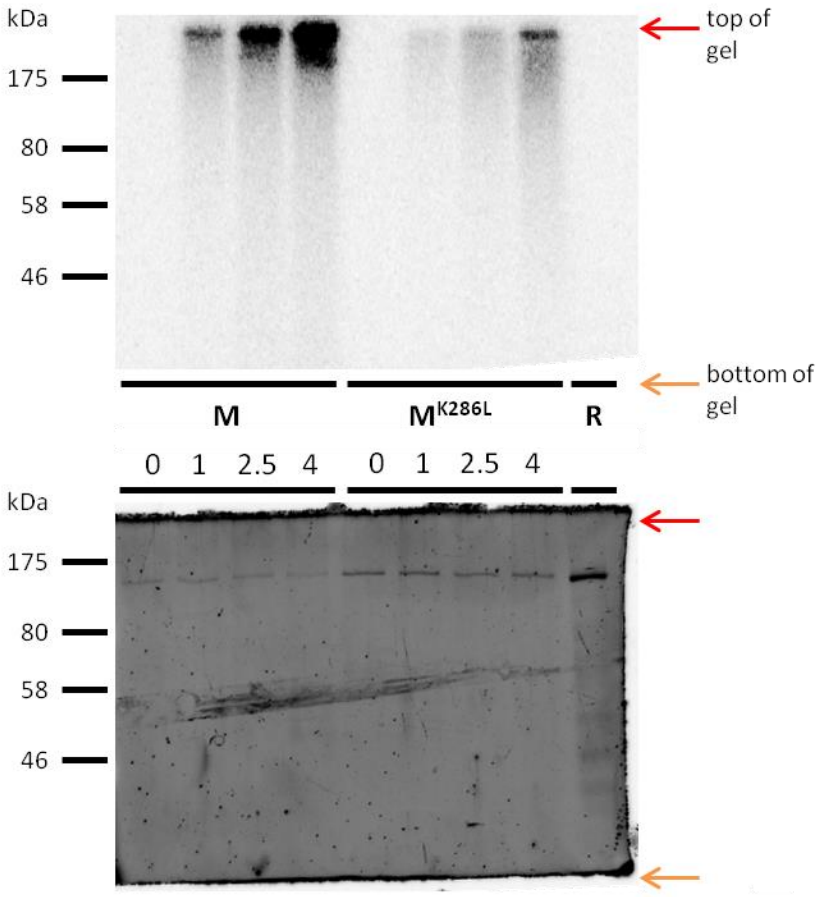



**Figure 5-3: A phospho-image of a resolved TLC plate showing the reaction time course from M, $M^{K286L}$, Myosin and buffer only [γ-$^{32}$P]ATP assays**

Each assay contained 0.4 µCi [γ-$^{32}$P]ATP in 5 µM ATP in hydrolysis assay buffer A (50 mM Tris ($Cl^-$) pH 7.5, 50 mM NaCl, 10 mM MgCl2, 10% (w/v) glycerol) and 250 ng of M protein purified by nickel affinity and *Strep*Tactin® chromatography, 8 µg of Myosin or buffer only in a total volume of 50 µL. The assays were performed at 25°C for one hour and two microlitre samples of each reaction were spotted onto the TLC plate after 0, 1, 2.5 and 4 hours (0 and 4 only for Myosin and buffer only). Once dry, the plates were resolved with a solution of 0.25 M LiCl and 2.5% (v/v) formic acid. The TLC plate was then exposed to a PhosphorImager plate before being scanned in a Bio-Rad Molecular Imager to detect radiolabelled areas.

All reactions generated some free $^{32}$P and all reactions had some [γ-$^{32}$P]ATP remaining at the end of the assay, although Myosin had almost converted it all to free $^{32}$P. The mystery spot seen on the earlier TLC plate was again observed in samples that contained M protein and was also observed in the reactions containing $M^{K286L}$, and to a lesser extent buffer only. The spot was again not observed in the samples from the reaction containing Myosin.

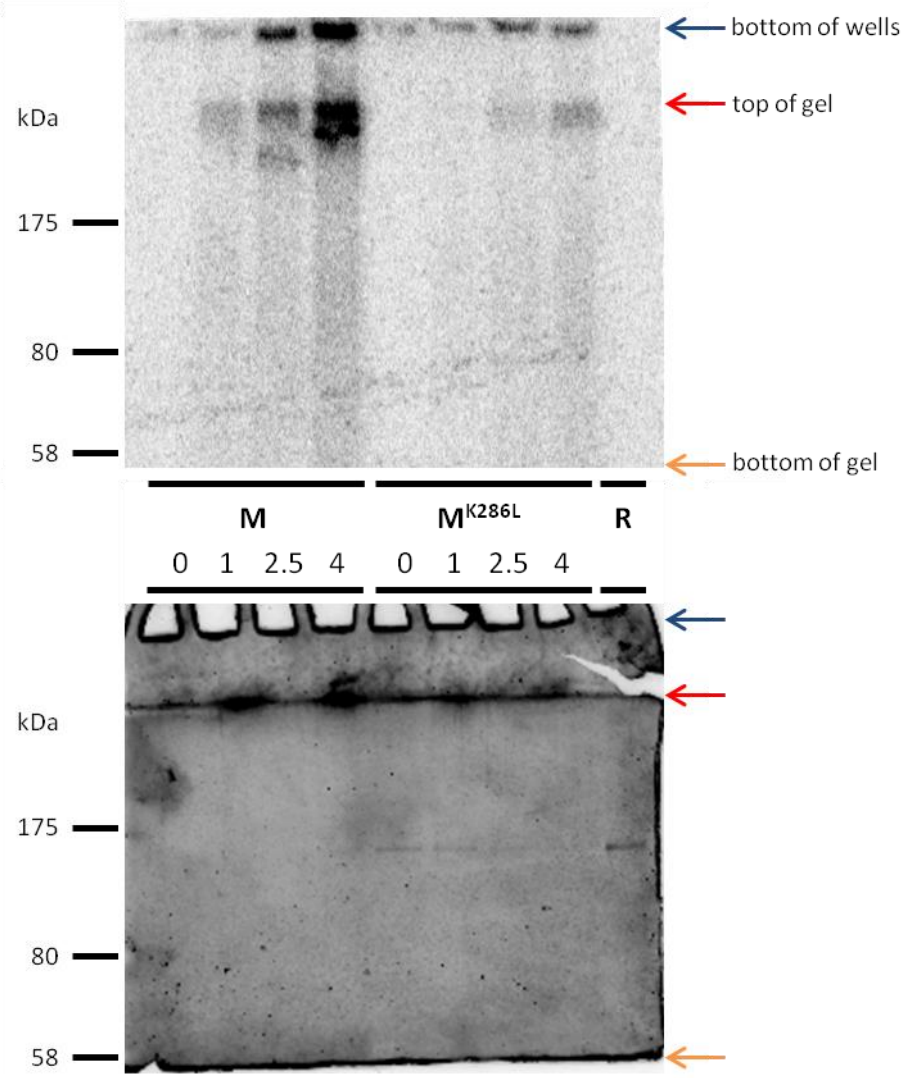



Protein ladder = NEB Prestained broad range; R = M protein input for size reference; 0, 1, 2.5, 4 = hours

**Figure 5-4: The phospho-image (top) of a 10% acrylamide gel (bottom) containing samples taken from [γ-$^{32}$P]ATP assays with M and M$^{K286L}$**

During the course of each assay 10 µL samples were taken after 0, 1, 2.5 and 4 hours and immediately mixed with 3X sample buffer containing 100 mM DTT before being frozen. Proteins within these samples were then separated via 10% SDS-PAGE. M protein as used in the assay was run in the gel as a size reference and as an unlabelled control (lane R). Once run, the gel was exposed to a PhosphorImager plate, after which the plate was scanned in a Bio-Rad Molecular Imager to identify any radiolabelled areas in the gel (top image). Gels were subsequently stained with the total protein SyproRuby® stain to visualise protein bands (bottom image).

Over the course of the four hour assay $^{32}$P signal increased in the reaction containing M protein, and to a lesser extent in the M$^{K286L}$ reaction. Again the $^{32}$P signal was not associated with M protein and actually accumulated at the top of the resolving gel. As observed previously, the M protein band (~150 kDa) disappeared over the course of the reaction. The same was also observed of M$^{K286L}$ band. As illustrated by its more intense staining, more M$^{K286L}$ than M protein was used in this assay.




Protein ladder = NEB Prestained broad range; R = M protein input for size reference; 0, 1, 2.5, 4 = hours

**Figure 5-5: The phospho-image (top) of an 8% acrylamide gel (bottom) containing samples taken from [γ-$^{32}$P]ATP assays with M and M$^{K286L}$**

During the course of each assay 10 µL samples were taken after 0, 1, 2.5 and 4 hours and immediately mixed with 3X sample buffer containing 100 mM DTT before being frozen. Proteins within these samples were then separated via 8% SDS-PAGE. M protein as used in the assay was run in the gel as a size reference and as an unlabelled control (lane R). Once run, the stacking gel was not removed from the gel and it was exposed to a PhosphorImager plate, after which the plate was scanned in a Bio-Rad Molecular Imager to identify any radiolabelled areas in the gel (top image). Gels were subsequently stained with the total protein SyproRuby® stain to visualise protein bands (bottom image).

Some of the $^{32}$P signal remained in the wells of the 4% stacking gel and some signal also remained at the interface of the 4 and 8% acrylamide layers. However, in the M samples taken after 2.5 and 4 hours it appears as though some signal was concentrated in distinct bands that have begun to move into the 8% resolving gel. A similar pattern may also be present in the M$^{K286L}$ samples from 2.5 and 4 hours, but the $^{32}$P signal strength was not amenable to its visualisation.

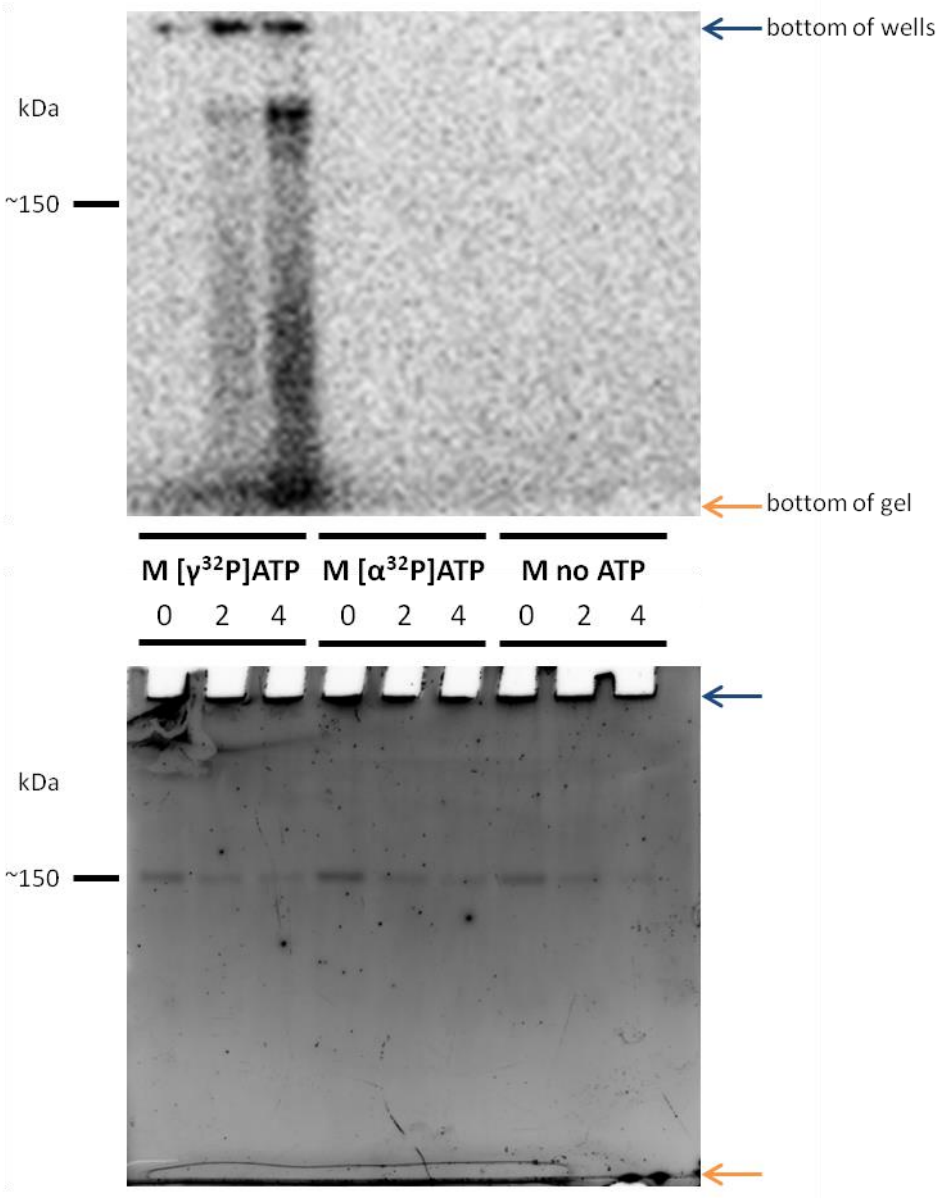

bottom of wells
kDa
~150
bottom of gel
M [γ32P]ATP
M [α32P]ATP
M no ATP
0 2 4 0 2 4 0 2 4
kDa
~150

**Figure 5-6: The phospho-image (top) of a 4-15% acrylamide gel (bottom) containing samples taken from assays of M with [γ-$^{32}$P]ATP and [α-$^{32}$P]ATP substrates and without ATP**

During the course of each assay 10 µL samples were taken after 0, 2 and 4 hours and immediately mixed with 3X sample buffer containing 100 mM DTT before being frozen. Proteins within these samples were then separated via 4-15% SDS-PAGE. Once run, the gel was exposed to a PhosphorImager plate, after which the plate was scanned in a Bio-Rad Molecular Imager to identify any radiolabelled areas in the gel (top image). Gels were subsequently stained with the total protein SyproRuby® stain to visualise protein bands (bottom image).

Assays contained M protein and [γ-$^{32}$P]ATP, M protein and [α-$^{32}$P]ATP and also M protein without any ATP. Increasing $^{32}$P signal was only observed in the gradient gel in samples from the reaction containing M protein and [γ-$^{32}$P]ATP, not from M with [α-$^{32}$P]ATP. M protein was lost over the course of the assay in the reactions that did not contain ATP.

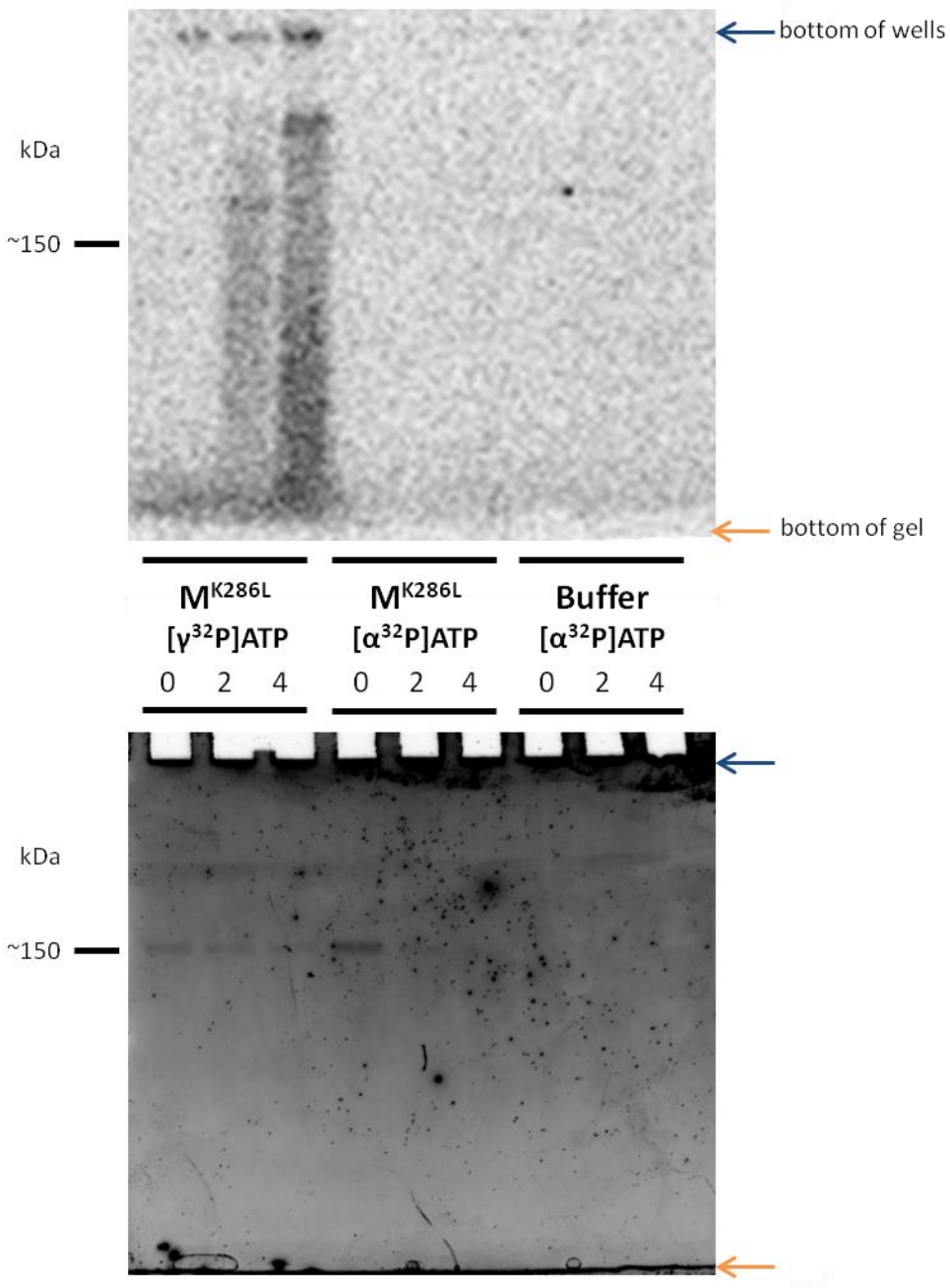

bottom of wells
kDa
~150
bottom of gel
M^K286L
[γ32P]ATP
M^K286L
[α32P]ATP
Buffer
[α32P]ATP
0 2 4
0 2 4
0 2 4
kDa
~150

**Figure 5-7: The phospho-image (top) of a 4-15% acrylamide gel (bottom) containing samples taken from assays of $M^{K286L}$ with [γ-$^{32}$P]ATP and [α-$^{32}$P]ATP substrates**

During the course of each assay 10 µL samples were taken after 0, 2 and 4 hours and immediately mixed with 3X sample buffer containing 100 mM DTT before being frozen. Proteins within these samples were then separated via 4-15% SDS-PAGE. Once run, the gel was exposed to a PhosphorImager plate, after which the plate was scanned in a Bio-Rad Molecular Imager to identify any radiolabelled areas in the gel (top image). Gels were subsequently stained with the total protein stain SyproRuby® to visualise protein bands (bottom image).

Assays contained $M^{K286L}$ protein and [γ-$^{32}$P]ATP, $M^{K286L}$ protein and [α-$^{32}$P]ATP and only [α-$^{32}$P]ATP. Increasing $^{32}$P signal was only observed in the gradient gel in samples from the reaction containing $M^{K286L}$ protein and [γ-$^{32}$P]ATP. The signal from the $M^{K286L}$ reaction was also lower than was observed from the corresponding M protein assays (Figure 5-6). No signal was observed in the gel from samples of the assay of [α-$^{32}$P]ATP without any protein.

## 5.3 Discussion

M protein is capable of hydrolysing ATP but the fate of the cleaved phosphoryl group is not yet known. Could this compound go on to phosphorylate an amino acid of M and be involved in its activation and/or signalling? This chapter sought to investigate biochemical properties additional to that of ATP hydrolysis of the M protein's nucleotide binding pocket by looking at the fate of the released inorganic phosphate. This was accomplished by incubating M proteins with $\gamma^{32}$P- and $\alpha^{32}$P- labelled ATP and analysing the products of the reactions by TLC and SDS-PAGE.

Free $^{32}$P, indicative of γ-phosphate release, was generated in reactions containing M proteins and a slow moving $^{32}$P labelled molecule, not consistent with the monomeric size of M, was also observed after SDS-PAGE in reactions containing M proteins. This $^{32}$P-labelled, in-gel phenomenon was never observed in reactions containing Myosin, but was reduced when nucleotide binding of the M protein was impaired. This was best illustrated by the increasing intensity of the $^{32}$P signal over the course of M protein assays and the reduced signal intensity from the reaction with the inactive $M^{K286L}$ mutant that has impaired nucleotide binding abilities (Figure 5-2 and Figure 5-3). Some residual activity from the $M^{K286L}$ mutant was not unexpected given that some nucleotide binding and ATP hydrolysis activity is observed from this mutant (Sornaraj, 2013; Williams, 2009; Williams *et al.*, 2011).

The slow migrating $^{32}$P-labelled signal could represent covalent attachment of the $^{32}$P molecule to a protein by phosphorylation. The signal could also be indicative of phosphorylated high molecular weight M protein complex, but this complex would need to be resistant to the denaturing conditions of the SDS-PAGE. Alternatively it could also represent phosphorylation of an unknown *P. pastoris* derived protein present after purification. No protein band consistent with the slow migrating $^{32}$P-labelled signal was observed in total protein stains of the gels, and thus if this phenomenon is protein phosphorylation any protein is below the threshold of detection of the SyproRuby® stain. Alternatively, the cleaved phosphate may not be associated with any protein; however, as Myosin reactions did not show any labelling, it is not free $^{32}$P, but a form that migrates slowly in SDS-PAGE. Future experiments should include an empty vector control or expression and purification a negative

control protein to firmly establish that this phenomenon is M protein dependent. Running samples on native- and SDS-PAGE may also help to determine whether this signal is in fact covalently attached to a protein, which will help to elucidate its identity. Also protease treatment prior to SDS-PAGE will determine if the $^{32}P$ signal is associated with protein.

The $^{32}P$ signal on the TLC plate of the $M^{K286L}$ and M [$\gamma^{32}P$]ATP reactions showed that the intensity of the free $^{32}P$ and the mystery spots were higher in the samples from the assay with $M^{K286L}$ than in the samples from the assay containing M protein. Despite this, the $^{32}P$ signals in the SDS-PAGE were much more intense in the samples from the M reaction than from those from the $M^{K286L}$ reaction. This is likely the result of more $M^{K286L}$ protein than M protein being added to the assay, as observed in the SyproRuby® stained SDS-PAGE. An alternative explanation is that the reduced biochemical activity of $M^{K286L}$ also impairs the ability of this protein to become stably autophosphorylated or to phosphorylate other proteins; therefore a greater amount of product is observed on the TLC plate because it is not associated with the protein itself. This is however a very speculative conclusion to draw and would require identification and quantification of autophosphorylation sites on any protein involved to confirm.

Some degradation of [$\gamma$-$^{32}P$]ATP was apparent in all TLC plates which contained [$\gamma$-$^{32}P$]ATP only reactions. $^{32}P$ labelling was subsequently observed in the acrylamide gels containing samples of these reactions. However, the signal intensity was much lower than that observed from reactions containing M protein, and even $M^{K286L}$ protein, and was therefore attributed to background breakdown of [$\gamma$-$^{32}P$]ATP in the reactions which are carried out at 25$^{o}$C. Breakdown of the [$\gamma$-$^{32}P$]ATP could also have occurred during the heating process that was used to prepare the protein/sample buffer mixture for SDS-PAGE; however, no $^{32}P$ signal was observed in any gel containing a sample from a Myosin reaction. This result could be attributed to the fact that Myosin was generally hydrolysed ATP faster than any of the M proteins and there was therefore little [$\gamma$-$^{32}P$]ATP left in the assays to degrade, but no background signal was observed in the [$\alpha$-$^{32}P$]ATP only reaction or from the assays of M and $M^{K286L}$ with [$\alpha$-$^{32}P$]ATP. The background $^{32}P$ signal could also be attributed to a minor contaminating

phosphatase. In the future, as a precaution, sodium fluoride and sodium orthovanadate could be added to the control reactions to inhibit any contaminating phosphatase (Kovács *et al.*, 2003). Whatever the cause, as the background signal was lower than any signal attributed to an M protein it was deemed inconsequential and it can be assumed that conversion of [γ-$^{32}$P]ATP into any of the products assayed by TLC or SDS-PAGE can be attributed to the function of the M protein.

All M protein bands on SDS-PAGE became less intense over the course of each assay and this was co-incident with the increasing intensity of the slower migrating $^{32}$P in-gel signal. It is possible that near-full length M proteins were being phosphorylated, shifting to a higher molecular weight and thus explaining this result. Such a shift in mobility as a consequence of phosphorylation has been observed previously for the CREB-binding protein (Kovács *et al.*, 2003). However, M degradation occurred in the absence of ATP and was therefore likely caused by the conditions of the assay and not a phosphorylation-mediated mobility shift. Unfortunately, quantitation of this loss was not possible in these reactions and it was impossible to determine if this loss was associated with a shift in mobility, but any suggestion that M protein loss (or mobility shift) over time is associated with incubation with ATP should be taken with some caution. However, this possibility cannot be completely ruled out and should be further investigated through identification of any proteins that co-migrate with the $^{32}$P signal bands observed in some of the acrylamide gels. This could be achieved with mass spectrometry (MS) techniques similar to those used by Nakagami *et al.* (2010) in this identification of phosphopeptides in the *A. thaliana* and rice genomes. Calf intestinal alkaline phosphatase (CIP) treatment of potentially phosphorylated protein samples could also be used to determine if the phenomenon is protein associated and to reverse any phosphorylation-mediated mobility shift (Kovács *et al.*, 2003). This signal in SDS-PAGE could also suggest that what is being observed in the M protein containing reactions is not proteinaceous. Such a possibility could be investigated by treating the reactions with a protease, like proteinase K, prior to SDS-PAGE. The presence of a signal after such treatment would rule out proteinaceous phosphorylation.

### 5.3.1 M is unlikely a nucleotide phosphatase

No $^{32}P$ signal was observed after SDS-PAGE analysis of samples taken from M or $M^{K286L}$ reactions with [$\alpha^{32}P$]ATP. This indicates that it is unlikely that M is capable of hydrolysing an adenosine monophosphate (AMP) substrate. Additionally this observation provides evidence that M cannot sequentially cleave terminal phosphates from nucleotide substrates to generate a nucleoside product and is therefore unlikely to be a nucleotide phosphatase. However, it does not preclude the possibility that the α-β phosphoanhydride bond is not being cleaved and gerenating the observed signal. Reactions should be performed with more radiolabelled adenosine substrates, such as the 2,8-$^{3}H$-labelled adenine nucleotide used by Fenyk *et al.* (2012) or [$\beta^{32}P$] labelled ATP, to furher investigate the substrates that M can ustilise. The use of such substrates would allow more and different products of the reaction to be visualised on TLC plates, with the former allowing visualisation of all of the hydrolysis products, right through to adenosine.

### 5.3.2 Identification of the mystery spot

A mystery $^{32}P$-labelled spot was identified on resolved TLC plates from reactions containing M protein. This spot travelled further from the origin than free $^{32}P$, as identified by comparison with the free $^{32}P$ generated in Myosin reactions. Interestingly this spot was never observed in reactions containing Myosin, the protein used as a positive control for ATP hydrolysis. The identity of this spot is unknown and in the future may be identified using MS techniques and/or nucleotide and phosphate standards run alongside the products from [$\gamma^{32}P$]ATP reactions during TLC analysis. It could represent an unknown intermediate of the hydrolysis reaction and its presence hints at a function for the M protein other than that of standard γ-phosphate ATP hydrolysis.

### 5.3.3 Potential phosphorylation sites on M

Assuming that the labelled product observed by SDS-APGE represents phosphorylated M, this calls into question whether a simple ATP hydrolysis reaction is performed by the NBS domain. Although M appears not to be a nucleotide phosphatase, phosphorylation-mediated signalling may still be a feature of the

activity of pre- and/or post-activated M protein. As suggested by Ridgen *et al.* (2000) it is possible that NB-LRR proteins, and therefore M, could form part of a His-Asp phosphorelay pathway. This would mean that critical functional residues like the histidine of the MHD motif and the aspartates of the extended Walker B motif could be transiently phosphorylated as part of a two-component signalling pathway. Does the histidine or one, or all, of the Walker B aspartates get phosphorylated to induce a conformational change within M? Could such a modification control the switch between inactive and active states, rather than nucleotide exchange?

Phosphorylation is known to control conformational changes within NB-LRR proteins. For example, mNLRC4 is phosphorylated at a structurally conserved serine and this modification is critical for assembly of the mNLRC4 inflammasome as it stabilises inter-domain interactions (Hu *et al.*, 2013b; Qu *et al.*, 2012). The receiver domains of response regulator proteins are also subject to phosphorylation-mediated conformational changes when they catalyse the phosphorylation of one of their own aspartate residues after an interaction with a phosphorylated histidine kinase (West & Stock, 2001). This makes two-component systems phosphorylation-mediated switches that activate transcriptional changes (Schaller *et al.*, 2011). Interestingly, the receiver domains of response regulator proteins can also catalyse the transfer phosphoryl groups from simple molecules, like phosphoramidate ($NH_2PO_3^{2-}$) and acetyl phosphate, *in vitro* (Lukat *et al.*, 1992). It would be worth determining whether M can also perform protein phosphorylation by using such molecules as alternative phospho-donor substrates to adenosine nucleotides in hydrolysis assays.

Nakagami *et al.* (2010) published a phosphoproteomics study in which multiple sites of phosphorylation were identified on a CC-and TIR-NB-LRR proteins from rice and *A. thaliana* (Nakagami *et al.*, 2010). Whilst some of these phosphorylated residues are not well conserved among the NB-LRR proteins, some are conserved among CC- and TIR-NB-LRRs from different plant species. One such phosphopeptide is a threonine residue of the P-loop motif (Nakagami *et al.*, 2010), another is a serine within the RNBS-B region. Rigden *et al.* (2000) also suggest the first conserved aspartate of the Walker B motif of NB-LRRs could be phosphorylated due to the similarity of this region with the CheY receiver domain. These examples of phosphorylation and the

potential for phosphorylation in NB-LRR proteins hint that such a biochemical process could be occurring in the M protein.

Through analogy, the histidine of the MHD motif is thought to interact with the β-phosphate group of the bound nucleotide (Riedl *et al.*, 2005). This residue is also thought to be the Sensor-2 residue in plant NB-LRR proteins and this motif, together with the WHD (or ARC2 sub-domain), has long been hypothesised to transmit the free energy of ATP hydrolysis to the N-terminal regions (signalling domains) of an R protein (Iyer *et al.*, 2004; Ogura and Wilkinson, 2001). The region is known to undergo conformational changes in Apaf-1 and mNLRC4 proteins after their oligomerisation and thus plays a critical role in their activation (Hu *et al.*, 2013b). If this histidine residue were to be phosphorylated as a result of ATP hydrolysis, the phospho-histidine could trigger a conformational change to initiate R protein activation. This could be achieved by release of the N-terminal domain or release of negative regulation imparted by the LRR and this ARC2 domain. It will be interesting to determine whether this histidine residue is being phosphorylated in M.

Alternatively, phosphorylation could play an essential role in maintaining M's inactive, autoinhibited state when its cognate effector is not present. Associated phosphatases/kinases could play important roles in regulation by removing and/or adding phosphoyl groups to regulate R protein activity in the presence or absence of the pathogen. Such a signalling pathway could have another role when the plant isn't being challenged by a pathogen, perhaps in regulating cell homeostasis like Apaf-1 (Kufer *et al.*, 2011). Could the compartmentalisation of cell death and pathogen response pathways help us understand signalling (Bai *et al.*, 2012)? Changes in phosphorylation could be a trigger for cell death/pathogen response pathways instead of normal homeostasis signalling? In fact, a phosphatase and a histidine kinase have been found to be associated with some plant NB-LRRs (de la Fuente van Bentem *et al.*, 2005; Kang *et al.*, 2008). CRT1 (compromise recognition of turnip crinkle virus) interacts specifically with the NB-ARC domain and has been found associated with SSI4, RPS2 and Rx NB-LRRs (Kang *et al.*, 2008). CRT1 is a member of the GHKL (gyrase, Hsp90, histidine kinase, MutL) ATPase/kinases superfamily and is distantly related to the chaperone Hsp90 (Kang *et al.*, 2008). These associations may

provide a link to the involvement of NB-LRRs in phosphorelay or similar signalling pathways.

### 5.3.4 Conclusions

Evidence of M-mediated phosphorylation hints that the biochemical mechanisms behind NB-LRR protein activation may not be as simple as the existing evidence suggests; the role of ATP hydrolysis in this process may be more intricate than is currently understood. Does ATP hydrolysis re-set the protein back to active state, or does the release of inorganic phosphate initiate downstream pathways, and/or does it provide the energy for a conformational change that releases the TIR domain to initiate downstream signalling, or both? Phosphorylation could play a role in many parts of the activation pathway, regulation, signal transduction and/or conformational change, but the phenomenon observed obviously requires much more work before its significance can be understood. Nevertheless, it has opened the door for exciting new avenues of research.

The above discussion revolves around evidence from STAND proteins of the mammalian and bacterial systems and NB-LRR proteins other than M. As mentioned in previous chapters it may be unwise to rely on such data to make conclusions on the biochemical activity of another NB-LRR, but it is valuable in directing experimentation. As such, direct identification of potential phosphorylated residues on M and, also direct identification of $^{32}P$ labelled spots on TLC plates with MS, will be required to help explain the biochemical activity/activities of M protein. Studies with other M protein mutants were not possible here due to time limitations, but should be undertaken in the future, as they make help to unravel the role of this biochemical function.

It is possible that further hydrolysis of ADP is occurring, or even that PPi (pyrophosphate) is being generated by cleavage the $\alpha$-$\beta$ bond of ATP. Multiple and simultaneous reactions maybe occurring at the NBs pocket that control different outcomes and these possibilities remain to be explored.

# 6 Final Discussion

The overall aim of this thesis was to refine, redefine or expand the plant NB-LRR protein activation model. The current molecular switch model suggests that the active state of an R protein is ATP-bound and that ATP hydrolysis resets the protein back to an ADP-bound inactive state. Results presented here challenge this model. ATP hydrolysis, rather than being required for resetting an R protein, may be necessary for the M protein to activate the HR and the released phosphoryl groups maybe involved in a downstream signalling event. Whilst aspects of this model could not be explored with R proteins other than M, it may be possible in the future to express soluble RPS2 protein with the baculovirus/insect cell system, and this would facilitate biochemical investigations of a different class of NB-LRR. The potential of downstream protein phosphorylation opens a new avenue for experimentation of R protein activation, and suggests that our current thinking maybe an over simplification of the innate immune response in plants.

## 6.1 The NB-LRR Activation Model is Evolving

With increasing evidence of significant functional and biochemical differences between plant NB-LRRs, it is becoming clear that, despite their sequence and functional similarities, there is not one common mechanism behind their activation. Keeping in mind that there are likely to be a number of different molecular mechanisms involved in R protein activation, it is useful to gain insights into mechanisms employed by structurally similar proteins in the mammalian and bacterial STAND classification. However, given our relatively shallow knowledge of R protein function, exploration of individual proteins and the biochemistry of the molecular mechanisms utilised by each plant NB-LRR will likely be required for the activation of each R protein to be understood.

MalT and Apaf-1 represent archetypal STAND proteins whose activities fit the classic molecular switch model of activation in which nucleotide exchange triggers activation and one round of ATP hydrolysis is used as a re-set switch (Marquenet & Richet, 2007; Proell *et al.*, 2008; Richet & Raibaud, 1989; Riedl *et al.*, 2005). CED4 is an example of an unconventional STAND. It is bound to ATP in its resting state and is maintained in an autoinhibited state through its interaction with Ced9. It also does not have any

detectable ATPase activity (Yan *et al.*, 2005). Like CED4, the tobacco N is bound to ATP in its resting state. Additionally its ATPase activity is enhanced in the presence of its effector, p50 (Ueda *et al.*, 2006). N is likely to represent a plant NB-LRR that does not fit the conventional molecular switch model.

For the tightly regulated mammalian STAND proteins, NOD1 and NOD2, there are fundamental differences in the role that ATP hydrolysis plays in their activation pathways. These NACHT family proteins contain extended Walker B regions. An initial ATP hydrolysis event is necessary for the activation and signalling activity of each protein, while a subsequent round of ATP hydrolysis is required for NOD2 deactivation, but not for NOD1. The first round of hydrolysis is catalysed by the first acidic residue of the extended Walker B motif and the second residue catalyses the hydrolysis that deactivates the signalling complex/platform of NOD2 (Zurek *et al.*, 2011a). Other NOD-like receptor proteins also contain extended Walker B regions that are capable of performing ATP hydrolysis. These include NLRC4 (Ipaf-1) and NALP12 (Proell *et al.*, 2008).

That ATP hydrolysis is not required for the ETI activation by I-2 and Mi-1 was demonstrated by Tameling *et al.* (2006) when they showed that autoactive mutants displayed a slower ATPase activity than the wild-type proteins. Conversely, that the autoactive $M^{D555V}$ mutant protein displays a higher maximal rate of ATP hydrolysis than non-activated M (Sornaraj, 2013), and $M^{D364E}$ has impaired HR activation and ATP hydrolysis ability, suggests that for the M protein, ATP hydrolysis is important for signalling activation.

If activation is dependent on ATP hydrolysis, NB-LRRs need to be strictly regulated to prevent promiscuous signalling and unintended cell death activation. Furthermore, if plant R proteins are constantly cycling rather than existing in a static state, this suggests that they have roles in the cell that maybe un-related to pathogen perception, or that this activity sets a basal level of immune defence (Bonardi *et al.*, 2011). Roles in cell homeostasis have been demonstrated by Apaf-1 and other STAND proteins (Kufer *et al.*, 2011). The existence of a threshold of NB-LRR activity being required for ETI activation has long been postulated (Chisholm *et al.*, 2006; Jones and Dangl, 2006) and the increasing reports of R protein overexpression triggering

spontaneous cell death in the absence of a pathogen supports this (see chapter 4; Bernoux *et al.*, 2011a; Rentel *et al.*, 2008; Swiderski *et al.*, 2009; Zhang *et al.*, 2004).

Current evidence suggests that intramolecular interactions between the LRR and the ARC2 domains are responsible for autoregulation and maintaining the 'off' conformation of NB-LRRs. But there is also evidence of other interacting proteins providing regulation for plant NB-LRRs and also other STAND proteins. Many chaperones and co-chaperones have been found associated with R proteins (including Hsp90, RAR1 & SGT1) and it is likely that R protein levels in the absence of a pathogen are tightly regulated by some, or all, of these interacting partners (Collier & Moffett, 2009; Ma *et al.*, 2013). A protein phosphatase is also known to interact with some R proteins and may assist in regulating their activity in the absence of a pathogen (de la Fuente can Bentem *et al.*, 2005). Moffett *et al.* in 2002 also suggest that some R proteins may have associated signalling partners that are involved in signal initiation. Helper NB-LRR proteins have recently been demonstrated to act downstream of some traditional NB-LRRs to transmit signals subsequent to their activation (Bonardi *et al.*, 2011).

Such interactions between traditional NB-LRR proteins that do require a functional P-loop, and so-called 'helper' R proteins that do not, may also represent a new form of negative regulation (Bonardi *et al.*, 2013; Williams *et al.*, 2014). In the absence of a pathogen, could interactions with these biochemically compromised proteins act to regulate otherwise signalling-competent NB-LRRs? In the absence of a pathogen, could interactions with such proteins prohibit cell death signalling and instead channel plant R protein activity to other roles within the cell?

Complementary with intermolecular and intramolecular interactions, the formation of oligomeric structures by R proteins could also play a role in autoregulation. Such regulation occurs in the Apaf-1 oligomer via the tightly packed CARD domains preventing access to the bound ADP molecule and thus inhibiting nucleotide exchange and/or hydrolysis (Riedl *et al.*, 2005). Whilst the formation of multimeric R protein structures has not been shown for all R proteins studied thus far, similar regulation could play an important role for those that have.

Conversely, by looking at R protein ATPase activity *in vitro* are we missing some regulatory proteins that normally inhibit intrinsic ATPase activity *in planta*? Only further investigations will tell. The introduction of plant protein extracts into biochemical assays may allow us to assess whether NB-LRR proteins have positively and/or negative regulatory partners.

## 6.2 Future Directions

Many unanswered questions remain in this area of plant NB-LRR protein activation. This portion of the final discussion will address some of the large areas of research that warrant further investigation. Some supplementary experiments have been discussed in the earlier results chapters and will not be re-examined here.

### 6.2.1 The effect of effectors

The impacts that effector proteins have on R protein biochemical function and activation remain some of the largest gaps in knowledge in activation modelling. Determining what effect, if any, these proteins will have on ATP hydrolysis rates and also nucleotide exchange will be an important step in our understanding of the steps that occur upon, and as a consequence of, recognition.

To trigger activation, R protein/effector interaction could change nucleotide binding preferences and/or increase or decrease the rate of ATP hydrolysis, or simply displace negatively regulating interacting proteins. It remains to be determined whether the direct interaction between M and AvrM triggers any such response, or if the molecular consequence of the interaction is something different entirely. *In vitro* attempts to influence nucleotide binding and/or ATP hydrolysis in the M protein with AvrM have so far been either inconclusive or unsuccessful (data not shown; Sornaraj, 2013; Williams, 2009). Perhaps the interaction is too weak or transient to be captured *in vitro*, or using the conditions necessary to keep the proteins soluble. Furthermore, it could be that an as yet unidentified adaptor protein, normally present in the plant cell, is required to stabilise the interaction. Could it also be that AvrM interacts with an intermediate in the reaction cycle of M? In our *in vitro* assay, M (or sufficient quantities of M) may not be cycling and thus we cannot capture the interaction

complex, or the biochemical consequences of it. Such a possibility will be difficult to study, but may be aided with the use of M protein mutants that mimic such transition states. The M/AvrM interaction has been observed in the Y2H system, but attempts in this thesis to demonstrate a specific *in vitro* interaction by immunoprecipitation have been inconclusive. M was seen to interact with virulent avrM as well as avirulent AvrM, an interaction which is not observed in the Y2H system (data not shown).

Future work in this area will be helped by improvements in the purity and yield of recombinant M protein which may be achieved through further optimisation of the *Strep*Tactin® chromatography step. Experiments are already planned to investigate the interaction using surface plasmon resonance (SPR) technology. If successful, these experiments will also measure the affinity and kinetics of an M/AvrM interaction.

### 6.2.2 Recombinant NB-LRR expression and crystal structures

Thus far, biochemical studies of plant NB-LRR proteins have been limited by an inability to express soluble recombinant protein. Mla27, M and L6 are the only full-length and near full-length proteins for which expression and purification of recombinant protein has been successful, however even with these proteins biochemical investigations have been limited. Attempts were made during this study to express the *A. thaliana* CC-NB-LRRs, RPS2 and RPP1B. Despite many different expression constructs being trialled, no expression was achieved with the heterologous *P. pastoris* expression system as developed for M and L6. Although only insoluble RSP2 protein was expressed by the baculovirus/insect cell system this gives hope that further optimisation of the expression construct will lead to soluble full-length RPS2. This system may also be the best hope for producing RPP1B protein, but experimentation with suitable expression constructs remains to be attempted.

Although there has been limited success expressing the large, full-length plant R proteins, several of the smaller TIR and CC domain structures have recently been solved (Maekawa *et al.*, 2011a; Ve *et al.*, 2011; Williams *et al.*, 2014). Hopefully a full-length NB-LRR structure will be determined sooner rather than later. This will undoubtedly answer many of the mechanistic questions surrounding activation, but

it is likely that it will also bring many more to the fore. Furthermore, given the differences observed amongst NB-LRR proteins, it is likely that more than one crystal structure will be needed to address the many differences that exist in the functioning of different plant NB-LRRs. Solving the structures of active and inactive R protein conformations will also help with this task.

### 6.2.3 Expansion of biochemical investigations

Many extensions to the biochemical assays performed in this thesis have already been discussed in their respective chapters, but some warrant further mention here. Perhaps the most important steps forward for investigating the role of ATP hydrolysis will come through the use of differently radiolabelled ATP nucleotides. Use of γ-, β- and α-labelled ATP molecules will allow us to determine whether M can cleave sequential phosphates from ATP, while also helping to identify the mystery products of the hydrolysis reactions. It will also be interesting to determine whether M can use non-adenosine nucleotides, like GTP, as substrates for hydrolysis and/or phosphorylation reactions. The addition of AvrM to phosphorylation assays will also be an important step in identifying if this mechanism can be influenced by the effector, and whether it is directly involved in R protein activation. Experiments such as these will require improvements to recombinant M protein purity and yield.

Evidence of products other than ADP in the ATP hydrolysis reactions, and the potential of protein phosphorylation, suggests that the activation mechanism may be far more complex than ever imagined. The next steps in investigating this potential phosphorylation phenomenon should include identification of the mystery product visualised on TLC plates by MS, and/or the use of nucleotide and phosphate standards on the TLC plates. A phosphorylation specific stain like ProQ Diamond and antibodies that recognise phosphorylated amino acids could also be used to determine the nature of the $^{32}P$ signal identified by SDS-PAGE. As before, improvements to recombinant protein purity and yield will be required before such analyses can be undertaken.

### 6.2.4 A quantitative measure of activity is needed

It has been the convention to judge the activity of an NB-LRR protein by visual assessment of the HR intensity that it generates, with or without an effector. Such an assay was used in this study; however, visual measurement of the HR has limited powers when it comes to discerning between degrees of NB-LRR activity. It is often difficult to distinguish between weak and normal activity and autoactive phenotypes, and to assess whether effector co-expression stimulates activity or autoactivity. Electrolyte leakage and trypan blue staining have all been previously employed as ways to quantify the HR, but these methods also have limited quantitative powers (Maekawa *et al.*, 2011a; van Ooijen *et al.*, 2008).

A more robust method of measuring activity levels would give greater power to the assignment of activity phenotypes. The power of studies on other STANDs like MalT and NOD1 and NOD2 is dependent on a quantifiable measure of the influence of the effector on STAND activity, enabling researchers to clearly differentiate between activity, autoactivity and inactivity phenotypes with *in vitro* assays (Zurek *et al.*, 2011b). The best case scenario for R proteins would be the development of a downstream activity-linked assay that could be used to measure autoactivity activity and effector-stimulated activity. It is especially important to develop and use a linked assay because the ATPase activity observed for M, as in other STAND proteins, is low and difficult to measure, even in the presence of effectors (Lee *et al.*, 2002; Marquenet & Richet, 2007).

The development of such an assay will be helped by the identification of downstream interacting proteins that are modified, or have their activity influenced by activation of an NB-LRR. Such proteins may include partner 'helper' NB-LRRs or transcription factors (Bonardi *et al.*, 2011; Chang *et al.*, 2013). If, as is suggested by the results in chapter 5, the M protein is capable of phosphorylation, phosphate addition to self or non-self proteins could be exploited as a means of quantifying activation. The introduction of plant protein extracts into assays containing M protein and radiolabelled nucleotides may allow the identification of downstream proteins that are phosphorylated by the activity of M. Such experiments would need to have strict controls due to the existence of many proteins also capable of performing

phosphorylation in plant cells. Any free radiolabelled nucleotide would also need to be removed before incubation with plant extracts for an accurate result to be gained. Nevertheless, downstream proteins identified in this manner could subsequently be introduced into *in vitro* activity-linked assays.

## 6.3 Expanded Activation Models for the M Protein

The current pool of evidence collectively suggests that NB-LRR proteins will use diverse mechanisms of activation. With this in mind, the following activation model for M is based only on biochemical and functional evidence from M protein studies and it is possible that this model will only be relevant for M protein activation. The crucial pieces of evidence used to formulate this model are as follows:

- M interacts directly with its cognate effector AvrM and this interaction is mediated by the LRR domain (Catanzariti *et al.*, 2010);
- Expressed alone, M's TIR domain does not homodimerise and is not autoactive (M. Bernoux, personal communication, 2012);
- Wild-type M is purified with bound ADP (Williams *et al.*, 2011);
- The ability of M to activate the HR is dependent on a functional nucleotide binding site, as demonstrated by the reduced occupancies of the inactive $M^{K286L}$ mutant (Williams *et al.*, 2011);
- Wild-type M can hydrolyse ATP (Sornaraj, 2013);
- The autoactive mutant $M^{D555V}$ is purified with more bound ATP than bound ADP, has a higher rate of hydrolysis than M (Williams *et al.*, 2011) and interaction with AvrM may increase this ATP hydrolysis rate (Sornaraj, 2013);
- Autoactive mutants can generate a more intense HR in the presence of AvrM (chapter 4; Sornaraj, 2013);
- Overexpression of M *in planta* generates spontaneous cell death (chapter 4);
- $M^{D364E}$ is totally occupied with ADP (Williams *et al.*, 2011), appears to have impaired ATPase activity and also develops a slower HR than M in co-infiltrations with AvrM (chapter 4);
- M is potentially undergoing auto-phosphorylation, or phosphorylating other proteins with the product(s) of the ATP hydrolysis reaction (chapter 5).

### 6.3.1 Activation mediated by ATP hydrolysis and phosphorylation

As a consequence of its intrinsic ability to hydrolyse ATP, the M protein continuously cycles between ATP and ADP bound forms within an unchallenged cell. This background level of signalling is below an HR activation threshold. Interaction with AvrM through the LRR domain increases the ATP hydrolysis rate, breaking the HR threshold. The mechanochemical energy released drives conformational changes, which cycles the protein to the ADP-bound state, ready for another round of activation. The phosphate product of ATP hydrolysis initiates a phosphorelay pathway, which culminates in induction of the HR and ETI.

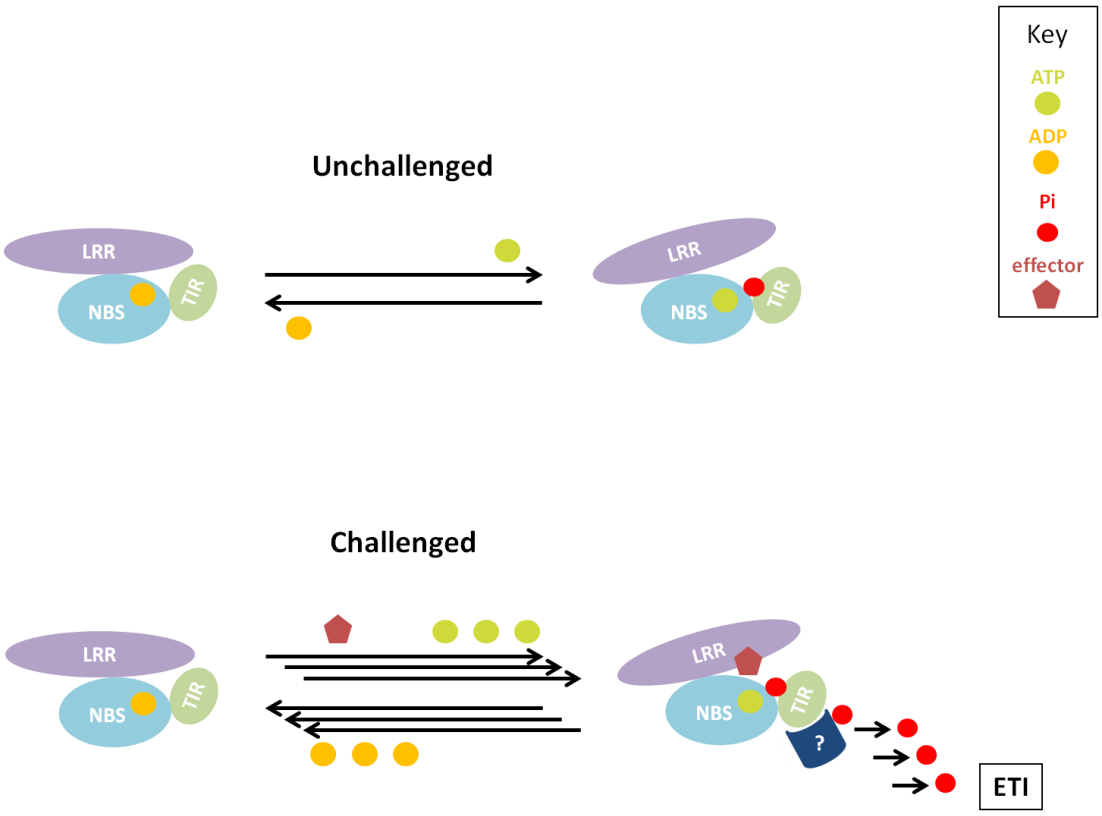



**Figure 6-1: ATP hydrolysis and phosphorylation mediate activation and ETI**

In the absence of a pathogen, in an unchallenged cell, the M protein slowly cycles between ATP- and ADP-bound states. This basal level of activity remains below the HR/ETI initiation threshold until M recognises AvrM and its ATP hydrolysis rate is enhanced. This increased activity breaks the HR/ETI initiation threshold and the phosphate products of the ATP hydrolysis reactions initiate signalling cascades.

## 6.4 Conclusions

As a whole, the plant NB-LRR field seems to be moving away from a 'one model fits all' mindset towards an acceptance that differences may exist between plant NB-LRR proteins. Whilst it was not possible to produce NB-LRR proteins other than M, this study has identified the baculovirus/insect cell system as a viable option for future such investigations. Expanding the number of R proteins that can be studied biochemically will be the way forward to understanding the many mechanisms of R protein activation and activity. By improving the purity and yield of the M protein after dual affinity purification, it is hoped that full-kinetic analysis of more forms of M can be undertaken. Such improvements will assist exploration into the phosphorylation events that maybe associated with M protein activity.

Whilst not providing conclusive evidence, this study has provided a tantalising sign of an expanded biochemical function for the M flax rust resistance protein. It is hoped that future work will determine the exact nature and function of this phenomenon, so that the M protein activation model can be comprehensively characterised. It is also hoped that such understanding can help shape the development of more effective and robust disease resistance in the ever increasing hostile environment in which our crop plants must compete and excel.

# Appendices

# Appendix A

**M coding sequence**

MSYLRDVATAVALLLDNLCCGRPNLNNDNEDTIQQTDSTSPVVDPSSSSQSMDSTSVVDAI
SDSTNPSASFPSVEYDVFLSFRGPDTRYQITDILYRFLCRSKIHTFKDDDELHKGEEIKVNLLRAI
DQSKIYVPIISRGYADSKWCLMELAKIVRHQKLDTRQIIIPIFYMVDPKDVRHQTGPYRKAFQ
KHSTRYDEMTIRSWKNALNEVGALKGWHVKNNDEQGAIADEVSANIWSHISKENFILETDE
LVGIDDHVEVILEMLSLDSKSVTMVGLYGMGGIGKTTTAKAVYNKISSHFDRCCFVDNVRA
MQEQKDGIFILQKKLVSEILRMDSVGFTNDSGGRKMIKERVSKSKILVVLDDVDEKFKFEDIL
GCPKDFDSGTRFIITSRNQNVLSRLNENQCKLYEVGSMSEQHSLELFSKHAFKKNTPPSDYET
LANDIVSTTGGLPLTLKVTGSFLFRQEIGVWEDTLEQLRKTLDLDEVYDRLKISYDALKAEAKEI
FLDIACFFIGRNKEMPYYMWSECKFYPKSNIIFLIQRCMIQVGDDGVLEMHDQLRDMGREIV
RREDVQRPWKRSRIWSREEGIDLLLNKKGSSQVKAISIPNNMLYAWESGVKYEFKSECFLNL
SELRLFFVGSTTLLTGDFNNLLPNLKWLDLPRYAHGLYDPPVTNFTMKKLVILVSTNSKTEWS
HMIKMAPRLKVVRLYSDYGVSQRLSFCWRFPKSIEVLSMSGIEIKEVDIGELKNLKTLDLTSCR
IQKISGGTFGMLKGLIELRLDSIKCTNLREVVADIGQLSSLKVLKTEGAQEVQFEFPLALKELSTS
SRIPNLSQLLDLEVLKVYGCNDGFDIPPAKSTEDEGSVWWKASKLKSLKLYRTRININVVDASS
GGRYLLPSSLTSLEIYWCKEPTWLPGIENLENLTSLVVDDVDIFQTLGGDLDGLQGLRSLETLTI
TEVNGLTRIKGLMDLLCSSTCKLEKLEIKACHDLTEILPCELHDQTVVVPSFEKLTIRDCPRLEVG
PMIRSLPKFPMLKKLDLAVANITKEEDLDVIGSLQELVDLRIELDDTSSGIERIASLSKLKKLTTLR
VKVPSLREIEELAALKSLQRLILEGCTSLERLRLEKLKEPDIGGCPDLTELVQTVVVCPSLVELTIR
DCPRLEVGPMIRSLPKFPMLKKLDLAVANIIEEDLDVIGSLEELVILSLKLDDTSSSSIERISFLSKL
QKLFRLRVKVSSLREIEGLAELKSLQLLFLKGCTSLERLWPDEQQLDNNKSMRIDIRGCKSLSV
DHLSALKSTLPPNVKIRWPDEKYK

**Key:** N-terminal hydrophobic region TIR domain NBS domain LRR domain

# Appendix B

**RPP1B coding sequence**

MGSAMSLGCSKRKATNQDVDSESRKRRKICSTNDAENCRFIQDESSWKHPWSLCANSVVN
DTKDTKSSALSLPSPPTSVSRIWKHQVFPSFHGADVRKTILSHILESFRRKGIDPFIDNNIERSK
SIGHELKEAIKGSKIAIVLLSKNYASSSWCLDELAEIMKCRELLGQIVMTIFYEVDPTDIKKQTG
EFGKAFTKTCKGKTKEYVERWRKALEDVATIAGYHSHKWRNEADMIEKIATDVSNMLNSFK
PSRDFNGLVGMRAHMDMLEQLLRLVLDEVRMIGIWGPPGIGKTTIARFLFNQVSDRFQLSA
IMVNIKGCYPRPCFDEYSAQLQLQNQMLSQMINHKDIMISHLGVAQERLRDKKVFLVLDEV
DQLGQLDALAKETRWFGPGSRIIITTEDLGVLKAHGINHVYKVGYPSNDEAFQIFCMNAFGQ
KQPHEGFDEIAREVMALAGELPLGLKVLGSALRGKSKPEWERTLPRLKTSLDGKIGSIIQFSYD
ALCDEDKYLFLYIACLFNKESTTKVEGLLGKFLDVRQGLHILAQKSLISIEDGNIYMHTLLEQFG
RETSRKQFIHHGYTKHQLLVGERDICEVLNDDTIDSRRFIGINLDLYKNVEELNISEKALERIHD
FQFVRINGKNHALHERLQGLIYQSPQIRSLHWKCYQNICLPSTFNSEFLVELDMSFSKLQKLW
EGTKQLRNLKWMDLSYSSYLKELPNLSTATNLEELKLRNCSSLVELPSSIEKLTSLQILDLHRCSS
LVELPSFGNATKLEILNLENCSSLVKLPPSINANNLQELSLTNCSRVVELPAIENATNLWKLNLL
NCSSLIELPLSIGTATNLKHLDFRGCSSLVKLPSSIGDMTNLEVFYLSNCSNLVELPSSIGNLRKLT
LLLMRGCSKLETLPTNINLKSLHTLNLIDCSRLKSFPEISTHIKYLRLIGTAIKEVPLSIMSWSPLA
HFQISYFESLKEFPHALDIITELQLSKDIQEVPPWVKRMSRLRALRLNNCNNLVSLPQLPDSLA
YLYADNCKSLERLDCCFNNPEIRLYFPKCFKLNQEARDLIMHTSTRNFAMLPGTQVPACFNH
RATSGDSLKIKLKESPLPTTLTFKACIMLVNEEMSYDLKSMSVDIVIRDEQNDLKVQCTPSYH
QCTEIYVLTEHIYTFELEVEEVTSTELVFEFTSVNESICKIGECGILQRETRSLRRSSSPDLSPESSR
VSSCDHC

**Key: Exon1 Exon2 Exon3 Exon4 Exon5**

# Appendix C

**RPS2 coding sequence**

MDFISSLIVGCAQVLCESMNMAERRGHKTDLRQAITDLETAIGDLKAIRDDLTLRIQQDGLEG
RSCSNRAREWLSAVQVTETKTALLLVRFRRREQRTRMRRRYLSCFGCADYKLCKKVSAILKSIGE
LRERSEAIKTDGGSIQVTCREIPIKSVVGNTTMMEQVLEFLSEEEERGIIGVYGPGGVGKTTLM
QSINNELITKGHQYDVLIWVQMSREFGECTIQQAVGARLGLSWDEKETGENRALKIYRALRQ
KRFLLLLDDVWEEIDLEKTGVPRPDRENKCKVMFTTRSIALCNNMGAEYKLRVEFLEKKHAW
ELFCSKVWRKDLLESSSIRRLAEIIVSKCGGLPLALITLGGAMAHRETEEEWIHASEVLTRFPAE
MKGMNYVFALLKFSYDNLESDLLRSCFLYCALFPEEHSIEIEQLVEYWVGEGFLTSSHGVNTIY
KGYFLIGDLKAACLLETGDEKTQVKMHNVVRSFALWMASEQGTYKELILVEPSMGHTEAPK
AENWRQALVISLLDNRIQTLPEKLICPKLTTLMLQQNSSLKKIPTGFFMHMPVLRVLDLSFTSI
TEIPLSIKYLVELYHLSMSGTKISVLPQELGNLRKLKHLDLQRTQFLQTIPRDAICWLSKLEVLNL
YYSYAGWELQSFGEDEAEELGFADLEYLENLTTLGITVLSLETLKTLFEFGALHKHIQHLHVEEC
NELLYFNLPSLTNHGRNLRRLSIKSCHDLEYLVTPADFENDWLPSLEVLTLHSLHNLTRVWGN
SVSQDCLRNIRCINISHCNKLKNVSWVQKLPKLEVIELFDCREIEELISEHESPSVEDPTLFPSLK
TLRTRDLPELNSILPSRFSFQKVETLVITNCPRVKKLPFQERRTQMNLPTVYCEEKWWKALEK
DQPNEELCYLPRFVPNTS

**Key:** N-terminal hydrophobic region CC region NBS domain LRR domain

# Appendix D

**ATP and ADP occupancies of Walker B mutant proteins compared to wild-type M, an autoactive mutant, $M^{D555V}$ and two inactive mutants, $M^{K286L}$ and $M^{K286L+D555V}$.**

| | Function *in planta* | ATP occupancy (%) | ADP occupancy (%) | N |
|---|---|---|---|---|
| $M^{*}$ | Autoctive | 2 ± 0.08 | 35 ± 3 | 4 |
| $M^{K286L*}$ | Inactive | 0.5 ± 0.1 | 3 ± 1 | 4 |
| $M^{D555V*}$ | Autoactive | 18 ± 4 | 5 ± 2 | 4 |
| $M^{K286L+D555V*}$ | Inactive | 0.8 ± 0.3 | 7 ± 3 | 4 |
| $M^{D363A}$ | Autoactive (very weak) | 0.4 ± 0.2 | 27 ± 9 | 4 |
| $M^{D364A}$ | Active (weak) | 4 ± 0.6 | 23 ± 3 | 4 |
| $M^{D363A+D364A}$ | Inactive | 0.4 ± 0.1 | 3 ± 1 | 3 |
| $M^{D364E}$ | Active | 1 ± 0.1 | 105 ± 5 | 4 |
| $M^{D366A}$ | Autoactive (weak) | 4 ± 0.4 | 12 ± 2 | 3 |

*from Williams *et al.*, 2011

**Empty vector ATP and ADP concentrations**

| | ATP (M) | ADP (M) | N |
|---|---|---|---|
| Empty vector | 2E-09 ± 3E-10 | 3E-08 ± 2E-08 | 4 |

# References

Aarts, N., Metz, M., Holub, E., Staskawicz, B.J., Daneils, M.J. & Parker, J.E. (1998). Different requirement for EDS1 and NDR1 by disease resistance genes defines at least two *R* gene-mediated signalling pathways in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.*, **95**: 10306-10311.

Ade, J., DeYoung, B., Goldstein, C. & Innes, R.W. (2007). Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease. *Proc. Natl. Acad. Sci. U. S. A.*, **104**: 2531-2536.

Albrecht M. & Takken F.L.W. (2006). Update on the domain architectures of NLRs and R proteins. *Biochem. Biophys. Res. Commun.*, **339**: 459-462.

Albrecht, M., Domingues, F.S., Schreiber, S. & Lengauer, T. (2003). Structural localisation of disease-associated sequence variations in the NACHT and LRR domains of PYPAF1 and NOD2. *FEBS Letters*, **554**: 520-528.

Aw, R. and Polizzi, K.M. (2013). Can too many copies spoil the broth? *Microbial Cell Factories*, **12**: 128-137.

Axtell, M. J. & 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.

Axtell, M.J., McNellis, T.W., Mudgett, M., Hsu, C.S. & Staskawicz, B.J. (2001). Mutational analysis of the *Arabidopsis RPS2* disease resistance gene and the corresponding *Pseudomonas syringae avrRpt2* avirulence gene. *Mol. Plant-Microbe Interact.*, **14**: 181-188.

Bai S., Liu J., Chang C., Zhang L., Maekawa T., Wang, Q., Xiao, W., Liu, Y., Chai, J., Takken, F.L.W., Schulze-Lefert, P. and Shen, Q. (2012). Structure-function analysis of barley NLR immune receptor MLA10 reveals its cell compartment specific activity in cell death and disease resistance. *PLoS Pathog*., **8**(6): e1002752. doi:10.1371/journal.ppat.1002752.

Belkhadir, Y., Subramaniam, R. & Dangl, J.L. (2004). Plant disease resistance protein signaling: NBS-LRR proteins and their partners. *Curr. Opin. Plant Biol.*, **7**: 391-399.

Bendahmane, A., Farnham, G., Moffett, P. & 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. *Plant J.*, **32**: 195-204.

Bernoux, M., Ellis, J.G. & Dodds, P.N. (2011a). New insights in plant immunity signaling activation. *Curr. Opin. Plant Biol.*, **14**: 1-7.

Bernoux, M., Ve, T., Williams, S., Warren, C., Hatters, D., Valkov, E., Zhang, X., Ellis, J.G., Kobe, B. & Dodds, P.N. (2011b). Structural and functional analysis of a plant resistance protein TIR domain reveals interfaces for self-association, signalling, and autoregulation. *Cell Host & Microbe*, **9**: 200-211.

Birch, P.R.J, Rehmany, A.P., Pritchard, L., Kamoun, S. & Beynon, J.L. (2006). Trafficking arms: oomycete effectors enter host plant cells. *Trends Microbiol.*, **14**: 8-11.

Boettner, M., Steffens, C., von Mering, C., Bork, P., Stahl, U. & Lang, C. (2007). Sequence-based factors influencing the expression of heterologous genes in the yeast *Pichia pastoris*- A comparative view on 79 human genes. *J. Biotechnol.*, **130**: 1-10.

Bonardi, V., Cherkis, K., Nishimura, M.T. & Dangl, J.L. (2012). A new eye on NLR proteins: focused on clarity or diffused by complexity? *Curr. Opin. Immunol.*, **24**: 41-50.

Bonardi, V., Tang, S., Stallmann, A., Roberts, M., Cherkis, K. & Dangl, J.L. (2011). Expanding functions for a family of plant intracellular immune receptors beyond specific recognition of pathogen effectors. *Proc. Natl. Acad. Sci. U. S. A.*, **108**: 16463-16468.

Botella, M.A., Parker, J.E., Frost, L.N., Bittner-Eddy, P.D., Beynon, J.L., Daniels, M.J., Holub, E.B & Jones, J.D.G (1998). Three genes of the *Arabidopsis RPP1 c*omplex resistance locus recognize distinct *Peronospora parasitica* avirulence determinants. *Plant Cell*, **10**: 1847-1860

Campbell, M. J. & Davis, R. W. (1999). On the *in vivo* function of the RecA ATPase. *J. Mol. Biol.*, **286**: 437-445.

Catanzariti, A., Dodds, P.N., Ve, T., Kobe, B., Ellis, J.G. & Staskawicz, B.J. (2010). The AvrM effector from flax rust has a structured C-terminal domain and interacts directly with the M resistance protein. *Mol. Plant-Microbe Interact.*, **23**: 49-57.

Catanzariti, A.M., Dodds, P.N., Lawrence, G.J., Ayliffe, M.A., Ellis, J.G. (2006). Haustorially-expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *Plant Cell*, **18**: 243-256.

Chang, C., Yu, D., Jiao, J., Jing, S., Schulze-Lefert, P. & Shen, Q. (2013). Barley MLA immune receptors directly interfere with antagonistically acting transcription factors to initiate disease resistance signalling. *The Plant Cell*, **25**: 1158-1173.

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

Clare, J. J., Rayment, F. B., Ballantine, S. P., Sreekrishna, K. & Romanos, M. A. (1991a). High-level expression of tetanus Toxin fragment C in *Pichia pastoris* strains containing multiple tandem integrations of the gene. *Bio/Technology*, **9**: 455-460.

Clare, J. J., Romanos, M. A., Rayment, F. B., Rowedder, J. E., Smith, M. A., Payne, M. M., Sreekrishna, K. & Henwood, C. A. (1991b). Production of epidermal growth factor in yeast: High-level secretion using *Pichia pastoris* strains containing multiple gene copies. *Gene*, **105**: 205-212.

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

Dangl, J.L. & Jones, J.D.G. (2001). Plant pathogens and integrated defence responses to infection. *Nature*, **411**: 826-833.

Danot, O. Marquenet, E., Vidal-Ingigliardi, D. & Richet, E. (2009). Wheel of life, wheel of death: a mechanistic insight into signalling by STAND proteins. *Structure*, **17**: 172-182.

Day, B., Dahlbeck, D. & Staskawicz, B.J. (2006). NDR1 interaction with RIN4 mediates the differential activation of multiple disease resistance pathways in *Arabidopsis*. *The Plant Cell*, **18**: 2782-2791.

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

de la Fuente van Bentem, S., Vossen, J.H., de Vries, K., van Wees, S.C., Tameling, W.I.L., Dekker, H., de Koster, C.G., Haring, M.A., Takken, F.L.W. & Cornelissen, B.J.C. (2005). Heat shock protein 90 and its co-chaperone protein phosphatase 5 interact with distinct regions of the tomato I-2 disease resistance protein. *Plant J.*, **43**: 284-298.

deCourcy-Ireland, E. (2007). A biochemical and functional analysis of autoactive M flax rust resistance proteins, Honours thesis, Flinders University, Adelaide, Australia.

Deslandes, L., Olivier, J., Theulieres, F., Hirsch, J., Feng, D.X., Bittner-Eddy, P., Beynon, J., Marco, Y. (2002). Resistance to Ralstonia solanacearum in Arabidopsis thaliana is conferred by the recessive RRS1-R gene, a member of a novel family of resistance genes. Proc. Natl. Acad. Sci U.S.A., 99: 2404-2409.

DeYoung, B.J. & Innes, R.W. (2006). Plant NBS-LRR proteins in pathogen sensing and host defence. *Nature Immunol.*, **7**: 1243-1249.

Dinesh-Kumar, S.P., Wai-Hong Tham & Baker, B.J. (2000). Structure-function analysis of the tobacco mosaic virus resistance gene *N*. *Proc. Natl. Acad. Sci. U. S. A.*, **97**: 14789-14794.

Dodds, P.N. & Rathjen, J.P. (2010). Plant immunity: towards an integrated view of plant-pathogen interactions. *Nat. Rev. Genet.*, **11**: 539-48.

Dodds, P.N., Lawrence, G.J., Catanzariti, A.M., Ayliffe, M.A. & Ellis, J.G. (2004). The *Melampsora lini AvrL567* avirulence genes are expressed in haustoria and their products are recognised inside plant cells. *Plant Cell*, **16**: 755-768.

Dodds, P.N., Lawrence, G.J., Catanzariti, A.M., Teh, T. Wang, C.A., Ayliffe, M.A. Kobe, B. & Ellis, J.G. (2006). Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance gene and flax rust avirulence genes. *Proc. Natl. Acad. Sci. U. S. A.*, **103**: 8888-8893.

Dodds, P.N., Lawrence, G.J., Pryor, T., Ellis, J. (2001). Six amino acid changes confined to the leucine-rich repeat β-strand/β-turn motif determine the difference between the *P* and *P2* rust resistance specificities in flax. *Plant Cell*, **13**: 163-178.

Earley, K., Haag, J.R., Pontes, O., Opper, K. Juehne, T., Song, K. & Pikaard, C.S. (2006). Gateway-compatible vectors for plant functional genomics and proteomics. *Plant J.*, **45**: 616-629.

Eitas, T.K. (2010). NB-LRR Regulation and Function in *Arabidopsis,* PhD thesis, University of North Carolina, Chapel Hill, U.S.A..

Ellis, J.G., Catanzariti, A.M., & Dodds, P.N. (2006). The problem of how fungal and oomycete avirulence proteins enter plant cells. *Trends Plant Sci.*, **11**: 61-63.

Ellis, J.G., Dodds, P.N., & Lawrence, G.J. (2007). Flax rust resistance gene specificity is based on direct resistance-avirulence protein interactions. *Annu. Rev. Phytopathol.*, **45**: 289-306.

Ellis, J.G., Lawrence, G.J., Luck, J.E. & Dodds, P.N. (1999). Identification of regions in alleles of the flax rust resistance gene *L* that determine differences in gene-for-gene specificity. *Plant Cell*, **11**: 495-506.

Eulgem, T., Rushton, P.J., Robatzek, S. & Somssich, I.E. (2000). The WRKY superfamily of plant transcription factors. *Trends Plant Sci.*, **5**: 199-206.

Fenyk, S., de San Eustaquio Campillo, A., Pohl, E., Hussey, P.J. & Cann, M.J. (2012). A nucleotide phosphatase activity in the nucleotide binding domain of an orphan resistance protein from rice. *J. Biol. Chem.*, **287**: 4023-4032.

Ferreira, R.B., Monteiro, S., Freitas, R., Santos, C.N., Chen, Z., Batista, L.M., Duarte, J., Borges, A., & Teixeira, A.R. (2007). The role of plant defence proteins in fungal pathogenesis. *Mol. Plant Path.*, **8**: 677-700.

Flor, H.H. (1956). The complementary genetic systems in flax and flax rust. *Adv. Gen.*, **8**: 29-54.

Flor, H.H. (1971). Current status of the gene-for-gene concept. *Annu. Rev. Phytopathol.*, **9**: 275-296.

Gabriëls, S.H.E.J., Vossen, J.H., Ekengren, S.K., van Ooijen, G., Abd-El-Haliem, A.M., van der Berg, G.C.M., Rainey, D.Y., Martin, G.B., Takken, F.L.W., de Wit, P.J.G.M. & Joosten, M.H.A.J. (2007). An NB-LRR protein required for HR signalling mediated by both extra- and intracellular resistance proteins. *Plant J.*, **50**: 14-28.

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. (2005). 'Protein Identification and Analysis Tools on the ExPASy Server', (In) John M. Walker (ed.): *The Proteomics Protocols Handbook*, Humana Press Inc., New Jersey, pp. 571-607.

Goodman, R.N. & Novacky, A.J. (1994). *The hypersensitive reaction in plants to pathogens.* Am. Phytopathol. Soc., St Paul, MN.

Graber, J.H., Cantor, C.R., Mohr, S.C & Smith, T.F. (1999). Genomic detection of new yeast pre-mRNA 3'-end-processing signals. *Nucleic Acids Res.*, **27**: 888-894.

Gurkan, C. & Ellar, D.J. (2003). Expression in *Pichia pastoris* and purification of a membrane-acting immunotoxin based on a synthetic gene coding for the *Bacillus thuringiensis* Cyt2Aa1 toxin. *Protein Expr. Purif.*, **29**: 103-16.

Gutierrez, J.R., Balmuth, A.L., Ntoukakis, V., Mucyn, T.S., Gimenez-Ibanez, S., Jones, A.M.E. & Rathjen, J.P. (2010). Prf immune complexes of tomato are oligomeric and contain multiple Pto-like kinases that diversify effector recognition. *Plant J.*, **61**: 507-518.

Hammond-Kosack, K.E. & Jones, J.G.D. (1997). Plant disease resistance genes. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **48**: 575-607.

Hammond-Kosack, K.E. & Parker, J.E. (2003). Deciphering plant-pathogen communication: fresh perspectives for molecular resistance breeding. *Curr. Opin. Biotechnol.*, **14**: 177-193.

Heath, M.C. (2000). Hypersensitive response-relate death. *Plant Mol. Biol.*, **44**: 321-334.

Higgins, D.R. & Cregg, J.M. (1998). 'Introduction: Distinctions between *Pichia pastoris* and other expression systems' in D.R. Higgins & J.M. Cregg (eds.), *Methods in Molecular Biology, Vol. 103* Pichia *Protocols*, 2nd edn., vol. 389, Humana Press Inc., New Jersey, pp. 1-15.

Howles, P., Lawrence, G., Finnegan, J., Mc Fadden, H., Ayliffe, M., Dodds, P., Ellis, J. (2005). Auto-active alleles of the flax *L6* rust resistance gene induce non race-specific rust resistance associated with the hypersensitive response. *Am. Phytopathol. Soc. Mol. Plant-Microbe Int.*, **18**: 570-582.

Hu, H., Gao, J., He, J., Yu, B., Zheng, P., Huang, Z., Mao, X., Yu, J., Han, G. & Chen, D. (2013a). Codon optimization significantly improves the expression level of a keratinase gene in *Pichia pastoris*. *PLoS ONE*, **8**: e58393. doi: 10.1371/journal.pone.0058393.

Hu, Z., Yan, C., Liu, P., Huang, Z., Ma, R., Zhang, C., Wang, R., Zhang, Y., Martinon, F., Miao, D., Deng, H., Wang, J., Chang, J. & Chai, J. (2013b). Crystal structure of NLRC4 reveals its autoinhibition mechanism. *Science*, **341**: 172-175.

Huang, C. Anderson, K.A., Damasceno, L.M., Ritter, G., Old, L.J. & Batt, C.A. (2010). Improved secretion of the cancer-testis antigen SSX2 in *Pichia pastoris* by deletion of its nuclear localization signal. *Appl. Microbiol. Biotechnol.,* **86**: 243-253.

Huang, H., Yang, P., Luo, H., Tang, H., Shao, N., Yuan, T. Wang, Y., Bai, Y. & Yao, B. (2008). High-level expression of a truncated 1,3-1,4-β-d-glucanase from *Fibrobacter succinogenes* in *Pichia pastoris* by optimization of codons and fermentation. *Appl. Microbiol. Biotechnol.*, **78**: 95-103.

Hwang, C.F. & Williamson, V.M. (2003). Leucine-rich repeat-mediated intramolecular interactions in nematode recognition and cell death signalling by the tomato resistance protein Mi. *Plant J.*, **34**: 585-593.

Inohara, N. & Nunez, G. (2003). Nods: Intracellular proteins involved in inflammation and apoptosis. *Nat. Rev. Immunol.*, **3**: 371-382.

Iyer, L.M., Leipe, D.D., Koonin, E.V. & Aravind, L. (2004). Evolutionary history and higher order classification of AAA+ ATPases. *J. Struct. Biol.*, **146**: 11-31.

Jia, Y., McAdams, S.A., Bryan, G.T., Hershey, H.P. & Valent, B. (2000). Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *EMBO J.*, **19**: 4004-4014.

Johal, G.S. & Briggs, S.P. (1992). Reductase activity encoded by the HM1 disease resistance gene in maize. *Science*, **258**: 985-987.

Jones, D.A. & Takemoto, D. (2004). Plant innate immunity- direct and indirect recognition of general and specific pathogen-associated molecules. *Curr. Opin. Immunol.,* **16**: 48-62.

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

Kang, H-G., Kuhl, J.C., Kachroo, P. & Klessig, D.F. (2008). CRT1, an *Arabidopsis* ATPase that interacts with diverse resistance proteins and modulates disease resistance to turnip crinkle virus. *Cell Host & Microbe*, **3**: 48-57.

Kemen, E., Kemen, A.C., Rafiqi, M., Hempel, U., Mendgen, K., Hahn, M. & Voegele, R.T. (2005). Identification of a Protein from Rust Fungi Transferred from Haustoria into Infected Plant Cells. Mol. Plant-Microbe Interact., 18: 1130-1139.

Kim, H. S., Desveaux, D., Singer, A. U., Patel, P., Sondek, J. & Dangl, J. L. (2005). The *Pseudomonas syringae* effector AvrRpt2 cleaves its C-terminally acylated target, RIN4, from *Arabidopsis* membranes to block RPM1 activation. *Proc. Natl. Acad. Sci. U.S.A.*, **102**: 6496-64501.

Kobe, B. & Deisenhofer, J. (1994). The leucine-rich repeat: a versatile binding motif. *Trends Biochem. Sci.*, **19**: 415-421.

Koonin, E.V. (1993). A common set of conserved motifs in a vast array of putative nucleic acid-dependent ATPases including MCM proteins involved in the initiation of eukaryotic DNA replication. *Nucleic Acids Res.*, **21**: 2541-2547.

Kovács, K.A., Steinmass, M., Magistretti, P.J., Halfon, O. & Cardinaux, J. (2003). CCAATT/Enhancer-binding protein family members recruit the coactivator CREB-binding protein and trigger its phosphorylation. *J. Biol. Chem.*, **278**: 36959-36965.

Krasileva, K.V., Dahlbeck, D. & Staskawicz, B.J. (2010). Activation of an *Arabidopsis* resistance protein is specified by the *in planta* association of its leucine-rich repeat domain with the cognate oomycete effector. *Plant Cell*, **22**: 2444-2458.

Kufer, T.A. & Sansonetti, P.J. (2011). NLR functions beyond pathogen recognition. *Nature Immunology*, **12**: 121-128.

Larkin M.A., Blackshields G., Brown N.P., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., Wilm A., Lopez R., Thompson J.D., Gibson T.J. & Higgins D.G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, **23:** 2947-2948.

Lawrence, G.J., Dodds, P.N. & Ellis, J.G. (2007). Pathogen profile: Rust of flax and linseed caused by *Melampsora lini*. *Mol. Plant. Pathol.,* **8**: 349-364.

Lee, P.C., Umeyama, T. & Horinouchi, S. (2002). *afsS* is a target of AfsR, a transcriptional factor with ATPase activity that globally controls secondary metabolism in *Streptomyces coelicolor* A3(2). *Mol. Microbiol.*, **43**: 1413-1430.

Leipe, D.D., Koonin, E.V. & 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. *J. Mol. Biol.*, **343**: 1-28.

Leister, R.T. and Katagiri, F. (2000). A resistance gene product of the nucleotide binding site-Leucine rich repeats class can form a complex with bacterial avirulence proteins *in vivo*. *Plant J*., **22**: 345-354.

Li, F., Yang, S., Zhao, L., Li, Q. & Pei, J. (2012). Synonymous codon usage bias and overexpression of a synthetic *xynB* gene from *Aspergillus niger* NL-1 in *Pichia pastoris*. *BioResources*, **7**: 2330-2343.

Liu, L., Spurrier, J., Butt, T.R. & Strickler, J.E. (2008). Enhanced protein expression in the baculovirus/insect cell system using engineered SUMO fusions. *Protein Expr. Purif.*, **62**: 21-28.

Luck, J.E., Lawrence, G.J., Dodds, P.N., Shepherd, K.W. & Ellis, J.G. (2000). Regions outside of the leucine-rich repeat domain of flax rust resistance proteins play a role in specificity determination. *Plant Cell*, **12**: 1367-1377.

Lukasik, E. and LW Takken, F.L.W (2009). STANDing strong, resistance proteins instigators of plant defence. *Curr. Opin. Plant Biol.*, **12**: 1-10.

Lukasik-Shreepaathy, E. (2011). Nucleotide-binding and molecular interactions of plant disease resistance proteins, PhD thesis, University of Amsterdam, Amsterdam, The Netherlands.

Lukat, G.S., McCleary, W., Stock, A.M & Stock, J.B. (1992). Phosphorylation of bacterial response regulator proteins by low molecular weight phospho-donors. *Proc. Natl. Acad. Sci. U.S.A.*, **89**: 718-722.

Lupas, A.N. & Jörg Martin, J. (2002). AAA proteins. *Curr. Opin. Struct. Biol.*, **12**: 746-753.

Ma, L., van den Burg, H.A., Cornelissen, B.J.C. & Takken, F.L.W (2013). 'Molecular basis of effector recognition by plant NB-LRR proteins' in G. Sessa (ed.), *Molecular Plant Immunity*, 1st edn., John Wiley & Sons, Inc., Iowa, pp. 23-40.

Mackey, D., Belkhadir, Y., Alonso, J.M., Ecker, J.R. & Dangl, J.L. (2003). *Arabidopsis* RIN4 is a target of the type III virulence effector AvrRpt2 and modulates RPS2-mediated resistance. *Cell*, **112**: 379-389.

Mackey, D., Holt, B.F. III, Wiig, A. & 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., Shen, Q., Micluta, M.A., Somssich, I.E., Takken, F.L.W. & Schulze-Lefert, P. (2011a). Coiled-coil domain-dependent homodimerisation of intracellular barley immune receptors defines a minimal functional module for triggering cell death. *Cell Host & Microbe*, **9**: 187-199.

Maekawa, T., Kufer, T.A. & Schulze-Lefert, P. (2011b). NLR functions in plant and animal immune systems: so far and yet so close. *Nat. Immunol.*, **12**: 817-826.

Marquenet, E. & Richet, E. (2007). How integration of positive and negative regulatory signals by a STAND signalling protein depends on ATP hydrolysis. *Mol. Cell*, **28**: 187-199.

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

Meyers, B.C., Dickerman, A.W., Michelmore, R.W., Sivaramakrishnan, S., Sobral. B.W. & Young, N.D. (1999). Plant disease resistance genes encode members of an

ancient and diverse proteins family within the nucleotide-binding superfamily. *Plant J.*, **20**: 317-332.

Meyers, B.C., Kozik, A., Griego, A., Kuang, H. & Michelmore, R.W. (2003). Genome-wide analysis of the NBS-LRR-encoding genes in *Arabidopsis*. *Plant Cell*, **15**: 809-834.

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

Muneyuki, E., Noji, H., Amano, T., Masaike, T. & Yoshida, M. (2000). $F_{(0)}F_{(1)}$-ATP synthase: general structural features of ATP-engine and a problem on free energy transduction. *Biochem. Biophys. Acta.*, **1458**: 467-481.

Nakagami, H., Sugiyama, N., Mochida, K., Daudi, A., Yoshida, Y., Toyoda, T., Tomita, M., Ishihama, Y. & Shirasu, K. (2010). Large-scale comparative phosphoproteomics identifies conserved phosphorylation sites in plants. *Plant Physiol.*, **153**:1161-1174.

Ni, M., Cui, D., Einstein, J., Nerasimhulu, S., Vergara, C.E. & Gelvin, S.B. (1995). Strength and tissue specificity of chimeric promoters derived from octopine and mannopine synthase genes. *Plant J.*, **7**: 661-676.

Nicholas, K.B., Nicholas H.B. Jr., and Deerfield, D.W. II. (1997). GeneDoc: Analysis and visualization of genetic variation, *EMBNEW.NEWS*, **4**:14. http://www.psc.edu/biomed/genedoc

Ogura, T. and Wilkinson, A.J. (2001). AAA+ superfamily ATPases: common structure-diverse function. *Genes Cells*, **6**: 575-597.

Oldroyd, G.E.D. and Staskawicz, B.J. (1998). Genetically engineered broad-spectrum disease resistance in tomato. *Proc. Natl. Acad. Sci. U.S.A.*, **95**: 10300-10305.

Proell, M., Riedl, S.J., Fritz, J.H., Rojas, A.M. & Schwarzenbacher, R. (2008). The Nod-Like Receptor (NLR) family: A tale of similarities and differences. *PLoS ON*E **3**: e2119. doi:10.1371/journal.pone.0002119

Qi, S., Pang, Y., Hu, Q., Liu, Q., Li, H., Zhou, Y., He, T., Liang, Q., Liu, Y., Yuan, X., Luo, G., Li, H., Wang, J., Yan, N. & Shi, Y. (2010). Crystal structure of the *Caenorhabditis elegans* apoptosome reveals an octameric assembly of CED-4. *Cell*, **141**: 446-457.

Qu, Y., Misaghi, S., Izrael-Tomasevic, A., Newton, K., Gilmour, L.L., Lamkanfi, M., Louie, S., Kayagaki, N., Liu, J., Kömüves, L., Cupp, J.E., Arnott, D., Monack, D. & Dixit, V.M. (2012). Phosphorylation of NLRC4 is critical for inflammasome activation. *Nature*, **490**: 539-542.

Rafiqi, M., Gan, P.H., Ravensdale, M., Lawrence, G.J., Ellis, J.G., Jones, D.A., Hardham, A.R., & Dodds, P.N. (2010). Internalization of flax rust avirulence proteins into flax and tobacco cells can occur in the absence of the pathogen. *Plant Cell*, **22**: 2017-2032.

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

Rairdan, G.J. & 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. & Moffett, P. (2008). The Coiled-coil and nucleotide binding domains of the potato Rx disease resistance protein function in pathogen recognition and signalling. *Plant Cell*, **20**: 739-751.

Rehmany, A., Gordon, A., Rose, L.E., Allen, R.L., Armstrong, M.R., Whisson, S.C., Kamoun, S., Tyler, B.M., Birch, P.R.J. & Beynon, J.L. (2005). Differential recognition of highly divergent downy mildew avirulence gene alleles by *RPP1* resistance genes from two *Arabidopsis* lines. *Plant Cell*, **17**: 1839-1850.

Rentel, M.C., Leonelli, L., Dahlbeck, D., Zhao, B. & Staskawicz, B.J. (2008). Recognition of the *Hyaloperonospora parasitica* effector ATR13 triggers resistance against oomycete, bacterial, and viral pathogens. *Proc. Natl. Acad. Sci. U. S. A.*, **105**: 1091-1096.

Resh, M.D. (1999). Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. *Biochim. Biophys. Acta*, **1451**: 1-16.

Richet, E. & Raibaud, O. (1989). MalT, the regulatory protein of the *Escherichia coli* maltose system, is an ATP-dependent transcriptional activator. *EMBO J.*, **8**: 981–987.

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

Rigden, D.J., Mello, L.V. & Bertioli, D.J. (2000). Structural modelling of a plant disease resistance gene product domain. *Proteins*, **41**:133-43.

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). *Molecular Cloning: A laboratory manual.* Cold Spring Harbour Laboratory Press, New York.

Saraste, M., Sibbald, P.R. & Wittinghofer, A. (1990). The P-loop- a common motif in ATP- and GTP-binding proteins. *Trends Biochem. Sci.*, **15**: 430-434.

Schaller, G.E., Shiu, S. & Armitage, J.P. (2011). Two-component systems and their co-option for eukaryotic signal transduction. *Curr. Biol.*, **21**: R320-R330.

Schmidt, S.A., Williams, S.J., Wang, C.A., Sornaraj, P, James, B., Kobe, B., Dodds, P.N., Ellis, J.G. & Anderson, P.A. (2007). Purification of the M flax-rust resistance protein expressed in *Pichia pastoris*. *Plant J.*, **50**: 1107-1117.

Scorer, C. A., Buckholz, R. G., Clare, J. J. & Romanos, M. A. (1993). The intracellular production and secretion of HIV-1 envelope protein in the methylotrophic yeast *Pichia pastoris*. *Gene*, **136**: 111-119.

Shen, Q., Wu, M., Wang, H.B., Naranmandura, H. & Chen, S.Q. (2012). The effect of gene copy number and co-expression of chaperone on production of albumin fusion proteins in *Pichia pastoris*. *Appl Microbiol Biotechnol.*, **96**(3): 763-772.

Slootweg, E. J., Spiridon, L.N., Roosien, J., Butterbach, P., Pomp, R., Westerhof, L., Wilbers, R., Bakker, E., Bakker, J., Petrescu, A., Smant, G. & Goverse, A. (2013). Structural determinants at the interface of the ARC2 and leucine-rich repeat domains control the activation of the plant Immune receptors Rx1 and Gpa2. J. *Plant Physiol.*, **162**: 1510-1528.

Sornaraj, P. (2013). Activation of the M flax-rust resistance protein, PhD thesis, Flinders University, Adelaide, Australia.

Staskawicz, B. J., Mudgett, M. B., Dangl, J. L. & Galan, J. E. (2001). Common and contrasting themes of plant and animal diseases. *Science,* **292**: 2285-2289.

Stock, A.M., Robinson, A.L. & Goudreau, P.N. (2000). Two-component signal transduction. *Annu. Rev. Biochem.*, **69**: 183-215.

Story, R.M. & Steitz, T.A. (1992). Structure of the recA protein-ADP complex. *Nature*, **355**: 374-376.

Swiderski, M.R., Birker, D. & Jones, J.D.G. (2009). The TIR domain of TIR-NB-LRR resistance proteins is a signalling domain involved in cell death induction. *Mol. Plant-Microbe Interact.*, **22**: 157–165.

Takemoto, D. & Jones, D. A. (2005). Membrane release and destabilization of *Arabidopsis* RIN4 following cleavage by *Pseudomonas syringae* AvrRpt2. *Mol. Plant-Microbe Interact.*, **18**: 1258-1268.

Takemoto, D., Rafiqi, M., Hurley, U., Lawrence, G.J., Bernoux, M., Hardham, A.R., Ellis, J.G., Dodds, P.N. & Jones, D.A. (2012). N-terminal motifs present in a subset of plant disease resistance proteins function in membrane attachment and contribute to the efficiency of disease resistance. *Mol. Plant-Microbe Interact.*, **25**: 379-92.

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

Takken, F.L.W., Albrecht, M. & Tameling, W.I.L. (2006). Resistance proteins: molecular switches of plant defence. *Curr. Opin. Plant Biol.*, **9**: 383-390.

Tameling, W.I.L., Elzinga, S.D.J., Darmin, P.S., Vossen, J.H., Takken, F.W., Haring, M.A. Cornelissen, B.J.C. (2002). The tomato *R* gene products I-2 and Mi-1 are functional ATP binding proteins with ATPase activity. *Plant Cell*, **14**: 1-11.

Tameling, W.I.L., Vossen, J.H., Albrecht, M., Lengauer, T., Berden, J.A., Haring, M.A., Cornelissen, B.J.C. & Takken, F.L.W. (2006). Mutations in the NB-ARC domain of I-2 that impair ATP hydrolysis cause autoactivation. *Plant Phys.*, **140**: 1233-1245.

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

The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.

Tornero, P., Chao, R.A., Luthin, W.N., Goff, S.A. & Dangl, J.L. (2002). Large-scale structure-function analysis of the *Arabidopsis* RPM1 disease resistance protein. *Plant Cell*, **14**: 435-450.

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

Ueda, H., Yamaguchi, Y. & Sano, H. (2006). Direct interaction between the tobacco mosaic virus helicase domain and the ATP-bound resistance protein, N factor during the hypersensitive response in tobacco plants. *Plant Mol. Biol.*, **61**: 31-45.

Urao, T., Yamaguchi-Shinozaki, K. & Shinozaki, K. (2000). Two-component systems in plant signal transduction. *Trends Plant Sci.*, **5**: 67-74.

van der Biezen, E.A. & Jones, J.D. (1998a). The NB-ARC domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. *Curr. Biol.*, **8**: R226-R227.

van der Biezen, E.A. & Jones, J.D.G. (1998b). Plant disease resistance proteins and the gene-for-gene concept. *Trends Biochem. Sci.*, **23**: 454-456.

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

van Ooijen, G., Mayr, G., Kasiem, M.M.A., Albrecht, M., Cornlissen, B.J.C. & Takken, F.L.W. (2008). Structure-function analysis of the NB-ARC domain of plant disease resistance proteins. *J. Exp. Bot.*, **59**: 1383-1397.

Ve, T., Williams, S., Valkov, E., Ellis, J.G., Dodds, P.N. & Kobe, B. (2011). Crystalisation, X-ray diffraction analysis and preliminary structure determination of the TIR domain

from the flax rust resistance protein. *Acta. Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.*, **67**: 237-240.

Walker, J.E., Saraste, M., Runswick, M.J. & Gay, N.J. (1982). Distantly related sequences in the $\alpha$- and $\beta$-subunits of ATP Synthetase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.*, **1**: 945-951.

Wang, C.I., Guncar, G., Forwood, J.K., the, T., Catanzariti, A.M., Lawrence, G.J., Loughlin, F.E., Mackay, J.P., Schirra, H.J., Anderson, P.A., Ellis, J.G., Dodds, P.N. & Kobe, B. (2007). Crystal structures of flax rust avirulence protein AvrL567-A and –D reveal the structural basis for flax disease resistance specificity. *Plant Cell*, **19**: 2898-2912.

Weaver, L.M., Swiderski, M.R., Yan Li & Jones, J.D.G. (2006). The Arabidopsis thaliana TIR-NB-LRR R-protein, RPP1A; protein localization and constitutive activation of defence by truncated alleles in tobacco and Arabidopsis. *Plant J.*, **47**: 829-840.

West, A.H. & Stock, A.M. (2001). Histidine kinases and response regulators proteins in two-component signalling systems. *Mol. Plant-Microbe Interact.*, **26**: 369-376.

Whisson, S.C., Boevink, P.C., Moleleki, L., Avrova, A.O., Morales, J.G., Gilroy, E.M., Armstrong, M.R., Grouffaud, S., van West, P., Chapman, S., Hein, I., Toth, I.K., Pritchard, L. & Birch, P.R.J. (2007). A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature*, **450**: 115-118.

Williams, S.J. (2009). Molecular insight into the activation of a plant disease resistance protein, PhD thesis, Flinders University, Adelaide, Australia.

Williams, S.J., Sohn, K.H., Wan, L., Bernoux, M., Sarris, P.F., Segonzac, C., Ve, T., Ma, Y., Saucet, S.B., Ericsson, D.J., Casey, L.W., Lonhienne, T., Winzor, D.J., Zhang, X., Coerdt, A., Parker, J.E., Dodds, P.N., Kobe, B. & Jones, J.D.G. (2014). Structural basis for assembly and function of a heterodimeric plant immune receptor. *Science*, **344**: 299-303.

Williams, S.J., Sornaraj, P., deCourcy-Ireland, E., Menz, R.I., Kobe, B., Ellis, J.G., Dodds, P.N. & Anderson, P.A. (2011). An autoactive mutant of the M flax rust resistance protein has a preference for binding ATP, whereas wild-type M protein binds ADP. *Mol. Plant-Microbe Interact.,* **24**: 897-906.

Yan, N., Chai, J., Lee, E.S., Gu, L., Liu, Q., He, J., Wu, J.W. & Kokel, D., Li, H., Hao, Q. Xue, D. & Shi, Y. (2005). Structure of the CED-4–CED-9 complex provides insights into programmed cell death in *Caenorhabditis elegans*. *Nature*, **437**: 831-837.

Yang, X., Chang, H.Y. & Baltimore, D. (1998). Essential Role of CED-4 oligomerisation in CED-3 activation and apoptosis. *Science*, **281**: 1355-1357.

Ye, Z., Lich, J.D., Moore, C.B., Duncan, J.A., Williams, K.L. & Ting, J.P.-Y. (2008). ATP binding by Monarch-1/NLRP12 Is critical for its inhibitory function. *Mol. Cell. Biol.*, **28**: 1841-1850.

Young, C.L., Britton, Z.T. & Robinson, A.S. (2012). Recombinant protein expression and purification: A comprehensive review of affinity tags and microbial applications. *Biotechnol. J.*, **7**: 620-634.

Zaitseva, J., Jenewein, S., Jumpertz, T., Holland, I.B. & Schmitt, L. (2005). H662 is the linchpin of ATP hydrolysis in the nucleotide-binding domain of the ABC transporter HlyB. *EMBO J.,* **24**: 1901-1910.

Zhang, Y., Dorey, S., Swiderski, M. & Jones, J.D. (2004). Expression of RPS4 in tobacco induces an AvrRps4-independent HR that requires EDS1, SGT1 and HSP90. *Plant J.,* **40**: 213-224.

Zurek, B., Bielig, H. & Kufer, T.A. (2011a). 'Cell-based reporter assay to analyse activation of Nod1 and Nod2' in J.P. Rast & J.W.D. Booth (eds.), *Immune Receptors: Methods and Protocols, Methods in Molecular Biology*, vol, 748, Springer Protocols, pp. 107-119.

Zurek, B., Proell, M., Wagner, R.N., Schwarzenbacher, R. & Kufer, T.A. (2011b). Mutational analysis of human NOD1 and NOD2 NACHT domains reveals different modes of activation. J. *Innate Immun.*, **18**: 100-111.