

# **Molecular Insight into the Activation of a Plant Disease Resistance Protein**

Simon J. Williams

A thesis submitted for the Degree of Doctor of Philosophy

School of Biological Sciences  
Faculty of Science and Engineering  
Flinders University, South Australia

September 2009

## TABLE OF CONTENTS

|  |      |
|--|------|
| LIST OF FIGURES AND TABLES .....   | VII  |
| DECLARATION .....  | VIII |
| ACKNOWLEDGMENTS .....  | IX   |
| ABSTRACT.....  | XI   |
| ABBREVIATION LIST .....  | XIII |
| CHAPTER 1: INTRODUCTION .....  | 1    |
| 1.1 Overview of Plant Innate Immunity.....                                 | 2    |
| 1.1.1 Plant innate immunity: a global view .....                           | 2    |
| 1.1.2 Amplitude of resistance .....  | 3    |
| 1.2 PAMP-Triggered Immunity .....  | 6    |
| 1.3 Pathogen Effectors .....   | 7    |
| 1.4 Effector-Triggered Immunity .....                                      | 10   |
| 1.4.1 R genes .....  | 10   |
| 1.4.1.1 The eLRR R genes .....   | 10   |
| 1.4.1.2 NBS-LRR.....   | 11   |
| 1.5 The Flax, Flax Rust System.....  | 12   |
| 1.5.1 The flax rust pathogen .....   | 12   |
| 1.5.2 Flax R genes and rust effector proteins.....                         | 13   |
| 1.6 Proposed Function of the NSB-LRR R Protein Domains.....                | 15   |
| 1.6.1 Leucine-rich repeat.....   | 15   |
| 1.6.2 Nucleotide binding site (NBS) .....                                  | 16   |
| 1.6.3 Amino terminal domain .....  | 20   |
| 1.7 Effector Perception by R Proteins .....                                | 21   |
| 1.7.1 Recognition of effector proteins by R proteins .....                 | 21   |
| 1.7.2 Indirect recognition: the guard and decoy hypothesis .....           | 21   |
| 1.7.3 Direct recognition .....   | 22   |
| 1.8 R protein Structure/Function .....                                     | 25   |
| 1.8.1 Structure studies of the NB-ARC region of Apaf-1 and CED-4.....      | 25   |
| 1.8.2 Intramolecular interactions regulate activation and signalling ..... | 26   |
| 1.8.3 A model of R protein activation .....                                | 27   |

|                   |   |           |
|-------------------|---|-----------|
| 1.8.4             | Functional studies of STAND proteins support the R protein molecular switch model ..... | 28        |
| <b>1.9</b>        | <b>Project Aims and Objectives .....</b>  | <b>31</b> |
| 1.9.1             | Aims .....  | 31        |
| <b>CHAPTER 2:</b> | <b>EXPERIMENTAL PROCEDURES .....</b>  | <b>33</b> |
| <b>2.1</b>        | <b>Materials .....</b>  | <b>34</b> |
| 2.1.1             | Culture media, solutions and buffers .....  | 34        |
| 2.1.2             | Yeast and bacterial stains used in this study .....                                     | 37        |
| 2.1.3             | Oligonucleotides .....  | 37        |
| 2.1.4             | Plasmids .....  | 38        |
| <b>2.2</b>        | <b>General Procedures .....</b>   | <b>39</b> |
| 2.2.1             | Preparation of electrocompetent and heat shock <i>E. coli</i> .....                     | 39        |
| 2.2.2             | Transformation of competent <i>E. coli</i> .....  | 39        |
| 2.2.2.1           | Electroporation of DH10B .....  | 39        |
| 2.2.2.2           | Heat Shock of BL21 pLysS .....  | 39        |
| 2.2.3             | Colony PCR .....  | 39        |
| 2.2.4             | Sodium do-decyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) .....            | 40        |
| 2.2.5             | Coomassie staining of proteins in polyacrylamide gels .....                             | 40        |
| 2.2.6             | Silver staining of proteins in polyacrylamide gels .....                                | 40        |
| 2.2.7             | Western blot .....  | 40        |
| 2.2.8             | Protein concentration .....   | 41        |
| 2.2.9             | Protein quantification .....  | 41        |
| <b>2.3</b>        | <b>Generation of Mutated Variants of M for Expression in <i>P. pastoris</i> .....</b>   | <b>41</b> |
| 2.3.1             | M truncated mutants .....   | 41        |
| 2.3.1.1           | Generation of M $\Delta$ TIR .....  | 41        |
| 2.3.1.2           | Generation of M $\Delta$ LRR .....  | 42        |
| 2.3.2             | Site-direct mutagenesis to produce rational point mutations in M .....                  | 42        |
| 2.3.3             | Preparation of electrocompetent <i>P. pastoris</i> .....                                | 43        |
| 2.3.4             | Electroporation of <i>P. pastoris</i> .....   | 43        |
| 2.3.5             | Screening and test expression of transformed <i>P. pastoris</i> .....                   | 43        |
| <b>2.4</b>        | <b>Growth and Expression of R proteins .....</b>  | <b>44</b> |
| 2.4.1             | Shaker flask expression .....   | 44        |
| 2.4.2             | Fermentation expression .....   | 44        |
| <b>2.5</b>        | <b>Cell Lysis .....</b>   | <b>45</b> |
| 2.5.1             | Cell disruptor .....  | 45        |
| 2.5.1.1           | <i>P. pastoris</i> .....  | 45        |
| 2.5.1.2           | Protein expression in <i>E. coli</i> .....  | 45        |
| 2.5.2             | French press .....  | 46        |

|  |  |           |
|--|--|-----------|
| <b>2.6</b>   | <b>Purification of M, L6 and M Variants.....</b>   | <b>46</b> |
| 2.6.1  | Gel filtration (GF).....   | 46        |
| 2.6.2  | Large scale purification .....   | 46        |
| 2.6.3  | Small scale purification .....   | 47        |
| <b>2.7</b>   | <b>Cloning, Expression and Purification of AvrM/avrM .....</b>   | <b>47</b> |
| 2.7.1  | Cloning of avrMC* into pET vector .....  | 47        |
| 2.7.2  | Transformation of BL21 pLysS with AvrM/avrM .....  | 48        |
| 2.7.3  | Growth and test expression .....   | 48        |
| 2.7.4  | Large scale expression .....   | 48        |
| 2.7.5  | Purification.....  | 49        |
| <b>2.8</b>   | <b>Biochemical Assays .....</b>  | <b>49</b> |
| 2.8.1  | ATPase assay.....  | 49        |
| 2.8.2  | ATP/ADP quantification assays using luminescence .....   | 49        |
| 2.8.3  | M protein quantification using VersaDoc imaging system.....  | 50        |
| <b>2.9</b>   | <b>Interaction Assays.....</b>   | <b>50</b> |
| 2.9.1  | Co-immunoprecipitation .....   | 50        |
| 2.9.1.1  | Immunoprecipitation using Protein G agarose beads.....   | 50        |
| 2.9.1.2  | Immunoprecipitation using protein A microbeads .....   | 51        |
| 2.9.2  | Nucleotide exchange assays .....   | 51        |
| <b>2.10</b>  | <b>Multiple Sequence Alignment of the NB-ARC Region of M, other R proteins, Apaf-1 and CED-4 .....</b> | <b>51</b> |
| <b>2.11</b>  | <b>Structure Model of the NB-ARC Domain of M .....</b>   | <b>52</b> |
| <b>CHAPTER 3: PURIFICATION OF THE M AND L6 FLAX RUST RESISTANCE PROTEINS EXPRESSED IN <i>PICHTIA PASTORIS</i>.....</b> |  | <b>53</b> |
| <b>3.1</b>   | <b>Introduction.....</b>   | <b>54</b> |
| <b>3.2</b>   | <b>Results .....</b>   | <b>55</b> |
| 3.2.1  | Three-step purification of M and L6 .....  | 55        |
| 3.2.2  | Small scale expression and one-step purification .....   | 59        |
| 3.2.3  | Potential issues in the functionality of the purified proteins.....                                    | 62        |
| 3.2.4  | Full-length M mutants can be purified from <i>P. pastoris</i> .....                                    | 66        |
| <b>3.3</b>   | <b>Discussion .....</b>  | <b>68</b> |
| 3.3.1  | Expression and purification of flax R proteins facilitates <i>in vitro</i> studies .....               | 68        |
| 3.3.2  | Physical properties of the flax R proteins cause self-association and aggregation .....                | 69        |
| 3.3.3  | Concluding remarks.....  | 70        |
| <b>CHAPTER 4: BIOCHEMICAL INVESTIGATION OF THE NB-ARC DOMAIN OF PURIFIED FLAX RUST R PROTEINS .....</b>                |  | <b>71</b> |

|   |  |            |
|---|--|------------|
| <b>4.1</b>  | <b>Introduction .....</b>  | <b>72</b>  |
| <b>4.2</b>  | <b>Results .....</b>   | <b>76</b>  |
| 4.2.1   | M and L6 proteins are purified with ADP bound.....   | 76         |
| 4.2.2   | Mutations within the NB-ARC domain that affect protein function have different nucleotide binding capacities   | 82         |
| 4.2.3   | ATP hydrolysis analysis of M proteins .....  | 87         |
| 4.2.4   | Do autoactive M proteins form oligomers? .....   | 91         |
| <b>4.3</b>  | <b>Discussion.....</b>   | <b>93</b>  |
| 4.3.1   | ADP bound is the “OFF” state .....   | 93         |
| 4.3.2   | ATP bound is the “ON” state .....  | 94         |
| 4.3.3   | Comparisons with published <i>in vitro</i> studies of R proteins .....   | 100        |
| 4.3.4   | Oligomerisation in the ATP bound state?.....   | 101        |
| 4.3.5   | Conclusion .....   | 102        |
| <br><b>CHAPTER 5: A STUDY OF M AND AVR M INTERACTION <i>IN VITRO</i>.....</b> |  | <b>103</b> |
| <b>5.1</b>  | <b>Introduction .....</b>  | <b>104</b> |
| <b>5.2</b>  | <b>Results .....</b>   | <b>106</b> |
| 5.2.1   | Purification of AvrM, avrM and avrMC* .....  | 106        |
| 5.2.2   | Co-immunoprecipitation experiments using purified recombinant proteins support the hypothesis of a direct protein-protein interaction between M and AvrM ..... | 111        |
| 5.2.3   | <i>In vitro</i> experiments fail to identify nucleotide exchange in M induced by AvrM interaction .....  | 117        |
| <b>5.3</b>  | <b>Discussion.....</b>   | <b>119</b> |
| 5.3.1   | Purification of flax rust effector proteins.....   | 119        |
| 5.3.2   | Co-IP studies support a direct interaction between M and AvrM.....   | 120        |
| 5.3.3   | The effect of the effector .....   | 120        |
| <br><b>CHAPTER 6: CONCLUSIONS AND FUTURE WORK .....</b>                       |  | <b>123</b> |
| <b>6.1</b>  | <b>Overview .....</b>  | <b>124</b> |
| <b>6.2</b>  | <b>Model of M Protein Activation .....</b>   | <b>125</b> |
| <b>6.3</b>  | <b>A Model Stimulates Experimental Questions for Future Research .....</b>   | <b>127</b> |
| 6.3.1   | Determining the <i>in planta</i> phenotypes .....  | 127        |
| 6.3.2   | Is ATP hydrolysis and/or oligomerisation involved in M protein function? .....   | 127        |
| 6.3.3   | Interaction and recognition .....  | 129        |
| 6.3.4   | Further characterisation of the NB-ARC domain, including other R proteins .....  | 131        |
| 6.3.5   | Intramolecular interactions and protein structure .....  | 132        |
| 6.3.6   | Intermolecular interactions and signalling .....   | 133        |

|                           |   |            |
|---------------------------|---|------------|
| 6.4                       | Conclusion .....  | 135        |
| <b>APPENDIX.....</b>      |   | <b>137</b> |
| Appendix 1:               | Sequencing Oligonucleotides .....   | 137        |
| Appendix 2:               | Gel Filtration Calibration Curves .....   | 138        |
| Appendix 3:               | Overview of the Truncation/Mutations of M Expressed in <i>P. pastoris</i> ..... | 139        |
| Appendix 4:               | Purification of M $\Delta$ LRR .....  | 140        |
| Appendix 5:               | Supplementary Information for ATP and ADP Quantification Experiments .....      | 141        |
| Appendix 6:               | AvrM/avrM/avrMC* Alignment .....  | 142        |
| Appendix 7:               | Improved Expression.....  | 143        |
| <b>BIBLIOGRAPHY .....</b> |   | <b>145</b> |

# List of Figures and Tables

|              |  |     |
|--------------|--|-----|
| FIGURE 1.1:  | GLOBAL VIEW OF THE PLANT IMMUNE SYSTEM   | 5   |
| TABLE 1.1:   | PATHOGEN RECOGNITION RECEPTORS (PRRS) AND THEIR PAMP TARGETS                               | 9   |
| FIGURE 1.2:  | A MODEL OF A FLAX CELL, WITH OR WITHOUT L5, AND ITS RESPONSE TO FLAX RUST CARRYING AVR567A | 14  |
| TABLE 1.2:   | NUCLEOTIDE BINDING SITE MOTIF CONSENSUS SEQUENCES  | 18  |
| FIGURE 1.3:  | AN NB-ARC ALIGNMENT  | 19  |
| FIGURE 1.4:  | THE MOLECULAR SWITCH MODEL FOR R PROTEIN ACTIVATION  | 30  |
| TABLE 2.1:   | CULTURE MEDIA  | 34  |
| TABLE 2.2:   | SOLUTIONS AND BUFFERS  | 35  |
| TABLE 2.3:   | YEAST AND BACTERIAL STRAINS  | 37  |
| TABLE 2.4:   | CLONING OLIGONUCLEOTIDES   | 37  |
| TABLE 2.5:   | PLASMIDS USED IN THIS STUDY  | 38  |
| FIGURE 3.1:  | EXPRESSION AND PURIFICATION OF RECOMBINANT M   | 57  |
| FIGURE 3.2:  | EXPRESSION AND PURIFICATION OF RECOMBINANT L6  | 58  |
| FIGURE 3.3:  | SMALL SCALE EXPRESSION AND PURIFICATION OF M FOR BIOCHEMICAL ANALYSIS                      | 61  |
| FIGURE 3.4:  | PURIFICATION OF M $\Delta$ TIR AND GEL FILTRATION ANALYSIS                                 | 63  |
| FIGURE 3.5:  | IMIDAZOLE AFFECTS M $\Delta$ TIR MONOMERIC STATE   | 64  |
| FIGURE 3.6:  | GEL FILTRATION COMPARISON OF M, L6, AND L6 $\Delta$ TIR                                    | 65  |
| FIGURE 3.7:  | PURIFIED M FULL-LENGTH MUTANTS   | 67  |
| TABLE 4.1:   | THE M AND M-MUTANTS ANALYSED <i>IN-VITRO</i>   | 75  |
| FIGURE 4.1:  | M AND L6 PROTEINS AFTER NIA PURIFICATION ARE ASSOCIATED WITH ADP                           | 79  |
| FIGURE 4.2:  | GEL FILTRATION FRACTIONS CONTAINING M AND M $\Delta$ TIR ALSO CONTAIN ADP                  | 80  |
| FIGURE 4.3:  | PERCENTAGE OF M AND M $\Delta$ TIR PROTEINS OCCUPIED WITH ATP OR ADP                       | 81  |
| FIGURE 4.4:  | PURIFICATION OF M AND M MUTANTS FOR ATP AND ADP QUANTIFICATION                             | 84  |
| FIGURE 4.5:  | ADP BINDING IS DEPENDENT ON A FUNCTIONAL P-LOOP  | 85  |
| FIGURE 4.6:  | PUTATIVE AUTOACTIVE MUTANTS ARE PURIFIED WITH ATP BOUND                                    | 86  |
| FIGURE 4.7:  | ATPASE ACTIVITY IN GF FRACTIONS OF M AND M-MUTANTS   | 88  |
| FIGURE 4.8:  | M AND M <sup>K286L</sup> PROTEINS HAVE INDISTINGUISHABLE HYDROLYSIS ACTIVITIES             | 89  |
| FIGURE 4.9:  | LONG WASH NIA PURIFICATION DOES NOT REMOVE BACKGROUND ATPASE ACTIVITY                      | 90  |
| FIGURE 4.10: | GF ANALYSIS OF M <sup>D555V</sup> AND M <sup>K286L+D555V</sup>                             | 92  |
| FIGURE 4.11: | A STRUCTURAL MODEL OF THE NB-ARC DOMAIN OF M   | 99  |
| FIGURE 5.1:  | PURIFICATION OF AVR <sub>M</sub>   | 108 |
| FIGURE 5.2:  | PURIFICATION OF AVR <sub>M</sub>   | 109 |
| FIGURE 5.3:  | PURIFICATION OF AVR <sub>M</sub> C*  | 110 |
| FIGURE 5.4:  | CO-IMMUNOPRECIPITATION OF M AND AVR <sub>M</sub>   | 114 |
| FIGURE 5.5:  | CO-IMMUNOPRECIPITATION OF M, AVR <sub>M</sub> AND AVR <sub>M</sub> C*                      | 115 |
| FIGURE 5.6:  | CO-IMMUNOPRECIPITATION OF M, M <sup>K286L</sup> , AVR <sub>M</sub> AND AVR <sub>M</sub> C* | 116 |
| FIGURE 5.7:  | NUCLEOTIDE EXCHANGE ASSAY  | 118 |
| FIGURE 6.1:  | A WORKING MODEL OF M ACTIVATION  | 126 |

# Declaration

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

Simon J. Williams

# Acknowledgments

It is with great pleasure that I recognise the important people in my life that have helped bring this thesis to fruition.

Throughout this study my supervisor Dr. Peter Anderson has been an excellent supervisor, superb mentor and great friend. Passion, determination and commitment are words that I would use to describe Peter's quest to further understand the molecular function of an R protein. During my time in the Anderson laboratory I believe I have taken on these attributes towards this topic and this has been critical in the completion of the experiments and presentation of this thesis. Peter (and the Anderson lab), some time ago, decided to take a somewhat difficult scientific path due to Peter's genuine interest and commitment to this work. It has required immense patience and experienced a number of low points; however, I believe our efforts as a laboratory group are beginning to be justified by the scientific breakthroughs that we are now generating. Also I'd like to say a huge thank you, for your advice, prompt feedback and editing assistance, throughout the thesis write-up stage.

I would also like to express thanks to my co-supervisors Dr. Ian Menz and Dr. Simon Schmidt. Without Simon this project would not have been possible and he has provided me with sound advice and encouragement throughout this study. Ian's knowledge of biochemistry is something that I aspire to and without his experience and advice much of this work would not have been achieved.

A number of collaborators have also helped make this work possible. I am particularly grateful to Dr. Peter Dodds for supplying me with various scientific materials, advice, opportunities, and encouragement. Also thank you to the Brisbane group; Professor Bostjan Kobe, Dr. Andi Wang and Thomas Ve, for their continued support throughout my thesis and the opportunities that have arisen due to their involvement in this project.

I am appreciative to the many members of the Anderson laboratory for providing an excellent working environment that has stimulated my growth as a researcher. I am particularly grateful for the help and support supplied by Pradeep Sornaraj and Emma deCourcy-Ireland. You both have shown guts and determination to undertake PhD projects that focus on R protein activation and a large amount of trust in the research system that we have been developing. I'm confident that this decision will prove a great one for your future. I am also very thankful to Emma for her assistance in the editing process.

Throughout my time as a student within the School of Biological Sciences I have made many great friends. I would like to mention Mel P, Bart, Drew and Nick for their assistance in my

experimental design and the numerous stimulating scientific discussions that have undoubtedly assisted in my development. I would also like to express gratitude to Keryn. We have shared countless cups of tea and chats throughout the entire PhD period. Your discussions frequently provided a welcome distraction from the realities of a PhD project and lot of laughter.

Thank you also to my many supportive friends, in particular Brad (Bunga); thanks for being a best mate and a lawyer/investment banker. Your generosity at the bar over the last four years and support of a struggling PhD student has been greatly appreciated (and noted), as has your continued friendship and support.

Thank you to my incredibly supportive family and extended family, in particular my parents John and Monica and sister Renée. Your never-ending, emotional and at times, financial support throughout this endeavour has enabled me to complete this thesis.

And finally to my partner in life, Jess. I'm unable to find the words that express my gratitude for your continued love and support over the last 4 years. Together I think we have proven that it is possible to maintain a relationship and two PhD's concurrently. Love you always.

# Abstract

A plant's ability to detect an invading pathogen and circumvent a subsequent disease state is essential for its survival. Disease resistance, and the mechanisms behind it, are thus of critical importance. The pioneering work of Harold Flor, using the interaction between flax and the flax rust fungus, *Melampsora lini*, demonstrated that this ability to detect and resist the infection of a specific pathogen rests with two critical genes; a resistance (*R*) gene in the plant and a corresponding avirulence (*Avr*) gene in the pathogen. This, so called 'gene-for-gene' model, has subsequently been shown to apply in many other plant-pathogen interactions and has spawned considerable research efforts directed towards understanding the molecular basis of host-pathogen interactions and the consequential disease resistance response. Using the flax-flax rust pathosystem, and utilising a biochemical approach, this research has endeavoured to further the current understanding of the molecular basis of the interaction between plants and pathogens, with a particular focus on R protein function.

Chapter 3 describes the production of soluble, recombinant flax R proteins, M and L6, using the *Pichia pastoris* expression system. These flax R proteins can be purified from total cell lysates utilising a number of chromatography techniques. Following nickel affinity chromatography, concentration of protein in the presence of imidazole, leads to aggregation. This, however, can be alleviated by lowering the imidazole concentration prior to the protein concentration step. This fine tuning of the purification protocol enabled the expression and enrichment of near full-length and truncated versions of M and L6, and rational point mutations of M.

Utilising this expression and purification system, Chapter 4 presents a detailed functional study of the flax M protein, with particular focus on mutations that cause autoactivity and inactivity. These mutations were generated in the NB-ARC region of M with predicted loss- or gain-of-function consequences, as determined from the results of the *in planta* phenotypes of analogous mutations in other R proteins, in particular the flax L6 protein. Nucleotide quantification of purified wild type M and L6 demonstrated that these proteins are associated with ADP. Analysis of proteins with mutations within the NB-ARC domain demonstrated that this ADP binding is dependent on a functional P-loop in the NB subdomain. Mutations within the MHD motif and motif VIII that are predicted to result in an autoactive *in planta* phenotype, have more ATP associated with purified protein preparations in comparison to wild type. Taken together, these results further support the model that R proteins act as a molecular switch, whereby the

inactive form of the protein is ADP bound, while the active conformation of the protein is ATP bound.

Prior to this study yeast two hybrid analysis had demonstrated that a direct interaction between M and AvrM occurs. To investigate the interaction, and the consequence of interaction, between M and AvrM proteins *in vitro*, an expression and purification protocol was generated for AvrM (and variants) in Chapter 5. Here, a direct protein-protein interaction was supported by co-immunoprecipitation of purified M and AvrM proteins. The interaction that M has with AvrM is dependent on a functional P-loop and therefore presumably requires the presence of a bound nucleotide. The preferred model of R protein activation suggests that interaction with an effector causes the R protein to exchange its bound nucleotide from ADP to ATP. To determine if AvrM could induce nucleotide exchange, ADP/ATP exchange assays were performed, however, the results of this study were inconclusive. It is possible that nucleotide exchange is not the mechanism of activation of the flax M protein; although, it is equally likely that the conditions in the *in vitro* assay were not conducive for exchange to occur or that other proteins are needed to facilitate the exchange event. Whilst this study adds further proof to the theory of a direct interaction between flax rust effectors and their corresponding R proteins, the molecular effect that this event has on the R protein is yet to be understood.

In summary, only a small number of biochemical investigations of R proteins have been published, nevertheless, they have provided highly revealing information regarding R protein function. Utilising an *in vitro* approach, the results from this thesis provide further insight into the function and interaction between flax R proteins, and their effectors. It is hoped that the techniques developed and presented in this thesis will assist, and inspire, future *in vitro* investigations of flax R protein molecular function, and thus contribute to a wider understanding of plant disease resistance.

# Abbreviation List

|                    |  |
|--------------------|--|
| ADP                | Adenosine 5'-diphosphate   |
| Apaf-1             | Apoptotic protease-activating factor 1   |
| Avr                | Avirulence   |
| ATP                | Adenosine 5'-triphosphate  |
| CC                 | Coiled coil  |
| CED-4              | Cell death protein 4   |
| CEX                | Cation exchange  |
| Co-IP              | Co-immunoprecipitation   |
| eLRR               | extracellular LRR  |
| ER                 | Endoplasmic Reticulum  |
| ETI                | Effector triggered immunity  |
| GF                 | Gel filtration   |
| HR                 | Hypersensitive response  |
| LRR                | Leucine-rich repeat  |
| MAPK               | Mitogen-activated protein kinase   |
| NB-ARC             | Nucleotide-binding adaptor shared by APAF-1, certain <i>R</i> gene products, and CED-4 |
| NBS                | Nucleotide binding site  |
| NOD                | Nucleotide oligomerisation domain  |
| NiA                | Nickel-metal ion affinity  |
| NLR                | NOD-LRR  |
| PAMPs              | Pathogen associated molecular patterns   |
| PRRs               | Pattern recognition receptors  |
| PTI                | PAMP triggered immunity  |
| R gene / R protein | Resistance gene/Resistance protein   |
| RLK-eLRR           | Receptor-like kinases with extracellular leucine-rich repeats                          |
| SAR                | Systemic acquired resistance   |
| STAND              | Signal transduction ATPases with numerous domains                                      |
| T3SS               | Type-III secretory system  |
| TIR                | Toll/Interleukin-1 Receptor  |
| TLR                | Toll-like receptors  |
| TMV                | Tobacco mosaic virus   |
| Y2H                | Yeast two hybrid   |



# Chapter 1: Introduction

From this introductory chapter the following manuscript was produced for an invited review, this is currently submitted and in review:

Williams, S. J., Anderson, P. A., Kobe, B., Ellis, J. G. & Dodds, P. N. (2009) The Molecular Basis of Rust Resistance in Flax, *The Americas Journal of Plant Science and Biotechnology*, in review.

## **1.1 Overview of Plant Innate Immunity**

Plants lack an adaptive immune system and rely on the capabilities of individual cells to detect and respond to invading pathogenic agents and thus prevent infection and disease. It is this innate immunity that enables them to survive and thrive in an environment where they encounter a broad range of potential pathogens with diverse life styles. A number of models have been proposed in recent years that attempt to explain how this innate immune system has developed during the evolutionary battle between pathogen and plant (Bent and Mackey, 2007, Jones and Dangl, 2006, Chisholm *et al.*, 2006). This chapter will introduce the data that supports, and in some cases refutes, these models, and will ultimately focus on a crucial plant protein that coordinates disease resistance.

### **1.1.1 Plant innate immunity: a global view**

A plant has several layers of defence that it can employ in order to resist disease from pathogenic agents. These layers can be broadly divided into passive and active defence. Passive defence mechanisms include structural, chemical and biological plant components, such as the leaf waxy cuticle, lignified cell walls and anti-microbial/fungal compounds. These non-targeted defences provide the front line of plant defence and most likely prevent the vast majority of potential pathogen-related infections. Active defence mechanisms are those that are induced by the plant and thus invoke the need for some form of plant cellular machinery to cover roles of pathogen surveillance, detection, signal activation and response by the plant cell. It is this coordinated machinery that controls what is termed the plant innate immune system.

Conceptually, a plant's innate immune system can be divided into two lines of defence. The first line is historically known as basal defence. The basal defence mechanism utilises a broad detection system that is targeted towards pathogen/microbe associated molecular patterns (PAMPs or MAMPs). As a result, this line of defence is now commonly referred to as PAMP or MAMP triggered immunity (PTI). Detection is, in most cases, facilitated by extracellular transmembrane pattern recognition receptors (PRRs) which survey the apoplast on the look out for PAMPs. Upon detection of a molecular pattern, these PRRs are activated and communicate signals across the plasma membrane to the inside of the plant cell, switching on defence related pathways to achieve effective resistance.

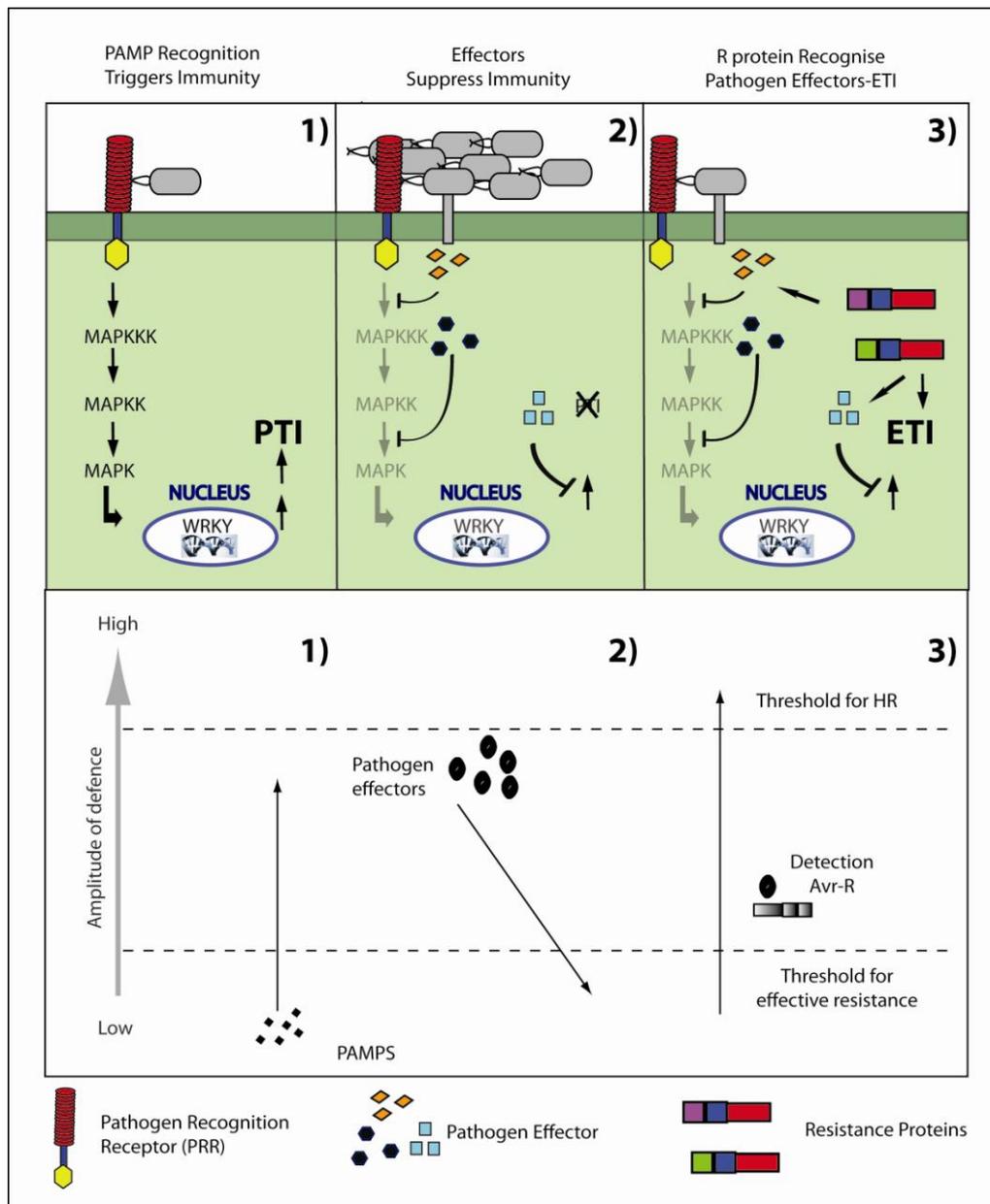
Certain, more specialised pathogens have devised infection strategies that utilise effector molecules to aid in plant colonisation. Effectors are secreted from the pathogen and function in the apoplast or enter the plant cell to affect intracellular targets. Effectors from bacteria have been demonstrated to act by suppressing aspects of PTI. In response, however, plants have

evolved a second line of defence where specialised host proteins, known as resistance (R) proteins, can detect these specific pathogen effectors. Aptly named effector-triggered immunity (ETI), but historically known as R gene mediated resistance, these R proteins detect the pathogen and coordinate the disease resistance response. As one would expect, natural selection pressures force pathogens to diversify or shed recognisable effectors to evade ETI. This, in turn, has led to the diversification of R gene specificities to recognise modified or alternative effectors. This cyclical evolutionary battle between plant and pathogen continues to generate many pathogen strains with different combinations of effectors, and variant genotypes of plants with different combinations of R genes. Modern agricultural practices have increased the stakes in this evolutionary battle. A number of elegant and informative explanations of plant innate immunity have been presented in recent reviews and Figure 1.1 demonstrates these pictorially (Jones and Dangl, 2006, Bent and Mackey, 2007, Chisholm et al., 2006).

### 1.1.2 Amplitude of resistance

ETI is a more targeted plant defence strategy and tends to induce a stronger response than PTI. ETI is typified by the hypersensitive response (HR), which often culminates in programmed cell death of the infected cell, and to a limited extent, neighbouring cells (Greenberg, 1997, Jones and Dangl, 2006, Tao et al., 2003). Early studies of HR identified a number of inducible defence mechanisms. Ion fluxes, such as  $\text{Ca}^{2+}$ , into the cytoplasm are known to play an early role. Calcium has been shown to be associated with the production of reactive oxygen species (ROS) (causing an oxidative burst), and phytoalexin production (Greenberg, 1997, Heath, 2000). Phytoalexins (plant antibiotics), and inducible defence-related genes, have been linked directly to limiting pathogen growth. Signalling molecules, such as salicylic acid (SA), jasmonic acid (JA), ethylene (ET), nitric oxide (NO) and ROS, have all been shown to contribute to plant defence (Hammond-Kosack and Parker, 2003). Interestingly, a number of these same pathways are involved in PTI, suggesting significant signalling overlap between PTI and ETI (reviewed by (Altenbach and Robatzek, 2007, Schwessinger and Zipfel, 2008)). PTI has been shown to be tightly regulated, and there is growing evidence to suggest that for activation of ETI this negative regulation is released. This could explain, at least in part, the difference in response amplitude between ETI and PTI (Schwessinger and Zipfel, 2008). The importance of cell death during ETI is a contentious issue; however, plants that have been infected by a pathogen that induce ETI have a heightened state of readiness for further attack by the same or different pathogen(s). This general, long lasting defence response is referred to as systemic acquired resistance (SAR) (Gaffney et al., 1993, Ryals et al., 1996, Verberne et al., 2003). This thesis is concerned

predominately with the processes involved in ETI with a focus on R protein activation; however PTI will be discussed in more detail below.



**Figure 1.1: Global view of the plant immune system**

This figure is adapted from (Chisholm et al., 2006, Jones and Dangl, 2006). It illustrates the potential evolution of resistance (in this case for a bacterial pathogen) combined with the amplitude of defence signalling within the plant. 1) Utilising its PRRs, the plant can recognise PAMPs and signal for the amplification of defence pathways for effective resistance (PTI). 2) Accordingly, pathogens have evolved genes and pathways to secrete proteins (effectors) into the cell where the virulence function enables the pathogen to evade or prevent PTI. 3) In order to maintain a competitive advantage, plants have evolved resistance genes, the products of which detect pathogen effectors and mediate signalling pathways leading to hypersensitive response and effector triggered immunity (ETI). Points 1, 2 and 3 are discussed in Sections 1.2, 1.3 and 1.4 respectively.

## 1.2 PAMP-Triggered Immunity

Plant PTI is facilitated by the perception capabilities of PRRs, which survey the extracellular space surrounding the cell wall on the look out for PAMPs. To date, only a handful of PRRs have been found (summarised Table 1.1), however, it is anticipated that this represents only a fraction of what exist (Nurnberger et al., 2004). The recognition of two common bacterial elements, flagellin and elongation factor TU (EF-TU), is controlled by different receptor-like kinases with extracellular leucine rich repeats (RLK-eLRR) (Chinchilla et al., 2006, Zipfel et al., 2006). LRRs are protein elements commonly found to facilitate intra- or inter-molecular protein-protein interactions (Kobe and Kajava, 2001). Their importance in pathogen perception and innate immunity in mammals, as a part of the Toll-like receptors (TLR), has been well studied. It is this LRR element found in plant PRRs and some R proteins, that is one of a number of similarities that can be found between human and plant innate immunity (reviewed by (Nurnberger et al., 2004, Rairdan and Moffett, 2007). The RLK-eLRR is, however, not the only class of PRRs. Other receptors identified so far, range from receptor-like proteins (RLP) with short cytoplasmic tails, to soluble proteins expressed extracellularly (Table 1.1). It is likely that such PRRs would associate with other RLKs or RLPs to facilitate further signalling (Zipfel, 2008). This has been demonstrated for the RLK-eLRR named BRI1-associated kinase 1 (BAK1). BAK1, by interaction with BRI1, mediates the signalling of brassinosteroid hormones, which control many aspects of plant growth and development. Studies have demonstrated that BAK1 is also involved in signalling of both FLS2 and EFR (Chinchilla et al., 2007, Zipfel, 2008). Fundamentally, the role of the PRR is to perceive the potential pathogen and instigate the relay of a signal into the plant cell where it can be acted on.

Once the signal is transmitted to the inside of the cell signalling pathways are required to transduce it to the nucleus, where reprogramming of transcription enables the necessary action to be taken. In *Arabidopsis thaliana* mitogen-activated protein kinase (MAPK) cascades have been demonstrated to transfer the signal after flagellum perception by FLS2, which leads to the activation of WRKY transcription factors that assist in transcriptional reprogramming of the cell (Asai et al., 2002). This cascade facilitates resistance to both bacterial and fungal pathogens (Asai et al., 2002), confirming MAPKs as the signal transducers of PTI. Interestingly, MAPK cascades are also involved in signal transfer in mammalian innate immunity, identifying yet another link between plant and mammalian innate immunity (Nurnberger et al., 2004).

PTI is an ever expanding field of research and a number of recent reviews explore aspects of perception, signal cascades and pathway cross talk between PTI and ETI in more detail (Bent and Mackey, 2007, Nurnberger et al., 2004, Zipfel, 2008). Although PTI provides an important level of protection for a plant, some pathogens have evolved mechanisms to overcome

it. In many cases such pathogens secrete a suite of molecules into the infected plant cell to circumvent PTI and cause disease.

### **1.3 Pathogen Effectors**

Plant pathogens have a wide range of infection strategies, however an emerging commonality between these pathogens is the secretion of proteins (including peptides and small compounds) that alter plant cell structure and function (Hogenhout et al., 2009, Panstruga and Dodds, 2009). These factors are collectively termed effectors. Effectors are generally thought of as pathogen elements that promote virulence during infection and have been found to function in both the apoplast and within the plant cell (reviewed by (Kamoun, 2006, Hogenhout et al., 2009)). In cases where the effector is recognised by an R protein, the effector is said to be an avirulence (Avr) protein and its intended virulence function is compromised. The term avirulence, or Avr, is a genetic term that was first used to describe the phenotypic consequence on the pathogen when resistance was triggered in the plant. More recently, the term effector has been used to describe pathogen-derived proteins which include the products of Avr genes (Hogenhout et al., 2009). For the purposes of this thesis, the term “effector” will be used to indicate the product of an avirulence or Avr gene.

A large proportion of effectors have virulence functions within the plant cell, so understanding the mechanisms used by pathogens to internalise these effectors is very important. Perhaps the most well characterised delivery system to date is that of gram negative bacterium *Pseudomonas syringae*. These bacteria use a type-III secretory system (T3SS) to transport effectors into the cell where they can then attack host targets. A gene cluster, called *Hrp* (hypersensitive response and pathogenicity), encodes the machinery for a T3SS which includes the formation of an infection pilus to enable effector transfer (reviewed by (Büttner and Bonas, 2002, Galan and Collmer, 1999, Staskawicz et al., 2001)). Oomycetes provide the most well studied eukaryotic example of effector translocation. Rather than using a host derived secretory system; these pathogens encode effectors with specialised sequences that facilitate their translocation into a plant cell. Those that undergo plant cell internalisation have two conserved sequence motifs, RxLR and EER, which have been implicated in translocation (reviewed by (Birch et al., 2006, Kamoun, 2006, Morgan and Kamoun, 2007)). Recently, both motifs were found to be required for effector delivery by the potato blight pathogen *Phytophthora infestans* (Whisson et al., 2007). The sequenced genomes of a number of oomycetes, including *P. sojae*, *P. ramorum*, and *P. infestans*, identified 400, 314 and 425 genes respectively, that encode secreted RxLR-EER classed-proteins (Tyler et al., 2006, Whisson et al., 2007). This

suggests that during infection, oomycetes have the potential to deposit a large and diverse array of effectors into the plant cell to aid in infection and help promote disease.

Research interests have now turned towards defining the potential virulence function of known effectors and also to define their host cell targets (Block et al., 2008). Suppression of host defence mechanisms seems an obvious target and a number of examples of this are emerging. He et al., (2006) demonstrated that two *P. syringae* effectors, AvrPto and AvrPtoB, were capable of suppressing early defence gene transcription and MAPK signalling (He et al., 2006). In fact, suppression of aspects of PTI appears to be a primary function of many characterised effectors (reviewed by (Jones and Dangl, 2006, Hogenhout et al., 2009)). It is also clear that effectors change aspects of plant development (Kay et al., 2007) and more recently a fungal effector was shown to suppress ETI (Houterman et al., 2008). The roles of pathogen effectors in host manipulation come from a range of functions including transcription regulation, protease activity and inhibition, protein degradation and protein phosphorylation (reviewed by (Block et al., 2008)). At present, however, a large number of cloned effectors still have not been characterised with regard to host cell targets and virulence function. Understanding the molecular function of many other effectors, and the common trends in their action (if any), will undoubtedly be essential if we are to fully understand the mechanisms used by different pathogens in plant infection and colonisation.

While effectors enhance infection, they also present potential targets for detection by their hosts. Consequently, an entire branch of plant immunity is based on effector perception for the activation of disease resistance. It is this function that is controlled by R proteins and is the focus of this thesis.

| PRR   | Structure  | Plant  | PAMP                         | Pathogen  | Reference   |
|---|--|--|------------------------------|-----------|---|
| FLS2<br>(Flagellin Sensing 2)                           | Receptor Like Kinase with extracellular LRR (RLK-eLRR)                                   | Arabidopsis<br>Tomato<br><i>N. benthamiana</i> | Flagellin                    | Bacteria  | (Chinchilla et al., 2006, Gomez-Gomez et al., 2001) |
| EFR<br>(EF-TU Receptor)                                 | (RLK-eLRR)   | Arabidopsis                                    | Elongation factor-TU (EF-TU) | Bacteria  | (Zipfel et al., 2006)                               |
| LeEIX 1 and LeEIX2<br>(Ethylene-inducing xylanase)      | Receptor Like Proteins (RLP) with short cytoplasmic tail                                 | Tomato   | Xylanase                     | Fungi     | (Ron and Avni, 2004)                                |
| chitin oligosaccharide elicitor binding protein (CEBiP) | A transmembrane protein with two extracellular LysM domains and a short cytoplasmic tail | Rice   | Chitin                       | Fungi     | (Kaku et al., 2006)                                 |
| CERK1   | A receptor like kinase with three LysM domains   | Arabidopsis                                    | Chitin                       | Fungi     | (Miya et al., 2007)                                 |
| b-glucan-binding protein (GBP)                          | Soluble protein (no transmembrane domain)  | Soybean  | Hepta-glucan                 | Oomycetes | (Fliegmann et al., 2004)                            |

**Table 1.1: Pathogen recognition receptors (PRRs) and their PAMP targets**

Outline of known PRRs including their plant origin and the PAMP targets and pathogens they defend against.

## 1.4 Effector-Triggered Immunity

Effector-triggered immunity (ETI) is an inducible defence mechanism mediated by plant disease resistance (R) gene products. The first genetic interpretation of ETI was described by Flor in the 1930s-1950s while investigating the genetic basis of resistance in flax (*Linum usitatissimum*) to the fungal pathogen *Melampsora lini* (Flor, 1956). This investigation led to the development of the gene-for-gene theory; whereby, plant resistance to a specific pathogen is achieved only when a *R gene* in the plant and a corresponding avirulence (*Avr*) gene within the pathogen are present (Flor, 1956, Flor, 1971). The absence of either gene, ultimately leads to the plant's susceptibility to infection and the development of disease symptoms. R genes have been identified that control resistance to an array of pathogens including bacteria, viruses, fungi, oomycetes, nematodes and insects. Identifying and cloning both resistance and avirulence genes and investigating the interplay between the gene products has become a major research objective for plant scientists, and is the focus of this thesis.

### 1.4.1 R genes

R genes have been reviewed extensively since the cloning of the first resistance gene, *Pto* (Martin *et al.*, 1993), which controls resistance to tomato bacterial speck disease. Currently over 40 isolated R genes from numerous plant species have been cloned (reviewed by (Dangl and Jones, 2001, Ellis *et al.*, 2000, Hammond-Kosack and Jones, 1997, Chisholm *et al.*, 2006)). With such a number of R genes now isolated and cloned, attention has now focussed towards understanding how the proteins they encode perceive effectors and activate signal pathways to achieve ETI. Bioinformatic tools have aided in this objective by assigning R genes into structural classes depending on their domain organisation. Despite some outliers, R genes can be divided into two broad classes; genes that encode products with extracellular leucine-rich repeat (eLRR), presumably involved in the recognition of effectors in the apoplast, and those that encode tri-domain intracellular nucleotide binding (NBS)-LRR proteins, shown to recognise internalised effectors protein (Chisholm *et al.*, 2006).

#### 1.4.1.1 The eLRR R genes

The R genes containing eLRRs can be further subdivided depending on the presence and origin of their intracellular signalling domain (Chisholm *et al.*, 2006). They share similarities in domain organisation with a number of the PRRs (see 1.2), and also draw comparison with pathogen receptors from mammalian innate immunity (Nurnberger *et al.*, 2004). Arguably, the most well characterised example of this broad class are the tomato *Cf*-genes that confer resistance to the

tomato leaf mould pathogen, *Cladosporium fulvum* (reviewed by (Rivas and Thomas, 2005)). This thesis, however, involves the study of an NBS-LRR protein, and it is this class of R proteins that will therefore be the main focus of this chapter.

#### **1.4.1.2 NBS-LRR**

The NBS-LRR proteins are generally trimodular and are distinguished by either a coiled-coil (CC) or Toll-interleukin 1 receptor (TIR)-like domain at the N-terminus. The NBS-LRR class is the most predominant class of plant disease R proteins, and the genes which encode NBS-LRR proteins account for approximately 150, 400 and over 500 genes in the *A. thaliana*, poplar and rice genomes, respectively (Monosi *et al.*, 2004, Tuskan *et al.*, 2006, Meyers *et al.*, 2003). It is clear that NBS-LRR proteins are central to disease resistance, as members of other R protein classes have been demonstrated to rely on NBS-LRR proteins for function and/or immunity. For example, the tomato cytoplasmic serine/threonine protein kinase resistance protein, Pto, requires Prf, an NBS-LRR-like protein, for resistance to strains of *P. syringae* (Salmeron *et al.*, 1996). NBS-LRR proteins have also been shown, in two separate studies, to act downstream of recognition in HR-mediated defence. Peart *et al.*, (2005) showed that NRG1, which encodes a CC-NBS-LRR protein, acts downstream of the TIR-NBS-LRR protein N to facilitate tobacco's resistance to strains of tobacco mosaic virus (TMV) (Peart *et al.*, 2005). They also demonstrated that NBS-LRR proteins are required in signal pathways that lead to HR and theorised this maybe a general feature of such R proteins. This idea was supported when *NRC1*, also encoding a CC-NBS-LRR protein, was cloned from tomato. *NRC1* was implicated in HR signalling by the eLRR R protein Cf4, responsible for tomato resistance to strains of *C. fulvum* (Gabriels *et al.*, 2007). *NRC1* is important for resistance to a range of diseases in tomato and has been found to act downstream of a number of R genes including *Pto*, *Cf-9*, *Rx* and *Mi*, as well as the PAMP receptor *LeEIX* (Gabriels *et al.*, 2007). This adds support to the hypothesis raised by Peart *et al.*, (2005) who suggest that any form of resistance protein may require downstream NBS-LRR proteins for cell death signalling (Peart *et al.*, 2005) and also provides a possible link between PTI and ETI (Gabriels *et al.*, 2007).

Given that many pathogens secrete their effector molecules into the cells of their host, R genes that control disease resistance to these pathogens that encode NBS-LRR proteins are expected to function within the host cell. This prediction has been confirmed by a number of localisation studies. *A. thaliana* R proteins RPS2 and RPM1, both CC-NBS-LRR proteins, localise to the inner leaflet of the plasma membrane (Axtell and Staskawicz, 2003, Boyes *et al.*, 1998) while the CC-NBS-LRR barley resistance protein Mla is cytosolic (Bieri *et al.*, 2004).

More recently, nuclear localisation of Mla has been demonstrated and this localisation been shown to be required for resistance against the powdery mildew fungus *Blumeria graminis* f. sp. *hordei* (Shen et al., 2007). A number of NBS-LRR R proteins including N, RRS1-R, Rx and RPS4 have also been shown to localise to the nucleus (Burch-Smith et al., 2007, Tameling and Baulcombe, 2007, Wirthmueller et al., 2007, Deslandes et al., 2003) spawning a number of recent reviews (Liu and Coaker, 2008, Shen and Schulze-Lefert, 2007). The *L* and *M* genes from flax encode a hydrophobic N terminal region that despite being predicted to act as a signal peptide (Schmidt et al., 2007a), may instead facilitate plasma membrane association or membrane anchoring (unpublished data). This is not unlike the case of RPP1A which has an N-terminal element that resembles a signal peptide but instead localises to cellular membranes, the endoplasmic reticulum (ER) and/or Golgi system (Weaver et al., 2006).

NBS-LRR proteins are undoubtedly critical components of the plant immune system and understanding the roles they play in the coordination of the resistance response is a crucial step in our understanding of plant immunity. This study focuses on the flax R proteins, M and L6, which are members of the NBS-LRR protein class, and are responsible for resistance to strains of flax rust (*M. lini*). Subsequent sections within this chapter will therefore focus primarily on the function of the NBS-LRR R proteins; however, prior to this, the flax-flax rust pathosystem will be introduced.

## **1.5 The Flax, Flax Rust System**

### **1.5.1 The flax rust pathogen**

Rust fungi obtain their nutrients exclusively from living plant cells and are thus defined as obligate biotrophs. The flax rust pathogen, *M. lini*, has an infection strategy that involves a number of steps and complex signalling mechanisms (reviewed by (Heath, 1997, Lawrence et al., 2007)). In short, fungal spores germinate on flax leaves and extend a germ tube out across the leaf until a stomatal entry point is discovered. An appressorium is then formed from which an entry peg extends down between the guard cells. Infection hyphae then grow into the mesophyll layer of the plant leaf, where haustorial mother cells subsequently form. From these cells, haustorial feeding structures penetrate the plant cell wall and invaginate the cell (Heath, 1997). The plant cell membrane is not breached during this process and the region formed between the plant and haustorial membrane is termed the extra-haustorial matrix. Haustoria are specialised feeding structures that are vital for both the acquisition of nutrients from the plant cell, and their conversion into useful metabolites to sustain the fungus (Hahn and Mendgen, 2001, Sohn et al., 2000). This, in effect, is the front line of rust infection and the genetic determinants that both the

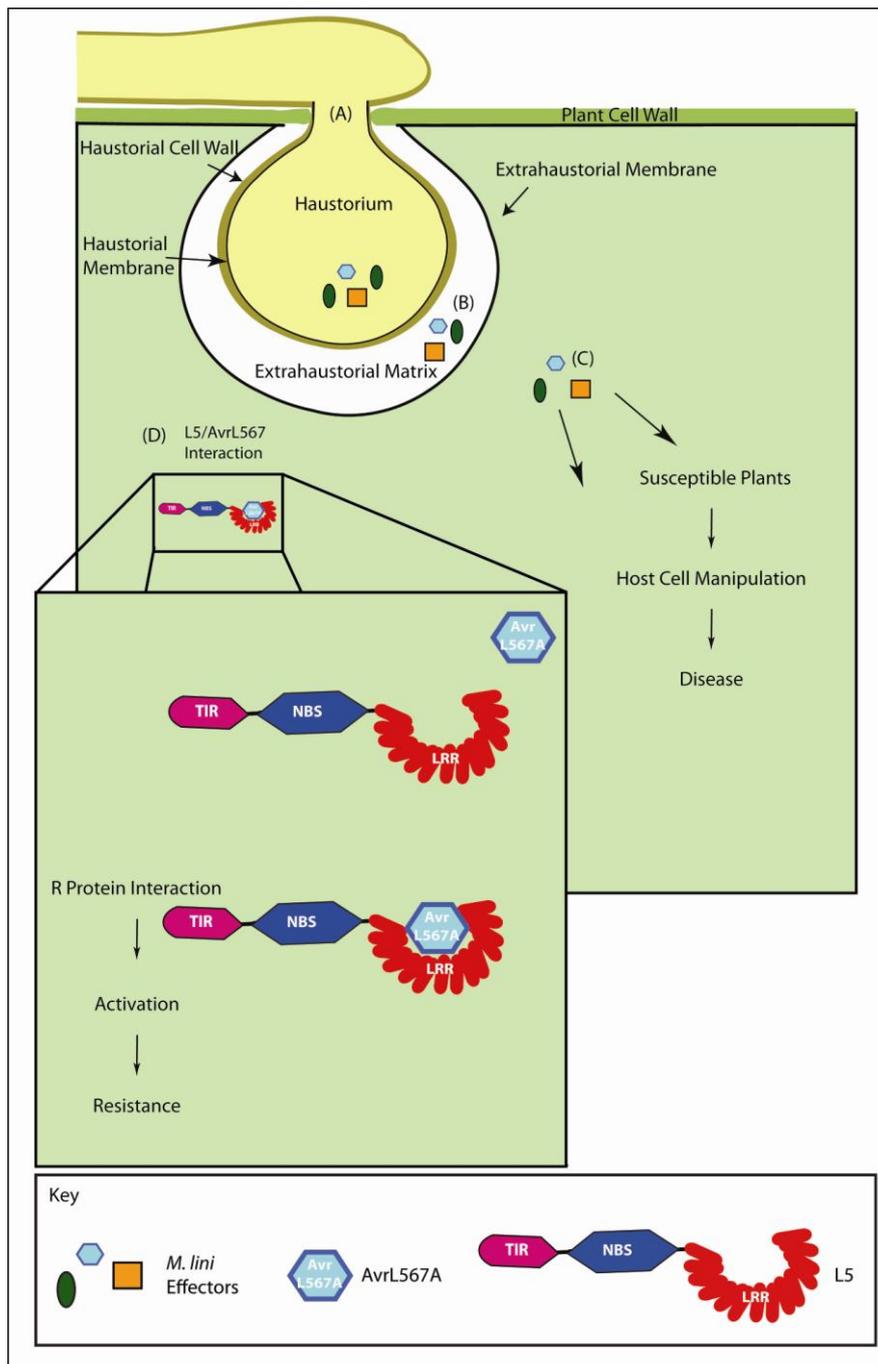
rust and its host possess are critical in determining the cross talk between pathogen and host, and thus the outcome of the interaction (Figure 1.2).

### 1.5.2 Flax R genes and rust effector proteins

Thirty one resistance genes, each conferring resistance to different strains of *M. lini*, have been reported in flax. They are confined to 5 separate loci, designated, K, L, M, N and P. A total of 19 different R genes have now been cloned with all encoding proteins of the Toll-Interleukin 1 Receptor-like, nucleotide binding site, Leucine rich repeat (TIR-NBS-LRR) class (Anderson et al., 1997, Dodds et al., 2001a, Dodds et al., 2001b, Ellis et al., 1999, Lawrence et al., 1995, Lawrence et al., 2009).

Genetic studies have defined approximately 30 *Avr* genes, of which alleles of four have been cloned, sequenced and functionally tested (Dodds and Thrall, 2009). These include *AvrL567*, *AvrM*, *AvrP123* and *AvrP4* (Catanzariti et al., 2006, Dodds et al., 2004). *Avr* genes are expressed in the rust haustoria and encode small soluble proteins with N-terminal signal peptides, although, unlike the R proteins, there is little sequence similarity between different effector proteins, and no identifying sequence motifs that give an insight into their potential function (Ellis et al., 2007a). They do, as mentioned, have an N-terminal signal peptide indicating that they are secreted from the haustoria into the extra-haustorial matrix. It is anticipated that this sequence would be cleaved from the mature protein following secretion. Once in the extrahaustorial matrix, effectors need to be translocated into the plant cell where they can interact directly, as outlined below, with their corresponding intracellular R protein. Although translocation of *AvrL567*, *AvrM* and *AvrP4* across the plant plasma membrane has been demonstrated the mechanism by which this occurs is unknown and is currently being investigated (Catanzariti et al., 2006, Dodds et al., 2004). Furthermore, the host targets of these effectors and their function during infection are at this stage unknown.

An extremely important outcome of the study of flax R proteins and their cognate rust effectors has been the discovery that they interact directly (Dodds et al., 2006) (see 1.7.3). This, coupled with the cloning and expression of flax R and flax rust effector proteins, makes the flax-flax rust pathosystem an excellent model to investigate elements of NBS-LRR protein function, including both interaction and activation.



**Figure 1.2: A model of a flax cell, with or without L5, and its response to flax rust carrying AvrL567A**

A) The haustorial feeding structure penetrates the cell wall and establishes an extra-haustorial matrix. B) Fungal effector proteins are secreted from the haustoria and gain entry to the plant cell by an, as yet, undetermined mechanism. C) In the absence of a corresponding resistance protein, the effectors presumably manipulate host targets to promote disease; however, both the targets and molecular function of effectors are unknown at this point. D) When a corresponding R protein is present, in this case L5, it interacts directly with the effector AvrL567A and activates resistance (see 1.7.3).

NB: The protein domain labels in L5 are abbreviated as in text.

## **1.6 Proposed Function of the NSB-LRR R Protein Domains**

The structural domains predicted within the NBS-LRR R proteins provide the first clue for interpretation of their function. Early hypotheses suggested that the C-terminal LRR is involved in pathogen perception; the central NB domain controls protein activation and the N-terminal CC or TIR enabled transduction of a signal, or signals, to activate the HR and other defence responses. Whilst this assignment of function to separate domains has provided a useful predictive framework for experimental design, it appears to be an oversimplification of how R proteins work, and more intricate interactions between domains are likely to control tertiary protein structure and function.

### **1.6.1 Leucine-rich repeat**

The LRR domain provides a structural platform for protein-protein interactions in a wide range of diverse proteins (Bella et al., 2008, Kobe and Kajava, 2001). The first structure of a LRR protein came from the human and porcine ribonuclease inhibitor (Kobe and Deisenhofer, 1993). It revealed that individual repeats, shown to range from 20-29 residues, make up a structural unit. The structural unit generally consists of  $\beta$  strand and  $\alpha$ -helix linked by either a  $\beta$ -turn or loop (Bella et al., 2008, Kobe and Kajava, 2001). The  $\beta$  strand comprises the characteristic repeat motif xxLxLxx (where x can be any amino acid and L indicates a conserved leucine, but can be replaced with a valine, isoleucine and phenylalanine). The overall structure resembles that of a curved solenoid, with parallel  $\beta$  strands on the concave side, and the helical or  $\beta$  turn elements on the convex side, whereby each repeat is a turn of the solenoid (Bella et al., 2008, Kobe and Kajava, 2001). The concave side is generally thought to be the ligand binding interface, and changes to the x residues exposed to the solvent would be predicted to alter binding site dynamics.

Consistent with this idea is the fact that the most variation between R genes and their closely related homolog's, resides within the LRR domain, particularly in sequences encoding the repeat motif (Ellis et al., 2000). With selection for amino acid variation occurring at these sites, it was predicted that the LRR domain provides the most likely site for effector binding. This, however, is the likely scenario for only a few effector-R protein interactions (see 1.7.3).

In a number of studies of NBS-LRR class of R proteins, the LRR has been implicated in numerous functions other than direct effector interaction. LRRs facilitate the direct interaction with the chaperones Hsp90 and Hsp75 and the co-chaperones, protein phosphatase 5, SGT1 and RAR1, which have all been demonstrated to be required in the positive regulation of R proteins in a pre-activated state (reviewed by (Padmanabhan et al., 2009, Shen and Schulze-Lefert, 2007)). Apart from providing interactions with other proteins, the LRR appears to be

important in maintaining intra-molecular interactions and has been implicated in both the negative and positive regulation of activation in a number of R proteins (see 1.8.2) (Hwang and Williamson, 2003, Hwang et al., 2000, Moffett et al., 2002, Rairdan and Moffett, 2006). With such an array of data presented regarding the function of the LRR domain in R proteins, it is clear that both structural and functional studies of purified R proteins would further clarify its role.

### 1.6.2 Nucleotide binding site (NBS)

NBS domains are common protein elements involved in the activation of proteins through the catalysis of nucleotide hydrolysis and/or the conformational changes in protein structures that are induced by nucleotide binding. R proteins have an NBS domain that contains the hallmark characteristics commonly found in nucleotide binding proteins (Meyers et al., 1999, Traut, 1994). Conserved sequence motifs within NBS domains facilitate catalytic and/or ligand binding activity. R proteins have a number of motifs known to be critical and highly conserved in ATP and GTP binding proteins, such as the P-loop (Kinase 1 or Walker A) and Kinase 2 (Walker B) (Saraste et al., 1990, Walker et al., 1982). Whilst the number of defined motifs varies depending on interpretation of sequences, ten motifs within the NBS region of R proteins have been identified and these are summarised in Table 1.2. Although the biochemical function of these motifs in R proteins remains mostly undefined, many have been demonstrated to have important *in-planta* function, with changes to key residues causing both loss- and gain-of-functions (elicitor-independent) phenotypes (Figure 1.3). Nucleotide binding is clearly critical for R protein function, as mutations to residues in motifs predicted to be involved in this function, particularly within the P-loop, cause loss-of-function.

Interestingly, conserved motifs in R proteins are also conserved in the mammalian apoptotic protease-activating factor 1 (Apaf-1) and *Caenorhabditis elegans* cell death protein-4 (CED-4). Consequently, the R protein NBS domain is most commonly termed the NB-ARC, standing for a nucleotide-binding adaptor shared by APAF-1, certain R gene products, and CED-4 (van der Biezen and Jones, 1998a). The NB-ARC nomenclature will be used in the place of NBS for the remainder of this thesis. R proteins have also been included in a much broader protein class known as signal transduction ATPases with numerous domains (STAND) (Leipe et al., 2004). The STANDs are generally signalling hubs with functions ranging from, mediators of cell death and inflammation, to regulators of transcription. Mandatory to inclusion in this classification is the NOD (nucleotide binding and oligomerisation domain) module, or NB-ARC domain in R proteins, which is closely related to that found in the AAA+ ATPases (Leipe et al., 2004). The STANDs include five major clades differentiated by key features and/or conserved motifs within the NOD module. Apaf-1, CED-4 and R proteins are linked to the AP-ATPase clade,

while the sister clade, NACHT (named after numerous nucleotide binding proteins with similar domains (Koonin and Aravind, 2000)), include the animal NLRs (NOD-LRR), a family of proteins which are involved in human innate immunity and inflammation. Similarities have historically been drawn between animal NLRs and R proteins due to their tridomain architecture, whereby a NOD domain is often linked with a C-terminal repeat domain and a proposed N-terminal signalling domain (Rairdan and Moffett, 2007). A number of conserved motifs are defined within the NOD module of the STAND proteins; however, the P-loop (Walker A) and Kinase II (Walker B) motifs are the most highly conserved. It is predicted that the NOD module will form a similar structural architecture in all STAND proteins, and it is proposed that the mechanisms used to activate these proteins are likely to be conserved (Danot *et al.*, 2009). The classification of R proteins in the STAND group is important as it enables parallels to be drawn from the biochemical and structural studies of other STAND proteins, (discussed further in section 1.8) providing a potential guide, without being too subjective, for future biochemical investigations into R protein function.

The first biochemical study of the NB-ARC domain of an R protein came from a pioneering study on the tomato R genes *I-2* and *Mi1*. Recombinant CC-NB-ARC (truncated for the LRR) proteins expressed in, and purified from, *E. coli*, were demonstrated to have ATP binding and hydrolysis capabilities (Tameling *et al.*, 2002). Importantly, when the invariant lysine (Figure 1.3) within the P-loop was mutated to an arginine, both binding and hydrolysis capabilities of the recombinant protein were significantly reduced. The same mutation in the R gene caused a loss of activity *in-planta*, indicating a role for ATP binding and/or hydrolysis in R protein function. ATP binding and hydrolysis activity associated with the P-loop motif has also been shown from *in-vitro* studies of the NB-ARC-LRR region of tobacco N resistance protein (Ueda *et al.*, 2006). The work by Tameling and co-workers and the mutational analysis of others (Figure 1.3) has confirmed that activation of R proteins, at least in part, is achieved by the function of the NB-ARC domain (Tameling *et al.*, 2002).

| Domains   | NB (~180aa) |                      |   |           |           |                          | ARC1 (~75aa) | ARC2 (~105) |                               |            |
|-----------|-------------|----------------------|---|-----------|-----------|--------------------------|--------------|-------------|-------------------------------|------------|
| Motifs    | hhGRExE     | P-loop               | RNBS-A  | Kinase II | RNBS-B    | RNBS-C                   | GLPL         | Motif VIII  | RNBS-D                        | MHD        |
| Consensus | hhGRExE     | <b>GVG<b>K</b>TT</b> | FLENIRExSKKHGLEHL<br>QKKLLSKLL<br>(FDLxAWVCVSQxF) | LLVLDDVW  | GSRIITTRD | YEVxxLSEDEA<br>WELFCKXAF | <u>GL</u> PL | <u>SYD</u>  | FLHIACFF<br>(CFLYCAL<br>FPED) | <u>MHD</u> |

**Table 1.2: Nucleotide binding site motif consensus sequences**

The NB-ARC spans ~360 amino acid residues and can be separated into three domains based on the crystal structure of Apaf-1 (Riedl et al., 2005). These are designated NB, ARC1 and ARC2. Consensus sequences for the plant NB-ARC-LRR resistance proteins have been interpreted from a number of studies (Leipe et al., 2004, Meyers et al., 1999, Meyers et al., 2003, Pan et al., 2000, Takken et al., 2006, van der Biezen and Jones, 1998a) and are summarised in (Takken et al., 2006). A Kinase-3 domain is absent from the plant R protein nucleotide binding domain and is therefore not shown above, however, the RNBS-B motif shares positional but not sequence similarity with known Kinase-3 domains.

The non-TIR and TIR classes of NBS-LRR genes vary in the RNBS-A and -D motif regions, hence the non-TIR consensus is in brackets.

Bold type in the consensus sequence row is the invariant amino acids of the P-loop- lysine; in the consensus the underlined residues are highly (almost invariantly) conserved and positions of functional interest.

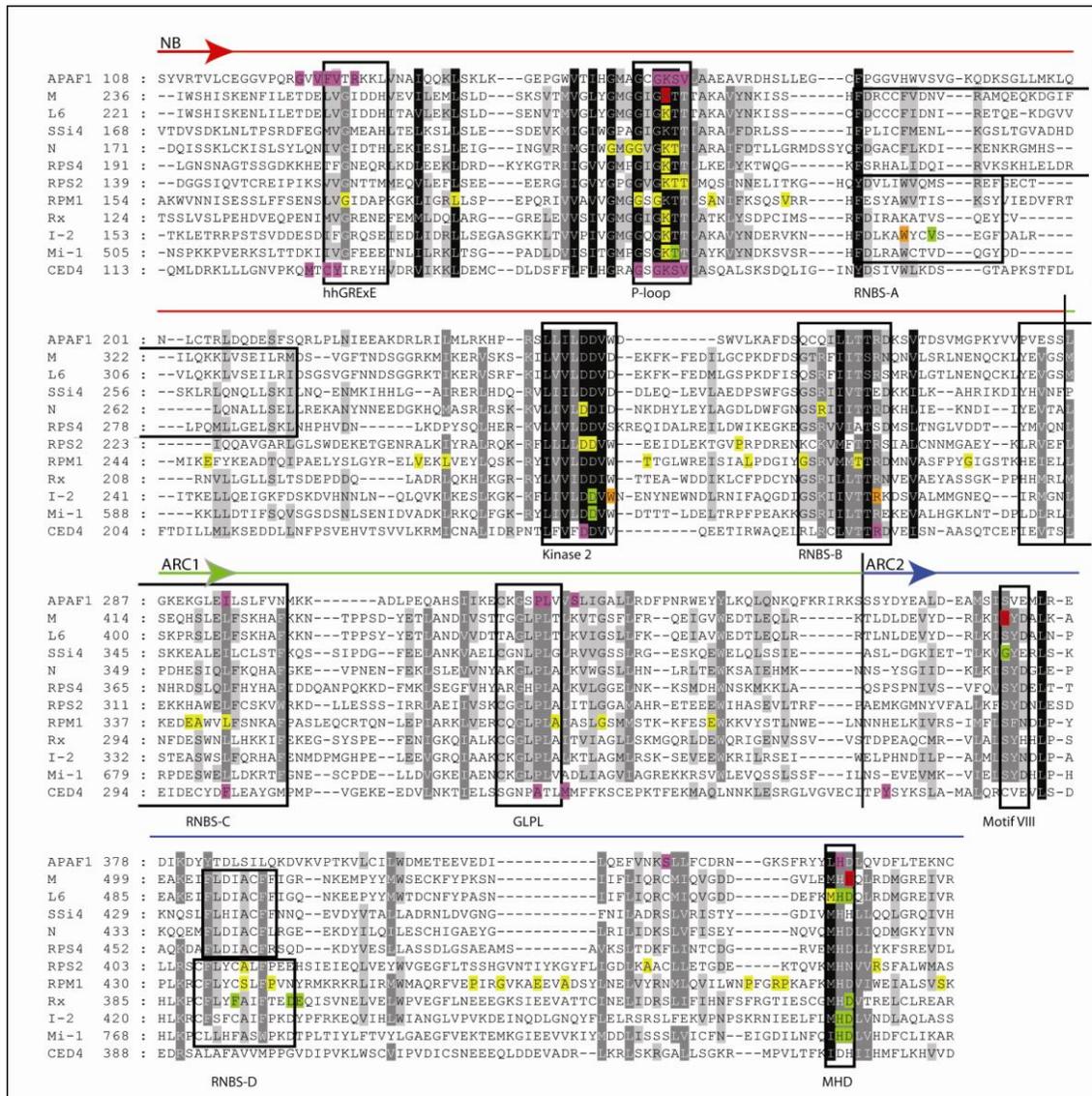


Figure 1.3: An NB-ARC alignment

This figure is adapted from (van Ooijen et al., 2008b). It shows a multi-sequence alignment of the NB-ARC region of numerous R proteins, Apaf-1 and CED-4. TIR-NB-ARC-LRR R proteins included in the alignment are M/L6 (flax), SSi4/RPS4 (*A. thaliana*) and N (tobacco). The non-TIR R proteins include RPS2/RPM1 (*A. thaliana*), Rx (potato) and I-2/Mi-1 (tomato). The predicted R protein subdomain borders, as determined by (Albrecht and Takken, 2006), are defined with a line that runs the length of the subdomain, NB: red; ARC1: green; ARC2: blue. Residues that are involved in coordinating the binding of either ADP or ATP in respective 3D structures of Apaf-1 (1z6t) or CED-4 (2a5y) were determined from the referenced source (Riedl et al., 2005, Yan et al., 2005) and using the ligand interaction program interface on the protein data bank website (<http://www.rcsb.org/pdb>), are highlighted in purple. Ten motifs are defined (see table 1.2) and are labelled/ boxed. Residues that have been demonstrated to be critical in respective R protein function for these proteins are highlighted; yellow indicates loss-of-function and green indicates gain-of-function. Residues highlighted in orange are predicted to cause loss of function. The sources of this information are listed as follows L6 (Howles et al., 2005), SSi4 (Shirano et al., 2002), N (Dinesh Kumar et al., 2000), RPS4 (Zhang et al., 2004), RPS2 (Axtell et al., 2001, Mindrinos et al., 1994, Tao et al., 2000), RPM1 (Grant et al., 1995, Tornero et al., 2002), Rx (Bendahmane et al., 2002), I-2, (de la Fuente van Bentem et al., 2005, Tameling et al., 2006, van Ooijen et al., 2008b), Mi-1 (van Ooijen et al., 2008a, van Ooijen et al., 2008b).

### 1.6.3 Amino terminal domain

As discussed above, the majority of R proteins contain either a CC or TIR domain at their N-terminus. The CC domain is a common structural domain involved in an array of different biological processes including protein-protein interactions (Martin et al., 2003). A number of studies have shown that the CC domain coordinates the interaction between an R protein and host proteins that are targeted by effectors (see section 1.7.2) (Ade et al., 2007, Mackey et al., 2002). The CC is also required for downstream signalling by the *A. thaliana* CC-NBS-LRR protein, RPS5 (Ade et al., 2007, Shen et al., 2007) and is subsequently implicated in both detection and signalling.

The TIR domain is so named due to its similarity with the cytoplasmic signalling domain utilised by Toll and Interleukin-1 (IL-1) like receptor proteins. In these mammalian receptors, the cytoplasmic domain provides a protein scaffold for protein-protein interactions and is indispensable for signalling (Martin and Wesche, 2002, Xu et al., 2000). These sequence similarities suggest that the plant TIR domain may provide the same function, and this has been supported by a number of functional studies. Deletions and point mutations within the TIR, shown to affect *Toll* and IL-1R signalling, also affect the N-mediated signalling events that lead to TMV resistance in tobacco (Dinesh-Kumar et al., 2000). Also, over-expression of the TIR domain, including a short C-terminal extension, of the flax R protein L10, and also *A. thaliana*'s RPS4 and RPP1A, caused an effector-independent necrotic response in *A. tumefaciens* transient assays in tobacco, and *Nicotiana benthamiana* leaves, respectively (Frost et al., 2004, Swiderski et al., 2009).

Homotypic interactions between TIRs and/or TIR dimerisation seem essential for signalling activation of Toll-like receptors (Takeda and Akira, 2005). TIR-mediated oligomerisation also appears to be involved in at least one member of the TIR-NBS-LRR class of R proteins. Transient expression experiments of the tobacco N disease resistance protein, followed by co-immunoprecipitation, demonstrated that the N protein oligomerises in the presence of the TMV elicitor p50 (Mestre and Baulcombe, 2006). This oligomerisation event was abolished in N protein containing a mutation of the conserved lysine of the P-loop motif. It is not known, however, if this effector triggered R protein oligomerisation is a general feature of the resistance response mediated by the TIR-NBS-LRR class. In the case of the tobacco N protein, the TIR may also play a regulatory role in activation through intramolecular interactions (Ueda et al., 2006) and also in pathogen perception, through interactions with host proteins (Burch-Smith et al., 2007, Caplan et al., 2008).

Collectively, these results indicate that the TIR is likely to play important roles in inter- and intra-molecular interactions and regulation of R protein activity. It also contributes to a growing list of similarities that exist between plant and animal innate immunity. It does, however, further expose a recognised need for the biochemical study of TIR and CC domains, in the context of intact full-length R protein, to elucidate the precise role these domains play in R-mediated resistance.

## **1.7 Effector Perception by R Proteins**

As discussed above, the presence of a particular R gene is vital for resistance to specific pathogens, but what role do the products of R genes play in pathogen detection? R proteins must recognise the presence of a foreign pathogen-derived effector protein. Such recognition must presumably lead to R protein activation, the result of which is the transduction of a signal, or signals, to promote defence responses. The remaining sections will tackle R protein function, covering, with examples, recognition, activation, and signalling.

### **1.7.1 Recognition of effector proteins by R proteins**

Arguably the most important mechanism to decipher in plant disease resistance is the way in which an R protein recognises a pathogen derived effector. Conceptually, the simplest mechanism of interaction between R and effector proteins is a direct one, whereby the R protein acts as a receptor, and the effector protein as a ligand. Whilst this is true in some cases, many R-effector interactions characterised to date recognise the effector's presence by monitoring the integrity of host proteins. In this section, evidence that demonstrates both types of recognition will be introduced.

### **1.7.2 Indirect recognition: the guard and decoy hypothesis**

The idea that an R protein may be guarding a host target of its corresponding effector protein, was first used to explain the dynamics between tomato Pto and Prf, which provide immunity to *P. syringae* carrying AvrPto (van der Biezen and Jones, 1998b). Whilst Pto has been demonstrated to interact directly with AvrPto (Scofield et al., 1996, Tang et al., 1996), Prf is also required for resistance. An array of evidence places Prf very early in the Pto-mediated resistance pathway (Martin et al., 2003, Rathjen et al., 1999). Pto forms a molecular complex with a unique N-terminal region of Prf, and Prf has been shown to have signalling, regulation and recognition capabilities (Mucyn et al., 2006). It is suggested that Prf monitors the integrity of Pto, a target of the virulence activity of AvrPto (van der Biezen and Jones, 1998b). A review by Dangl and Jones (2001) generalised this theory into the guard hypothesis. It predicts that the Avr-Pto product, functioning as an effector, targets a host protein to promote disease. The R protein functions by

protecting the integrity of the host target, thus providing the plant with an effective means of detecting the pathogen and activating an immune strategy (Dangl and Jones, 2001). Support for the guard hypothesis has come from a number of studies in which *P. syringae* is the infectious agent, and effectors target an array of “guarded” host proteins or guardees, (reviewed most recently in (Jones and Dangl, 2006)). This list now includes a *P. syringae* protease AvrPphB which cleaves the *A. thaliana* host protein PBS1 to activate the CC-NBS-LRR protein RPS5 (Ade et al., 2007). PBS1 and RPS5 are required for resistance to AvrPphB, and their interaction is mediated by the CC domain of RPS5 (Ade et al., 2007). Another example is the *A. thaliana* R protein, RPM1, which interacts through its N-terminal domain with the guardee RIN4, (Mackey et al., 2002). RIN4 is the target for the pathogen effector protein AvrRpm1 from *P. syringae*

Recently, the guard model has come under some scrutiny. It has been proposed that the guardee may in actual fact be a decoy protein that mimics an effector target (van der Hoorn and Kamoun, 2008). The “decoy” model has been generated in an attempt to cover some potential inconsistencies associated with the guard model. The authors argue that some effectors have multiple host cell targets, and that pathogen virulence activity may not require alteration of a guardee. A guardee would therefore have opposing selection pressures depending on the presence or absence of an R gene. When an R gene is present, the guardee has selection pressures towards interacting with an effector, however in the absence of an R gene, natural selection would drive the guardee to decrease its binding affinity with an effector. The decoy model has been put forward to potentially solve this evolutionary discrepancy as a decoy protein would evolve to always maintain its interaction with an effector and not be compromised by this interaction (van der Hoorn and Kamoun, 2008). The main difference between a decoy and a guardee is that the decoy is not required in host resistance, and that alteration of the decoy does not result in an enhanced fitness to the pathogen when the R protein is not present (van der Hoorn and Kamoun, 2008). The guard and decoy models may not be mutually exclusive, and both models provide an excellent framework to further decipher the indirect interaction between R and Avr proteins in the future.

### 1.7.3 Direct recognition

To date, four cases of direct interaction between effectors and NB-ARC-LRR R proteins have been reported, all using the yeast two hybrid (Y2H) system. The R genes involved in these interaction tests include *RRS1-R* from *A. thaliana*, *L5*, *L6* and *L7* from flax, *Pi-Ta* from rice and *N* from tobacco and their corresponding pathogen and the pathogen effectors, *Ralstonia solanacearum*, effector Pop2, *M lini*, effector AvrL567, *Magnaporthe grisea*, effector Avr-Pita and tobacco mosaic virus, effector p50, respectively (Deslandes et al., 2003, Dodds et al., 2006, Jia et

al., 2000, Ueda et al., 2006). In all cases, the LRR domain is required for interaction with the effectors. Two of these interactions have been supported by *in vitro* protein binding assays, however in both cases the interaction could only be demonstrated using truncations of full-length proteins containing only the LRR or NBS-LRR domains (Jia et al., 2000, Ueda et al., 2006).

The LRR has been clearly demonstrated to provide pathogen strain-dependent specificity within the gene-for-gene system. The general role of LRRs in protein-protein interactions (see 1.6.1), implicates involvement in effector binding in R proteins that interact directly with their effectors (Dodds et al., 2001b, Ellis et al., 1999, Holt et al., 2003). Within the flax-flax rust system, both genetic and, more recently, *in vitro* interaction studies have strongly implicated the LRR domain with a function of direct recognition of rust effector proteins. Genetic analysis of the *L* locus has shown that there are 12 alleles (encoding entirely TIR-NBS-LRR proteins), conferring at least 10 rust resistance specificities to different strains of *M. lini* (Ellis et al., 2007a, Ellis et al., 1999). Variants, *L6* and *L11*, differ only in sequences within the LRR, yet they recognise distinct strains of *M. lini* (Ellis et al., 1999). Further analysis demonstrated that the majority of those polymorphisms occurred within the repeat motif at the predicted solvent exposed x positions within the LRR consensus, with different polymorphisms being essential for either *L6* or *L11* specificity (Ellis et al., 2007b). Domain swap experiments between homologues at the *P* locus in flax, also demonstrate the importance of the LRR in specificity. Six amino acid differences between *P2* and *P* were found to be sufficient to distinguish between their resistance specificities. All of the polymorphic residues were localised to variable x residues of the repeat motif in the LRR domain (Dodds et al., 2001b).

It is clear from these results that the LRR domains of flax R proteins play a pivotal role in effector recognition, but what information on this subject can be obtained from the analysis of rust effectors? The AvrL567 gene family encodes twelve highly diverse sequence variants, seven of which return a necrotic response, when infiltrated using the *A. tumefaciens* delivery system into flax leaves containing *L5*, *L6* or *L7*. This degree of diversity, and difference in specificity, is the likely result of an evolutionary battle, where the rust and its host try to avoid and maintain recognition, respectively (Dodds et al., 2006). Y2H experiments with the different AvrL567 variants and *L5*, *L6* and *L7* support the gene-for-gene specificity shown in the *in planta* study, and demonstrate a direct interaction between the gene products (Dodds et al., 2006). Interestingly, a resistance inactive mutation in *L6* involving an invariant lysine within the P-loop (also called walker A motif), discussed below as a critical motif found in all ATP binding proteins, prevents interaction with the AvrL567 effector in the Y2H analysis. This suggests that the *L6* protein requires a functional nucleotide binding pocket to interact with the AvrL567 effector protein. The integrity of the P-loop has also been shown to be critical in maintaining intra-molecular

interactions in the potato virus X resistance protein, Rx (Moffett et al., 2002). It is therefore conceivable that a disruption in protein structure caused by a mutation in the P-loop motif of L6 may be responsible for the loss of interaction. Similar results have been demonstrated in Y2H experiments involving M and AvrM (P. Dodds, personal communication).

Structural determination of AvrL567A and D proteins revealed that almost all the side chains of the polymorphic residues in the AvrL567 variants were mapped to the outer solvent exposed region of the protein (Wang et al., 2007). Importantly, AvrL567A is recognised by both L5 and L6; however, AvrL567D was only detected by L6, and AvrL567C by neither (Dodds et al., 2006). From this, Wang et al., (2007) made targeted mutations to four polymorphic residues in AvrL567 (residues 50, 56, 90, 96) believed to be critical in this specificity. Using both Y2H, and an *in planta* HR assay, they demonstrated that differing combinations of the four mutations could alter the specificity of the effector. The data also suggested that multiple contacts at distant points in the AvrL567 protein are required for interaction, however single amino acid changes were sufficient to both stabilise and destabilise the interaction (Wang et al., 2007). Armed with this information, a model of the interaction between the L5 LRR domain, and the AvrL567A effector protein was generated, satisfying all the critical interactions that determined specificity (Wang et al., 2007). The structure of the L5 LRR is modelled on other known LRR structures. It is important therefore to stress that the model of L5/AvrL567 interaction is highly speculative.

Resistance specificity in flax is, however, not always confined to the LRR (Luck et al., 2000). Analysis of alleles at the *L* locus in flax demonstrated that alleles with the same LRR region but different TIR and NBS regions can encode different specificities (Ellis et al., 1999, Luck et al., 2000). Clearly interaction between flax rust effectors and flax R proteins is not exclusively confined to the LRR

Ueda et al., (2006) reported that the tobacco N protein directly recognises TMV p50 using Y2H and *in vitro* studies (Ueda et al., 2006). Two recent studies, however, cast some level of doubt over this conclusion. Firstly, co-localisation studies have shown that the TIR domain of N is critical for the interaction and it alone can associate with the p50 effector (Burch-Smith et al., 2007). However, Y2H analysis and *in vitro* pulldown could not demonstrate that the interaction was in-fact direct. A follow-up study has since demonstrated that the interaction between the p50 and N is mediated by a tobacco host protein called NRIP1 (Caplan et al., 2008).

It is clear therefore that care should be taken when interpreting the L5/AvrL567A interaction model. It is likely that many subtleties exist in the R/effector protein interaction in the flax-flax rust system, and for other R/effector systems for that matter. For any predicted interaction between R and effector proteins to be tested, protein purification, followed by structural and biochemical analysis, including *in vitro* R/effector binding studies, is required.

## 1.8 *R* protein Structure/Function

### 1.8.1 Structure studies of the NB-ARC region of Apaf-1 and CED-4

While the structure of the NB-ARC region of an *R* protein has not been solved, the crystal structure of the NB-ARC domains of Apaf-1 and CED-4 have (Riedl et al., 2005, Yan et al., 2005). The NB-ARC region of Apaf-1, in conjunction with the N-terminal caspase recruitment domain (CARD), was solved in its inactive conformation. The NB-ARC domain was most similar to structures solved for the AAA+ ATPases protein family (Riedl et al., 2005). The NB-ARC region could be separated into four structurally distinct domains. The NB domain produces a three layered  $\alpha/\beta$  domain, also described as a  $\beta$ -sheet flanked by  $\alpha$ -helices (Albrecht and Takken, 2006) and the ARC domain was further separated into 3 ARC subdomains; ARC1 generates a four-helix bundle, while ARC2 produces a winged helix fold and the ARC3 domain a second helical domain. ADP was found deeply buried within the structure, bound between the  $\alpha/\beta$  domain, helical domain one and the winged helix domain, with the critical binding residues identified in Figure 1.3 (Riedl et al., 2005).

The structure of CED-4 was solved in a complex with CED-9, a constitutive inhibitor of CED-4. CED-9 binds to an asymmetric dimer of CED-4, however, it recognises only one of the two CED4 proteins (Yan et al., 2005). The CED-4 proteins are ATP bound, which, like ADP in Apaf-1, is deeply buried between the NB, ARC1 and ARC2 domains. It is of interest, however, that very different domain conformations are adopted by Apaf-1 and CED-4 (Takken *et al.*, 2006, Yan *et al.*, 2005). It is suggested that CED-4 is frozen in an active state with ATP bound, although, activation is masked by the binding of the CED-9 inhibitor (Danot et al., 2009).

Alignment studies of *R* proteins and the NB-ARC region of Apaf-1, predicts that *R* proteins contain three subdomains within the NB-ARC domain; NB, ARC1 and ARC2 (Figure 1.3) (Albrecht and Takken, 2006). For the tomato *R* protein, I-2, mapping studies of the NB-ARC region with Apaf-1 and CED-4, and more recent modelling studies with Apaf-1 as the template, have been used to highlight a number of residues likely to be important in *R* protein function (Takken et al., 2006, van Ooijen et al., 2008b). A large number of loss- and gain-of-function (autoactive) mutations map to areas within the NB-ARC region shown to be important in ATP/ADP interaction in the respective Apaf-1 and CED-4 structures (Figure 1.3). This further supports the role of ATP and ADP binding in *R* protein function (Takken *et al.*, 2006). Of particular interest is the cluster of autoactive mutations in conserved motifs within the ARC2 region (Figure 1.3). A recent structure/function study of one such motif, the MHD (so named after its consensus), implicates it as an important regulator of *R* proteins (van Ooijen et al., 2008b). The motif is located towards the C-terminal end of the NB-ARC domain and is the most well

characterised motif in R proteins regarding autoactivity. Manipulations to the conserved histidine or aspartate within the MHD motif have been reported to result in an autoactive phenotype in a number of R proteins (Bendahmane et al., 2002, Howles et al., 2005, van Ooijen et al., 2008b). The MHD is predicted to play a sensory role, coordinating nucleotide binding and controlling the interaction between the subdomains within the NB-ARC domain (van Ooijen et al., 2008b). In the Apaf-1 structure the highly conserved histidine, within this equivalent motif, is shown to form a hydrogen bond with the  $\beta$ -phosphate of the bound ADP (Riedl et al., 2005).

### 1.8.2 Intramolecular interactions regulate activation and signalling

The structures of Apaf-1 and CED-4 highlight that intramolecular interactions between protein domains are almost certainly involved in maintaining both inactive and active conformations. The potato *Rx* gene (CC-NB-ARC-LRR), which confers resistance to the potato virus X (PVX), has been at the forefront in understanding such interactions in R proteins. Transient expression assays involving the co-expression of separated domains of *Rx*, (CC-NB-ARC and LRR) and (NB-ARC-LRR and CC) both demonstrated an Avr dependent HR, signifying the reconstitution of resistance, presumably by the reconstitution of a functional R protein (Moffett et al., 2002). Using co-immunoprecipitation reactions, the domains separated by construct design could interact physically, however the CC-NB-ARC and LRR were shown to interact only in the absence of the Avr protein. This indicates that the effector protein is capable of disrupting the CC-NB-ARC and LRR interaction, which may suggest that the LRR is regulating activity (Moffett et al., 2002). This idea was supported in separate studies of *A. thaliana* proteins where in the cases of RPS2, RPS5 and RPP1A, the LRR is likely to be involved in negative regulation of R protein activation. This was the conclusion from the expression experiments of R genes without their LRR that resulted in an autoactive phenotype (Ade et al., 2007, Tao et al., 2000, Weaver et al., 2006). Ironically, the LRR has also been argued to positively regulate resistance for the tomato R protein Mi1.2 (Hwang and Williamson, 2003, Hwang et al., 2000). These differing arguments on domain function further emphasise the need for structural and *in vitro* functional analysis of purified R proteins.

The Moffett group have shown that physical association between the CC-NB-ARC and LRR requires the ARC1 subdomain, but is not dependent on nucleotide binding (Rairdan and Moffett, 2006). The ARC2 subdomain is believed to play an auto-inhibitory role in *Rx* and is also required for activation (Rairdan and Moffett, 2006). The importance of the ARC2 domain in STAND protein activation has also been reported. In a recent structural review of the crystal structures available for STAND proteins, the conformation changes that underlie protein activation was a major focus (Danot et al., 2009). From this study, It is suggested that the ARC2

domain (WHD) undergoes a 180°C rotation between the closed and open conformations (Danot et al., 2009). This evidence suggests that ARC2 relays the elicitor perception signal to the rest of the NB-ARC domain for protein re-organisation and activation. This idea is somewhat consistent with the localisation of a large number of loss- and gain-of-function mutations within the ARC2 domain in R proteins, discussed above. However, such a hypothesis requires further testing *in vitro* for any STAND-like protein, including plant R proteins.

### 1.8.3 A model of R protein activation

As discussed above, the NB-ARC domain of R proteins is critical for function and is likely to provide the necessary components required for protein activation. As stated, truncated recombinant versions of the R proteins I-2, Mi1 and N have demonstrated ATP binding and hydrolysis capabilities (Tameling et al., 2002, Ueda et al., 2006). Further analysis of I-2 demonstrated the potential mechanism of activity through the biochemical investigation of two mutations shown to display autoactive phenotypes *in planta* (Tameling et al., 2006). To elaborate, utilising their *E. coli* recombinant expression system, CC-NB-ARC domains of I-2 carrying an aspartate to glutamate change within the kinase 2 motif, and a serine to phenylalanine change within the RNBS-A motif, were expressed and purified (Figure 1.3). The two mutant proteins were found to have similar ATP binding kinetics, however, their hydrolysis activity was compromised when compared to the non-mutated CC-NB-ARC protein (Tameling et al., 2006). This suggested that hydrolysis may not be a requirement for R protein activation and changes in the identity of the bound nucleotide may in-fact control the activation of the protein. The same kinase 2 mutation in the *A. thaliana* protein RPS5 also caused an autoactive phenotype (Ade et al., 2007).

Tameling and co-workers proposed that the NB-ARC domain of R proteins function as a molecular switch to control activation (Takken et al., 2006, Tameling et al., 2006). The functional switch is anticipated to be in an “off” or inactive state with ADP bound, and an “on” or active state when ATP is bound. In such a model, effector perception would theoretically induce nucleotide exchange from ADP to ATP, thus forming a molecular switch. Hydrolysis of the ATP molecule to ADP may enable the R protein to be reset, facilitating possible signal amplification (Figure 1.4). This model, however, differs from that proposed from the only other biochemical study of an NBS-LRR R protein (Ueda et al., 2006). Here the resting state of N is proposed to be ATP bound, and interaction with the TMV elicitor, p50, promotes ATP hydrolysis. Ueda et al., (2006) suggest that hydrolysis is what triggers the defence response (Ueda et al., 2006). So which model is correct? Given that these researchers presented work on different R proteins from different structural classes (CC/TIR), it is possible both models are correct and there exist a fundamental difference

in the activation of R proteins from these classes. However, a number of biochemical studies of STAND proteins have supported the molecular switch model involving ADP/ATP exchange (Figure 1.4), with suggestions that this mode of activation may be generally conserved among this broad class of proteins (reviewed by (Danot et al., 2009)).

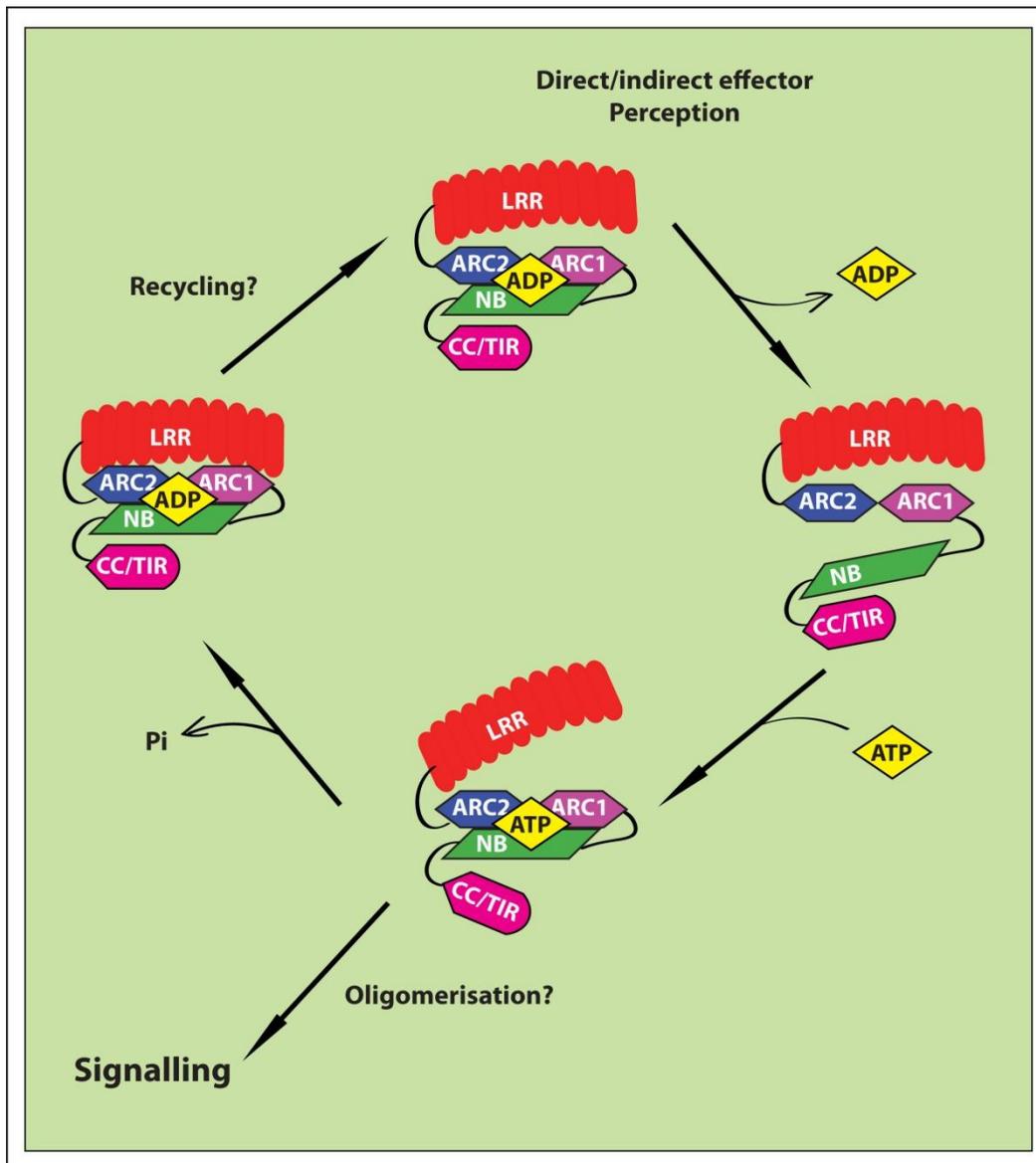
#### **1.8.4 Functional studies of STAND proteins support the R protein molecular switch model**

As previously mentioned the NOD module (NB-ARC domain in R proteins) is fundamental for the inclusion of a protein into the STAND class and assists its classification within the STAND sub groups. In a recent review it has been predicted that the NOD module forms an architecture that is generally conserved throughout the STAND members, and also suggests that the subsequent mechanics behind the function of the STAND proteins could also be generally conserved (Danot et al., 2009). At this stage, the structures of only four members of the STAND class are available and clearly more structures are required before major generalisations can be made. However, we can tentatively interpret the mechanics of protein activation from biochemical studies of STAND proteins and apply this knowledge to the context of R protein activation. To this end, arguably the most advanced biochemical studies of STAND proteins has come from the investigations of Apaf-1 and MalT, the latter an *E. coli* transcription activator. Biochemical characterisation of both of these proteins is further advanced than that of R proteins and therefore the experimental evidence to support models of activation that lead to signalling are more compelling.

MalT is an *E. coli* transcription activator of the maltose regulon. Like R proteins, the model of MalT activation has been shown to involve an ADP bound autoinhibited state, and an ATP bound active state (Marquenet and Richet, 2007). The evidence for this comes from the purification of the wild type MalT protein in an ADP bound state, while a mutation in the second aspartate of the kinase-2 motif, which is the same respective residue mutated and analysed in I-2 (Tameling *et al.*, 2006), was ATP bound (Marquenet and Richet, 2007). A functional assays demonstrated that this ATP bound form of MalT was hyper-activated (Marquenet and Richet, 2007). MalT is normally activated by the chemical, maltotriose, which was demonstrated experimentally to promote ADP to ATP exchange and the formation of an oligomer. The addition of maltotriose also increased the rate of ATP hydrolysis in wild type MalT, however, the hyper-activated mutant was unable to hydrolyse ATP. Whilst ATP hydrolysis is therefore not critical for transcription activation by MalT, it is critical for the control of activity. It is suggested that ATP hydrolysis returns a STAND protein from an active to an inactive state, and that it may therefore be involved in protein recycling (Marquenet and Richet, 2007).

Apaf-1 is an extremely important mammalian cell death protein which complexes with cytochrome c in the presence of deoxyadenosine triphosphate (dATP) or ATP to form an oligomeric apoptosome. The apoptosome is capable of recruiting and activating procaspase-9, which, in turn activates a caspase-related cell death pathway (Bao and Shi, 2007, Riedl et al., 2005). Some conjecture surrounds the nucleotide bound state of the autoinhibited form of Apaf-1, with structural and biochemical studies suggesting it is dADP/ADP bound (Bao et al., 2007, Riedl et al., 2005), while others present evidence suggesting it is bound to dATP (Kim et al., 2005). Nonetheless, the activated apoptosome constitutes a complex that contains seven Apaf-1 proteins in the dATP bound state bound with cytochrome c (Bao and Shi, 2007, Kim et al., 2005, Zou et al., 1999).

From the experimental evidence presented for both MalT and Apaf-1 activation, it is tempting to speculate upon the similarities in the model proposed for the activation of an R protein (Tameling et al., 2006). At this point some key features of the model in R proteins have been demonstrated experimentally, however, evidence to further support or reject such a model are undoubtedly necessary. This will come from further attempts to express and purify R proteins to enable their biochemical study.



**Figure 1.4: The molecular switch model for R protein activation**

This figure is adapted from (Takken et al., 2006, Takken and Tameling, 2009, Tameling et al., 2006). In the absence of a pathogen effector, the R protein exists in an inactive, tightly regulated, ADP bound conformation (1). Direct or indirect interaction with an effector protein (2) causes the protein to change conformation, stimulating a more open structure (3). The open conformation is then able to exchange ADP for ATP, resulting in the formation of an active ATP bound conformation that may undergo oligomerisation (4). It is the ATP bound conformation that is active and signals the defence response. Hydrolysis of the bound ATP enables the protein to return to an autoinhibited ADP bound state (1). This could potentially enable protein recycling as a possible method for signal amplification.

## 1.9 ***Project Aims and Objectives***

Our understanding of the functions of R protein domains and how R proteins are activated has progressed significantly over recent years, there is, however, still much to be determined. The ability to perform biochemical analysis has been confounded by a well recognised difficulty in expressing and purifying functional full-length or near full-length recombinant R protein. This study aimed to help rectify this deficiency and describes a method for the production of near full-length flax R proteins, M and L6. This method has facilitated a biochemical investigation of the flax R protein, M; to further elucidate aspects of its function and interaction with the corresponding flax rust effector protein, AvrM.

### 1.9.1 **Aims**

1. Develop and refine recombinant protein expression in, and purification from, *Pichia pastoris*, to enable the production of near full-length flax M and L6 proteins, to facilitate biochemical investigations.
2. Using purified recombinant M protein, investigate the mechanisms of activation and regulation within the NB-ARC domain.
3. Determine the nature of the interaction between the flax M protein and the flax rust effector protein AvrM *in vitro*.
4. Reconstitute *in-vitro*, the activation of the M protein in response to AvrM, and investigate any affect this may have on the nucleotide binding dynamics of the M protein.



## **Chapter 2: Experimental Procedures**

This chapter outlines the methodology used for all experiments pertaining to this thesis

## 2.1 Materials

All chemicals were ultragrade and purchased from either Amresco or Sigma-Aldrich. Buffers and broths used for culturing micro-organisms were obtained from Oxoid. DNA modifying enzymes including restriction enzymes, ligases and phosphatases were purchased from New England Biolabs, while the high fidelity DNA polymerase (Fusion taq) was obtained from Finnzymes. DNA purification kits were obtained from Promega, while oligonucleotides were purchased from GeneWorks (Pty Ltd, Adelaide, Australia). All chromatography media, columns and chromatography equipment were sourced from GE Life Sciences.

### 2.1.1 Culture media, solutions and buffers

All culture media, solutions and buffers were prepared in milliQ purified water. A summary of the culture media (Table 2.1) and the solutions/buffers used for protein or DNA manipulation (Table 2.2) are detailed below.

**Table 2.1: Culture media**

| Medium                                       | Composition   |
|--|---|
| Luria-Bertani (LB) <sup>a</sup>              | Tryptone (10 gl <sup>-1</sup> ), yeast extract (5 gl <sup>-1</sup> ), NaCl (10 gl <sup>-1</sup> )   |
| ½ Salt LB <sup>a</sup>                       | Tryptone (10 gl <sup>-1</sup> ), yeast extract (5 gl <sup>-1</sup> ), NaCl (5 gl <sup>-1</sup> )  |
| Yeast Peptone Dextrose (YPD) <sup>a</sup>    | Yeast extract (10 gl <sup>-1</sup> ), bacto-peptone (20 gl <sup>-1</sup> ), dextrose (20gl <sup>-1</sup> )  |
| BMGY (buffered glycerol-complex medium)      | Yeast extract (10 gl <sup>-1</sup> ), bacto-peptone (20 gl <sup>-1</sup> ), 100 mM potassium phosphate pH 6.0, yeast nitrogen base (13.4 gl <sup>-1</sup> ), biotin (40 µgl <sup>-1</sup> ) <sup>b</sup> and 1% glycerol (v/v)  |
| BMMY (buffered methanol-complex medium)      | Yeast extract (10 gl <sup>-1</sup> ), bacto-peptone (20 gl <sup>-1</sup> ), 100 mM potassium phosphate pH 6.0, yeast nitrogen base (13.4 gl <sup>-1</sup> ), biotin (40 µgl <sup>-1</sup> ) <sup>b</sup> and 0.5% methanol (v/v)  |
| PTM1 trace salts <sup>b</sup>                | CuSO <sub>4</sub> ·5H <sub>2</sub> O (6.0 gl <sup>-1</sup> ), KI (0.8 gl <sup>-1</sup> ), MnSO <sub>4</sub> ·H <sub>2</sub> O (3.0 gl <sup>-1</sup> ), Na <sub>2</sub> MoO <sub>4</sub> ·2H <sub>2</sub> O (0.2 gl <sup>-1</sup> ), H <sub>3</sub> BO <sub>3</sub> (0.2 gl <sup>-1</sup> ), CaSO <sub>4</sub> ·2H <sub>2</sub> O (0.5gl <sup>-1</sup> ), ZnCl <sub>2</sub> (20gl <sup>-1</sup> ), FeSO <sub>4</sub> ·7H <sub>2</sub> O (65 gl <sup>-1</sup> ), biotin (0.2gl <sup>-1</sup> ), H <sub>2</sub> SO <sub>4</sub> (5mL <sup>-1</sup> ) |
| Basal salt media (fermentation) <sup>a</sup> | Phosphoric acid 85 % (26.7 mL <sup>-1</sup> ), CaSO <sub>4</sub> ·2H <sub>2</sub> O (0.93gl <sup>-1</sup> ), K <sub>2</sub> SO <sub>4</sub> (18.2 gl <sup>-1</sup> ), MgSO <sub>4</sub> ·7H <sub>2</sub> O (14.9 gl <sup>-1</sup> ), KOH (4.13 gl <sup>-1</sup> ) and glycerol (40 gl <sup>-1</sup> )   |
| Glycerol feed solution                       | Glycerol <sup>a</sup> 50% (w/v), 12 mL <sup>-1</sup> PTM1 trace salts   |

|                        |   |
|------------------------|---|
| Methanol feed solution | Methanol 100% (w/v), 12 mL <sup>-1</sup> PTM1 trace salts |
|------------------------|---|

<sup>a</sup> Autoclaved to sterilise for 20 minutes at 121°C

<sup>b</sup> Filter Sterilised through a 0.22µm Filter

**Table 2.2: Solutions and buffers**

| Solution/Buffer                                      | Constituents   |
|--|--|
| <b>SDS-PAGE gels, Coomassie and western analysis</b> |  |
| 3xSDS-PAGE loading dye                               | 150 mM Tris-HCl pH 6.8, 6 % SDS (w/v), 30% glycerol (w/v), 0.05% bromophenol blue, 100 mM DTT  |
| 5xSDS-PAGE running buffer                            | 125 mM Tris, 960 mM glycine, 0.5 % SDS (w/v) (pH should be 8.3, do not adjust)   |
| Fixing solution                                      | Acetic acid 10 % (v/v), methanol 40 % (v/v)  |
| Coomassie blue solution                              | Acetic acid 10 % (v/v), methanol 40 % (v/v), Coomassie brilliant blue R-250 0.1% (w/v)   |
| Destain solution                                     | Acetic acid 10 % (v/v), methanol 10 % (v/v)  |
| 1xTransfer buffer (western blot)                     | 25 mM Tris, 125 mM glycine   |
| TBS-T buffer   | 20 mM Tris pH 7.4, 0.1 % Tween 20 (v/v), 150 mM NaCl   |
| Blocking buffer                                      | 5 % (w/v) skim milk powder in TBS-T  |
| <b>L6, M and M variant purification buffers</b>      |  |
| 100x Protease inhibitors mix                         | 100 mM phenyl-methanesulfonyl fluoride (PMSF), 100 mM benzamidine, 100 mM p-aminobenzamidine, 500 mM ε-amino-n-caproic acid.                       |
| Cell wash buffer                                     | 20 mM Tris pH 7.5, 150 mM NaCl, 10mM EDTA  |
| CEX lysis buffer                                     | 20 mM MES pH 7, 100 mM NaCl, 10% (w/v) glycerol, 5 mM EDTA, 0.25 mM Triton X100 (TX-100), 1 mM dithiothreitol (DTT), and 1x protease inhibitor mix |
| CEX equilibration buffer                             | 10 mM MES pH 6.5, 10% (w/v) glycerol, 0.25 mM TX-100, 1 mM betamercaptoethanol (BME), NaCl to I=130 mM   |
| CEX wash buffer A                                    | 10 mM MES pH 7, 0.25 mM TX-100, 1 mM BME, NaCl to I=130 mM   |
| CEX wash buffer B                                    | 10 mM HEPES pH 8, 0.25 mM TX-100, 1 mM BME, NaCl to I=130 mM   |
| CEX elution buffer                                   | 10 mM CAPS pH-10.4, 0.25 mM TX-100, 1 mM BME, NaCl to I=130 mM   |

|  |   |
|--|---|
| Neutralisation buffer                          | 50 mM Tris pH-7.2, 0.25 mM TX-100, NaCl to I=130 mM   |
| Cell lysis buffer                              | 50 mM Tris pH 8, 150 mM NaCl, 10% (w/v) glycerol, 0.25 mM TX-100, 20 mM BME and 1x protease inhibitor mix           |
| NiA buffer A                                   | 20 mM Tris pH 8, 130 mM NaCl, 0.25 mM TX-100, 5 mM BME  |
| NiA buffer B                                   | As NiA buffer A, with 250 mM imidazole  |
| GF buffer                                      | 20 mM Tris pH 7.5, 150 mM NaCl, 10% (w/v) glycerol, 10 mM magnesium acetate, 1 mM DTT                               |
| <b>AvrM and avrM purification buffers</b>      |   |
| Effector lysis buffer                          | 20 mM HEPES pH 7.5, 300 mM NaCl, 10 mM imidazole, 20 mM BME and 1x protease inhibitor mix                           |
| Effector equilibration buffer                  | 20 mM HEPES pH 7.5, 300 mM NaCl, 10 mM imidazole, 5 mM BME  |
| Effector wash buffer                           | 20 mM HEPES pH 7.5, 300 mM NaCl, 65 mM imidazole, 5 mM BME  |
| Effector elution buffer                        | 20 mM HEPES pH 7.5, 300 mM NaCl, 250 mM imidazole, 5 mM BME   |
| <b>Assay buffers</b>                           |   |
| ATPase assay buffer                            | 50 mM Tris-HCl pH 7.5, 50 mM NaCl, 10% (w/v) glycerol, 10 mM MgSO <sub>4</sub> , 1.5 mM DTT                         |
| Pyruvate kinase buffer                         | 125 mM Tris-acetate pH 7.4, 5 mM phosphoenolpyruvic acid (PEP), 2.5 mM MgSO <sub>4</sub> , 125 U/ml pyruvate kinase |
| Thin layer chromatography (TLC) running buffer | 0.25 mM LiCl, 5 % formic acid   |

I= Total ionic strength of the buffer

NB: Reducing agents DTT/BME and protease inhibitors were added to buffers directly prior to use.

### 2.1.2 Yeast and bacterial stains used in this study

The yeast and bacterial strains used for protein expression and cloning are summarised in Table 2.3

**Table 2.3: Yeast and bacterial strains**

| Yeast                           | Use                |
|---------------------------------|--------------------|
| <i>P. pastoris</i> X-33         | Protein expression |
| Bacteria                        |                    |
| <i>E. coli</i> DH10B            | Regular cloning    |
| <i>E. coli</i> BL21 pLysS (DE3) | Protein expression |

### 2.1.3 Oligonucleotides

The oligonucleotides used for cloning purposes in this study are listed in Table 2.4 (see Appendix 1 for sequencing primers)

**Table 2.4: Cloning oligonucleotides**

| Oligonucleotide name | Sequence (5'-3')  |
|----------------------|---|
| MrLRRstopnot         | ACTTGC GGCCGCTCACTCCCAAGCATATAGC                                      |
| Mf1571               | CGATGATGTTGATGAGAAG   |
| Mf $\Delta$ TIR      | CGTTTTCGAAAGAATGGGACATCATCATCACCATCATCAC<br>CATCATCAAGGAGCTATAGCAGATG |
| Mr1928               | TAGCCTATCATAAACCTCATCAAGG   |
| MfD364E              | GTCGTTCTCGATGAAGTTGATGAGAAGTTT  |
| MrD364E              | AAACTTCTCATCAACTTCATCGAGAACGAC  |
| MfK286L              | GGAATAGGCCTGACAACACTACTGC   |
| MrK286L              | GCAGTAGTTGTCAGGCCTATTCCG  |
| MfD555V              | GAAATGCACGTCCAACCTTAG   |
| MrD555V              | CTAAGTTGGACGTGCATTTCTAAC  |
| MfS492R              | GATAGGCTAAAAATAAGATACGATGCCTTG  |
| MrS492R              | CAAGGCATCGTATCTTATTTTTAGCCTATC  |
| avrMfC-terminal stop | GAAATCTCCAGACATGTGAAGATATAAGAG  |
| avrMrC-terminal stop | CTCTTATATCTTCACATGTCTGGAGATTTTC                                       |

### 2.1.4 Plasmids

The plasmids used in this study for protein expression are listed in Table 2.5

**Table 2.5: Plasmids used in this study**

| Plasmid Name                   | Description   | Contain  |
|--------------------------------|---|--|
| pPICz (zeocin resistance)      | Contains the M and L6 cDNA, minus the first 21 and 28 amino acids, respectively. An N-terminal 9x histidine tag facilitates metal ion affinity purification. For cloning details refer to Schmidt et al, (2007) | For protein expression of; M, M <sup>K286L</sup> , M <sup>D555V</sup> , M <sup>K286L+D555V</sup> , M <sup>S492R</sup> , M $\Delta$ TIR, M $\Delta$ LRR, L6 and L6 $\Delta$ TIR |
| pET15b (ampicillin resistance) | Contains the AvrM and avrM cDNA, minus the signal peptides, amino acids 29-343 and 29-314 respectively. Both were cloned into the <i>EcoRI</i> site separately.   | AvrM, avrM and avrMC* for protein expression   |

## **2.2 General Procedures**

### **2.2.1 Preparation of electrocompetent and heat shock *E. coli***

Electrocompetent *E. coli* (DH10B) and heat shock competent *E. coli* (BL21 pLYSs) were prepared in accordance with Sambrook et al., (1989).

### **2.2.2 Transformation of competent *E. coli***

#### **2.2.2.1 Electroporation of DH10B**

Plasmid or ligated DNA (1  $\mu$ l, 1-100 ng/ $\mu$ l) was mixed with a 20  $\mu$ l aliquot of electrocompetent *E. coli* cells and incubated for 5 minutes on ice. Cells were electroporated in accordance with the manufacturer's protocol (Cell Porator, BRL). Following electroporation, cells were transferred into non-selective LB and incubated at 37°C for 45mins. Cells were then plated onto the appropriate selective media and incubated for 16-20 hours at 37°C.

#### **2.2.2.2 Heat Shock of BL21 pLysS**

Plasmid DNA (1 $\mu$ l, 100ng/ $\mu$ l) was added to 50  $\mu$ l of heat shock competent *E. coli* and incubated on ice for 30 minutes. Heat shock was performed at 42°C for 45 seconds. Following heat shock, 50 $\mu$ l of non-selective LB was added to the reaction and the culture was incubated at 37°C for 1 hour with shaking. Cells were then plated onto selective media (100  $\mu$ g/ml ampicillin and 50  $\mu$ g/ml chloramphenicol for BL21 pLysS) and incubated for 16-20 hours at 37°C.

### **2.2.3 Colony PCR**

Yellow tips were used to pluck single antibiotic resistant colonies from the plate and the bacteria were mixed with 20  $\mu$ l of sterile water. This mixture was heated to 98°C for 5 minutes, and 5  $\mu$ l of this was used as the template for polymerase chain reaction (PCR). PCRs were typically 20  $\mu$ l reactions containing template DNA, 1-0.5 units of standard Taq polymerase (New England Biolabs), 50 ng of forward and reverse primers and 1 mM dNTPs. All PCRs were performed in a Gene Amp DNA system (PerkinElmer) PCR machine. The reactions for colony PCR were typically; 94°C for 2 minutes, then 25 cycles of 94°C for 10 seconds, 54°C for 10 seconds, 72°C for 45 seconds and finally 72°C for 7 minutes. Annealing temperatures and extension times varied depending on primer combinations and product size.

### 2.2.4 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was performed largely as described previously (Laemmli, 1970). Typically 40  $\mu$ l of protein sample was mixed with 20  $\mu$ l of 3x SDS-PAGE sample buffer and boiled for 5 minutes. Proteins were separated on 10% polyacrylamide gels. Electrophoresis was performed in 1x SDS-PAGE running buffer for 1-1.1 hours at 170 volts using Biorad 'Mini-Protean® Tetra Cell' gel electrophoresis unit. On each gel, a prestained molecular weight marker (NEB) was used to enable the molecular weights of proteins of interest to be estimated. Gels were then used for either Coomassie, silver or Sypro Ruby staining, or transferred onto nitrocellulose membrane for western analysis.

### 2.2.5 Coomassie staining of proteins in polyacrylamide gels

Following SDS-PAGE gels were fixed in fixing solution for 30 minutes. Gels were then stained for 1 hour in Coomassie stain followed by destaining for 2-3 hours in destain solution.

### 2.2.6 Silver staining of proteins in polyacrylamide gels

The silver staining procedure was adapted from the following method (Blum et al., 1987). In brief, SDS-PAGE gels were fixed for 30 minutes followed by a series of washes (3-5 x 5 minutes) with milli-Q H<sub>2</sub>O to remove any residual fixer. Gels were incubated for 2 minutes in 0.2 g l<sup>-1</sup> sodium thiosulphate, washed (3 x 30 seconds) with milli-Q H<sub>2</sub>O and incubated for 25 minutes in silver nitrate (2 g l<sup>-1</sup>). Gels were washed (3 x 60 seconds) before being developed in 100 ml of developer (sodium carbonate 30 g l<sup>-1</sup>, 37% (v/v) formaldehyde, 2% (v/v) (0.2 g l<sup>-1</sup>) sodium thiosulphate). Development (a maximum of 10 minutes) was stopped using 0.5 M EDTA; gels were stored in water.

### 2.2.7 Western blot

Following SDS-PAGE, gels were equilibrated in cold transfer buffer for 10 minutes before being loaded into a BioRad Mini Trans-Blot® apparatus. Proteins were transferred from SDS-PAGE gels to a nitrocellulose membrane (Hybond-ECL, GE Biosciences) as previously described (Towbin et al., 1979); transfer was performed at 60 volts for 90 minutes. Following transfer, membranes were washed in blocking buffer for 1 hour. Membranes were then incubated for 45 minutes with primary antibodies (in blocking buffer). The primary antibodies used in this study included: The anti-M antibody, a rabbit polyclonal affinity purified antibody (Schmidt *et al*, 2007) was used at a 1:2500 dilution. The anti-His antibody, a rabbit polyclonal affinity purified antibody (Rockland) was used at a 1:5000 dilution. For all experiments two 15 minute wash steps in blocking buffer were followed with a 30 minute incubation with horseradish peroxidase-conjugated

secondary anti rabbit immunoglobulin G (Rockland) at a 1:10,000 dilution. The membranes were washed twice for 15 minutes in TBS-T. This was followed by a 2 minute incubation in Supersignal® west pico chemiluminescent substrate (Thermoscientific). Detection was facilitated by exposure to KODAK X-Omat-XK1 film, and film development using KODAK RP X-Omat developer/fixer or the image was captured (light output) using a VersaDoc gel documentation instrument (Biorad).

### 2.2.8 Protein concentration

Amicon Ultra Centrifugal Devices (Millipore) were generally washed in the same buffer that proteins were being concentrated in. Proteins to be concentrated were added to the concentrator and centrifuged up to 3,500 g in a bench top centrifuge at 4°C. Buffer exchanges into assay buffers were also performed during concentration.

### 2.2.9 Protein quantification

Protein concentrations were determined using the Bradford reagent (BioRad) and assayed as described by Bradford (1976).

## 2.3 Generation of Mutated Variants of *M* for Expression in *P. pastoris*

### 2.3.1 *M* truncated mutants

#### 2.3.1.1 Generation of *M*ΔTIR

The TIR was removed from the *M* cDNA by PCR. To achieve this, a large oligonucleotide (MfΔTIR, Table 2.4) was designed to include a *Bst*BI restriction enzyme site, followed by a start codon, then a 9x histidine tag and a stretch of 19 nucleotides that annealed at the intron 1 junction point between the sequence encoding the TIR and NB-ARC domains, starting at amino acid Q225. *M* cDNA without a histidine tag in the pPICz vector (designated 1M, Schmidt et al., 2007) was used as the template for the polymerase chain reaction (PCR). The MfΔTIR in combination with the reverse primer Mr1928, which anneals 3' of a *Bst*BI site located within the NB-ARC encoding domain, generated a fragment of approximately 800 bp. Amplification was carried out with a high fidelity DNA polymerase, phusion DNA polymerase (Finnzymes), in accordance with the manufacturer's recommendations. Cycles were as follows: 98°C for 2 minutes, then 25 cycles of 98°C for 10 seconds, 54°C for 30 seconds, 72°C for 30 seconds and finally 72°C for 7 minutes. This PCR product and the 7M construct (described by Schmidt et al., 2007b) were both digested with *Bst*B1 in accordance with the manufacturer's protocols. Following restriction enzyme digestion, a MoBio ultraclean™ PCR clean-up DNA purification kit

was used to clean up the DNA of both digests. The digested 7M DNA was then de-phosphorylated (CIP treatment) and cleaned up again prior to ligation with the PCR generated *BstB1* digested insert. 1  $\mu$ l of the ligation reaction was then used to transform electrocompetent *E. coli*. The zeocin (Invitrogen) resistant colonies were screened using colony PCR and plasmid DNA from positive clones was sequenced to verify the correct insert.

### 2.3.1.2 Generation of M $\Delta$ LRR

The removal of the LRR from the 7M expression construct was also done by PCR. The PCR reaction used to generate this small insert involved a reverse primer, defined as MrLRRstopnot (Table 2.4). This primer contained an inbuilt stop codon and a *NotI* site. The forward primer used in the reaction was MF1571 (Table 2.4) which annealed down stream of an *SspI* site in the M gene. Amplification was carried out with high fidelity DNA polymerase, phusion DNA polymerase (Finnzymes), in accordance with manufacturer's recommendations. Cycles were as follows: 98°C for 2 minutes, then 25 cycles of 98°C for 30 seconds, 54°C for 15 seconds, 72°C for 45 seconds and finally 72°C for 7 minutes. PCR products and 7M plasmid DNA were digested with *NotI* and *SspI* in accordance with the manufacturer's protocols. The DNA was then cleaned up as described above and de-phosphorylated (CIP treatment) and cleaned again prior to ligation with the generated insert. 1  $\mu$ l of the ligation reaction was then used to transform electrocompetent *E. coli*. The zeocin (Invitrogen) resistant colonies were screened using colony PCR and plasmid DNA sequenced to confirm correct cloning.

### 2.3.2 Site-direct mutagenesis to produce rational point mutations in M

In order to study biochemical effects of specific residues in the M protein, site-direct mutagenesis was used to introduce point mutations into the M gene. Miss-match primers were designed to introduce these desired changes (Table 2.4). A PCR using high fidelity DNA polymerase, phusion DNA polymerase (Finnzymes), with a long extension time enabled the amplification of the entire pPICz plasmid including the M gene incorporating the desired point mutation(s). Table 2.4 includes all miss-match primers. The template used for all single mutation was M cDNA, however, for the double mutant M<sup>K286L+D555V</sup>, the M<sup>K286L</sup> template was used. Therefore, the M<sup>K286L+D555V</sup> mutant required two independent rounds of site-directed mutagenesis for its production. Typically 100 ng of template was used in the PCR reaction and conditions were as follows: 98°C for 2 minutes, then 25 cycles of 98°C for 10 seconds, between 50-60°C (T anneal = T<sub>m</sub>-2°C) for 30 seconds, 72°C for 5 minutes and finally 72°C for 7 minutes. PCR products were then subjected to *DpnI* digest to remove any remaining template. The PCR products were then electroporated into electrocompetent *E. coli*. Single colonies were selected and plasmid DNA

was subjected to full gene sequencing to ensure the desired change had been incorporated and that no point mutations were present. In this way the following mutations M<sup>K286L</sup>, M<sup>D555V</sup>, M<sup>K286L+D555V</sup> and M<sup>S492R</sup> were generated in the pPICz vector, all lacking the 5' sequence encoding the first 21 amino acids of M to aid in purification of soluble protein (Schmidt et al., 2007b).

### 2.3.3 Preparation of electrocompetent *P. pastoris*

Overnight cultures of *P. pastoris* (strain X33) were grown in 10 ml of YPD at 30°C. 5 ml of this culture was used to inoculate 250 ml YPD media, which was grown to an OD<sub>600</sub> of 0.6-0.8. Cells were harvested by centrifugation at 3000 g for 10 minutes at 4°C in a Sorvall GSA rotor. The cells were washed in 500 ml of sterile ice cold water, followed by a 250 ml wash before being resuspended in 20 ml of ice cold 1 M sorbitol. The cells were then centrifuged at 1,500 g in a bench top centrifuge (Jouan CR312) and the pellet resuspended to form a pipette-able paste, using a small volume of 1 M sorbitol. Cells were maintained on ice and used for transformation on the same day.

### 2.3.4 Electroporation of *P. pastoris*

5-10 µg of vector DNA containing the construct of interest was linearised by digestion with *PmeI* (NEB) in accordance with manufacturer's protocol. Digested DNA was cleaned and concentrated by ethanol precipitation (Sambrook et al., 1989) and resuspended in 10 µl of water. 80 µl of electrocompetent *P. pastoris* was mixed with 10 µl of digested plasmid and transferred to an ice-cold 0.2 cm electroporation cuvette (BioRad). Electroporations were conducted with a GenePulser II (Biorad) at a charging voltage of 1500 volts, capacitance 25 µF and the resistance set at 200Ω. 1 ml of 1 M sorbitol was added to cells immediately after electroporation, which were then incubated for 1 hour at 30°C without shaking. A further 1 ml of 2x YPD was added to the cells and they were incubated for a further 1.5 hours at 30°C with shaking. Positive transformants were selected for by plating onto 100 µg/ml zeocin YPD agar, and incubating for 48-60 hours at 30°C.

### 2.3.5 Screening and test expression of transformed *P. pastoris*

Typically, 8 well separated *P. pastoris* colonies, growing on YPD-zeocin transformation plates, were selected and re-streaked to single colonies on YPD (100 µg/ml zeocin) agar and simultaneously used to inoculate 10 ml of BMGY (100 µg/ml zeocin). YPD plates were incubated for 48-60 hours at 30°C before being refrigerated. BMGY cultures were incubated with shaking for 48 hours, harvested by centrifugation (3000 g, 10 minutes) and resuspended in 10 ml BMMY (100 µg/ml ampicillin). Cultures were incubated for 72 hours at 15°C in BMMY, with the addition

of methanol to a final concentration of 0.5% (v/v) every 24 hours. Cells were pelleted by centrifugation and stored at -80°C or used fresh for test expression.

Test expression screens involved resuspending the cell pellets in an equal volume of cell lysis buffer and acid washed glass beads. Cells were subjected to 5 minutes of vortexing at high speed at 4°C and lysates were cleared by centrifugation at 14,000 g. As determined by Bradford assay, 30 µg of total protein from crude lysates was mixed with SDS-PAGE sample buffer and subjected to SDS-PAGE separation and western analysis. Transformants were selected on the basis of western signal intensity and correct protein size.

## **2.4 Growth and Expression of R proteins**

### **2.4.1 Shaker flask expression**

A *P. pastoris* colony that had been demonstrated by test expression to be successfully expressing a desired protein was inoculated into 10 ml of BMGY (100 µg/ml zeocin). These were incubated with shaking (200 rpm) at 30°C overnight. This starter culture was then used to inoculate between 100-500 ml BMGY containing 100 µg/ml ampicillin, which was incubated with shaking at 30°C for 48 hours. Cells were harvested by centrifugation at 3,000 g for 10 minutes in a Sorvall GSA rotor, the supernatant discarded, and the cells resuspended in BMMY induction media containing 100 µg/ml ampicillin. After incubation at 15°C for 72 hours (with the addition of 0.5% methanol every 24 hours) cells were harvested as above. Cells were washed with wash buffer (Table 2.2), and snap frozen in liquid nitrogen and stored at -80°C or used immediately. Typically, 6 g of wet cell pellet was obtained per 100 ml of culture media.

### **2.4.2 Fermentation expression**

A 10 ml culture of BMGY (100 µg/ml zeocin) was inoculated from a *P. pastoris* colony and incubated with shaking at 30°C for 24 hours. This culture was used to inoculate 200 ml BMGY (100 µg/ml zeocin) which was used to seed the fermentation culture. 2.5 L of Basal Salts fermentation media (Stratton *et al*, 1998) was added to the fermentor vessel (New Brunswick Scientific) and autoclaved. After the vessel was cooled to room temperature, the media was equilibrated to pH 5.0 with ammonium hydroxide. 12 ml of trace salts (Stratton *et al*, 1998) and 100 mg ampicillin was added to the fermentor. During fermentation, the media was maintained at pH 5.0. The fermentor was inoculated with the 200ml BMGY culture and a 40% dissolved oxygen rate was maintained (aeration and oxygen feed) over the duration of the fermentation. 18-24 hours post inoculation (when residual media glycerol was exhausted), a glycerol feed of 10 ml per hour (2 ml trace salts per 500 ml glycerol) was added to the culture. The length and rate of

glycerol feed can be determined at the discretion of the user, however, typically a 5% glycerol feed was maintained for 48 hours. At completion of the glycerol feed, the cells were starved for 6-10 hours before induction with methanol. This starvation step ensures that all carbon sources are used before induction. The culture was chilled to 12°C and an initial methanol feed rate was set at 1-1.5% (2-3 ml/hour) for 12-16 hours with a gradual step-wise increase to between 5 and 6% (at the discretion of operator) over the following 8 hours. The methanol feed rate was calculated on the basis of the methanol consumption rate, to avoid excess methanol build up. Methanol concentrations at or above 5% are toxic to the cells. Methanol feed was generally maintained for 48 hours at the upper methanol feed rate. Cells were harvested from a culture that was approximately 4L after glycerol and methanol feeds by centrifugation at 4,000 g at 4°C for 15 minutes in a Sorvall GSA rotor and washed with wash buffer (Table 2.2). Aliquots of 50 g or 100 g of cells were snap frozen in liquid nitrogen and stored at -80°C.

## **2.5 Cell Lysis**

### **2.5.1 Cell disruptor**

#### **2.5.1.1 *P. pastoris***

*P. pastoris* cells from fermentation expression were resuspended in a 5x volume of CEX Lysis buffer. Cells were lysed by three sequential passes through a pre-cooled cell disruptor (5,000-10,000 psi, without exceeding 15,000) (EmulsiFlex-05 Homogenizer, Aventis). Un-lysed cells, cell debris and nuclear material were removed from the lysate by centrifugation at 10,000g for 15 minutes at 4°C in a Sorvall GSA. The supernatant was then spun at 100,000 g for 45 minutes at 4°C in a Beckman Ti60 or Ti45 rotor. Alternatively, the supernatant was filtered through a pre-filter and then a 0.45 µm (Millipore) filter. The resulting lysate was then applied to various chromatography columns for purification of the expressed protein.

#### **2.5.1.2 Protein expression in *E. coli***

*E. coli* cells expressing AvrM, avrM and avrMC\* were resuspended in 7x the pellet volume of effector lysis buffer. Cells were lysed by three sequential passes through a pre-cooled cell disruptor (10,000-15,000 psi). Cell lysates were cleared by centrifugation at 100,000 g for 1 hour at 4°C. The resulting lysate was then applied to various chromatography columns for purification of the expressed protein.

## 2.5.2 French press

*P. pastoris* cells from shaker flask expression (typically 3-6g) were resuspended in 3x the pellet volume of Cell Lysis buffer. Cells were lysed by three sequential passes through a pre-cooled French Press (Aminco) (5000-8000 psi). Un-lysed cells, cell debris and nuclear material were removed from the lysate by centrifugation at 10,000 g for 15 minutes at 4°C in a Sorvall GSA. Then Supernatant was ultracentrifuged at 100,000 g for 45 minutes at 4°C in a Beckman Ti60 or Ti45 rotor. Alternatively, the supernatant was filtered through a pre-filter and then a 0.45 µM (Millipore) filter. The resulting lysate was then applied to various chromatography columns for purification of the expressed protein.

## 2.6 Purification of M, L6 and M Variants

Purification steps were performed using the AKTA Explorer FPLC (Pharmacia, GE Scientific).

### 2.6.1 Gel filtration (GF)

Gel filtration (also known as size exclusion chromatography) was used for the analysis of crude lysates, or as a downstream purification step of the M, L6 and M variants. A number of GF columns were used; S200 HR (Superdex 200 HR 10/30 analytical), S200PG (HiLoad™16/60 Superdex 200™ prep grade) and S400 (Sephacryl 400™ HR). All columns originated from Amersham biosciences/Pharmacia. Prior to sample loading, columns were equilibrated with gel filtration buffer, with some variations in buffer type depending on the experimental design. The equilibration buffer was the same as the running buffer. Flow rates ranged between 0.3-2 ml/min, depending on the manufacturer's protocol. All columns were calibrated with proteins of known molecular weight (Sigma) to produce standard curves and enable size estimations of the proteins of interest from elution profiles (Appendix 2).

### 2.6.2 Large scale purification

*P. pastoris* cells expressing the M or L6 gene were lysed in a MES-based buffer (2.4.1). A cation exchange column (CEX) (500 ml, SP-Sepharose Big beads, Pharmacia) was equilibrated with equilibration buffer (Table 2.2), in preparation for the first step in the purification protocol. Lysates (pH 6.5) were loaded onto the column and washed extensively with 5 column volumes (c/v) of CEX wash buffer A and 5 c/v of CEX wash buffer B. An increase in pH was used to elute the protein from the resin (elution buffer). The elution volume from CEX was generally 0.6-0.7x the total lysate volume. The CEX eluate was neutralised with neutralisation buffer (Table 2.2) and 5 mM imidazole was added in preparation for the second purification step of nickel-metal ion affinity (NiA) purification chromatography.

The Ni column was equilibrated with Ni buffer A and 5 mM imidazole. Neutralised CEX elutes were then loaded onto the Ni column. The Ni column was then washed with 10 column volumes (c/v) of Ni buffer A containing 35 mM imidazole. Protein was eluted from the column with NTA buffer B, which contained 250 mM imidazole. SDS-PAGE, Coomassie staining and western analysis were performed on fractions of interest from CEX and NiA purification. Protein eluted from the NiA column could be concentrated (2.2.8) and further separated by gel filtration.

### 2.6.3 Small scale purification

*P. pastoris* cells expressing the M, L6 or variants of the M protein were lysed in a cell lysis buffer for purification using NiA. Cell lysates were titrated with a high pH lysis buffer to 7.7-8. The size of the NiA column used depended on the amount of cells lysed. Typically, a 3 ml column was used to purify a 3 or 6 g pellet. The NiA column was pre-equilibrated with NiA buffer A containing 10 mM imidazole prior to loading of the cleared cell lysates. After the cell lysate was loaded, the column was washed with NiA buffer A containing 55 mM imidazole in an attempt to remove non-specific and weakly bound proteins. Two wash strategies were used; a short wash involved 10 c/v, while a long wash involved 100 c/v. M, L6 or M-variant proteins were eluted with a 250 mM imidazole elution step (NiA buffer B). The total elution collected was 5 c/v, or 15 ml from a 3 ml column. Protein was then concentrated for use in biochemical assays or subject to further purification/analysis utilising gel filtration. During the course of the project it was found that the most effective method to purify flax R proteins involved their elution from NiA directly into GF or ATPase assay buffer to limit the proteins exposure to high imidazole concentrations. A reduction in imidazole to 75 mM total concentration at this point helped to prevent protein aggregation induced by protein concentration while exposed to high concentrations of imidazole.

## 2.7 Cloning, Expression and Purification of AvrM/avrM

The AvrM and avrM gene, minus the 5' sequence encoding the signal peptide were cloned into the pET15b vector and sourced from Dr. Peter Dodds (CSIRO, Plant Industry, Canberra). Both cDNAs were cloned in frame with a sequence that encoded an n-terminal 6x histidine tag followed by a thrombin cleavage site (Appendix 6).

### 2.7.1 Cloning of avrMC\* into pET vector

A C-terminal truncation was designed in the avrM gene by engineering a stop codon at the equivalent position to that of AvrM (Appendix 6). To achieve this, mismatch primers, avrMF- and avrMR-C-terminal stop (Table 2.4) were designed to incorporate a premature stop codon at the required position. The PCR contained 50 ng of avrM pET template, 100ng of miss-match primers and utilised high fidelity Phusion DNA polymerase with appropriate buffers. The reaction

conditions were as follows: 98°C for 2 minutes, then 25 cycles of 98°C for 10 seconds, 60°C for 30 seconds, 72°C for 5 minutes and finally 72°C for 7 minutes. PCR products were then subjected to a *DpnI* digest to remove remaining template. The PCR products were then transformed into electrocompetent *E. coli* cells. Clones were analysed by full gene sequencing to ensure the desired change had been incorporated and that no other point mutations had been added.

### 2.7.2 Transformation of BL21 pLysS with AvrM/avrM

AvrM, avrM and avrMC\* pET constructs were transformed into BL21 pLysS cells using heat shock transformation (2.2.2.2). Cells were plated onto LB plates with 100 µg/ml ampicillin and 50 µg/ml chloramphenicol. Positive transformants were used in test expression experiments.

### 2.7.3 Growth and test expression

Typically, four positive transformants were selected for test expression. Single colonies were used to inoculate 2 ml of LB broth containing 100 µg/ml ampicillin and 50 µg/ml chloramphenicol and incubated overnight at 37°C with shaking at 200rpm. The following day, 50 µl of the overnight culture was used to inoculate 1 ml of fresh LB. This was incubated under the same conditions as the overnight culture but for 2 hours. To induce protein expression, isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM. Cultures were incubated for a further 2 hours, at which point, 20 µl of the culture was analysed by SDS-PAGE, Coomassie staining and western analysis.

### 2.7.4 Large scale expression

1L cultures were generally used to generate protein expressing cells for subsequent protein purification. A 20 ml starter culture (LB containing 100 µg/ml ampicillin and 50 µg/ml chloramphenicol) was grown overnight at 37°C with shaking at 200 rpm. The following morning, 4 x 250 ml of LB plus antibiotic was inoculated each with 5 ml of starter culture. These cultures were grown under the same conditions as the overnight culture until the OD<sub>600</sub> reached between 0.6-0.8. To induce protein expression, IPTG was added to a final concentration of 1 mM and the cultures were incubated at 18°C overnight. Expressing cells were harvested the following morning by centrifugation at 4,000g at 4°C for 15 minutes in a Sorvall GSA rotor and washed with wash buffer (Table 2.2) and snap frozen in liquid nitrogen and stored at -80°C ready for extraction (2.5.1.2).

### 2.7.5 Purification

Cleared cell lysates containing the protein of interest were subject to NiA chromatography. Typically, lysates were loaded over a 3 or 5 ml column that had been pre-equilibrated with effector equilibration buffer. The column was washed with 30 c/v of effector wash buffer, and eluted with 5 c/v of effector elution buffer. Protein collected in the elution was concentrated and further purified and buffer exchanged using S200 PG GF.

## 2.8 Biochemical Assays

### 2.8.1 ATPase assay

Samples of purified protein were concentrated and buffer exchanged into ATPase buffer. Reaction volumes were adjusted to 50  $\mu\text{l}$  with the addition of 5  $\mu\text{l}$  of [ $\alpha^{32}\text{P}$ ] ATP (300 mCi/ $\mu\text{l}$ ) to give a final ATP concentration of either 5  $\mu\text{M}$  or 20  $\mu\text{M}$ . Assays were generally performed in triplicate for each protein sample. Assays were run at 25°C and samples were taken at various time points up to 1 hour. Samples of the reaction were stopped by spotting 2  $\mu\text{l}$  of assay mix onto PEI cellulose thin layer chromatography (TLC) plates (Merck). Plates were run in a TLC chamber using TLC buffer (Table 2.2) until the buffer front reached the end of the plate. The amount of [ $\alpha^{32}\text{P}$ ] ATP and [ $\alpha^{32}\text{P}$ ] ADP was quantified from the TLC plate using a Phosphor Imager instrument (Molecular Imager FX, Biorad).

### 2.8.2 ATP/ADP quantification assays using luminescence

Following purification by NiA and/or GF, purified M or L6 protein samples were heated to 98°C for 5 minutes (performed in triplicate) followed by centrifugation at 14,000 g for 2 minutes at 4°C to precipitate the protein and release bound ATP or ADP. 100  $\mu\text{l}$  of each sample was added to 150  $\mu\text{l}$  of pyruvate kinase buffer. For ADP measurements, 125 U/ml of pyruvate kinase (PK) was added to enable the conversion of ADP into ATP (ATP is the readable form of nucleotide in the assay). ADP samples were incubated at room temperature for 30 minutes. These samples were then boiled for 5 minutes and centrifuged at 14,000 g for 2 minutes at 4°C to remove PK. Luminescence assays were performed that involve the following reaction:  $\text{ATP} + \text{luciferin} + \text{O}_2 \xrightarrow{\text{luciferase}} \text{AMP} + \text{Oxyluciferin} + \text{CO}_2 + h\nu \text{ (light)}$ , where in saturating levels of luciferin, light output (luminescence) is proportional to ATP concentration. 5  $\mu\text{l}$  of sample was added to 100  $\mu\text{l}$  of ATP assay mix, prepared at a 1:10 dilution in the dilution buffer supplied (Sigma). The light output was measured, in mV, using a 1250 luminometer (BioOrbit). 5  $\mu\text{l}$  of a known concentration of ATP was then added to the sample to act as an internal standard. The ATP concentration in the sample was calculated as follows; the output of light from the sample, divided

by that of the standard, multiplied by the ATP concentration of the standard. For the ADP (ATP + PK) sample, the ADP concentration was measured as the difference between the sample for which PK was added and that where it was not.

### **2.8.3 M protein quantification using VersaDoc imaging system**

To compare protein concentrations between protein preparations and different M variants for biochemical analysis, a gel-based protein quantification system was used. Equal volumes of proteins of interest were separated by SDS-PAGE. Gels were stained overnight in the dark with the fluorescent stain Sypro Ruby (Biorad). After destaining in 10% (v/v) methanol, 7% (v/v) acetic acid for 1 hour, the gels were visualised and the image captured using the VersaDoc™ imaging system, using the following Sypro Ruby settings; 520LP UV TRANS with 1xgain and 1x1bin with an exposure time of between 30-60 seconds. Protein bands were quantified using the Quantity one software package (Biorad). Quantification involved the measuring of an adjusted volume (Intensity\*area) which, enables one to determine the intensity and the area of the protein bands. A bovine serum albumin standard curve was run on each gel and used to determine protein concentration of full-length M proteins run on the same gel and stained under the same conditions.

## **2.9 Interaction Assays**

For interaction assays, approximately 2-5 molar excess of effector proteins (AvrM or avrMC\*) compared to M proteins was used.

### **2.9.1 Co-immunoprecipitation**

M proteins were purified as detailed in 2.6.3; however, 0.5mM ATP and 10mM Magnesium acetate were added to lysis and all purification buffers including GF buffer. Pull-down experiments were performed with antibodies specific to the M protein using the anti-M antibody. The anti-M antibodies were purified as described by Schmidt et al., (2007b).

#### **2.9.1.1 Immunoprecipitation using Protein G agarose beads**

Reactions containing proteins of interest were incubated in GF buffer at 25°C for 15 minutes to enable protein/protein association. Protein G-PLUS agarose (Santa Cruz Biotechnology) (50 µl) was added to each sample for pre-clearing (to reduce non-specific binding) and incubated at 4°C for 30 minutes. The agarose was removed by centrifugation and the supernatant added to a fresh tube. 2 µg of affinity purified anti-M antibody was then added to each reaction which was incubated for 60 minutes at 4°C. Protein G agarose (50 µl) was added to the reaction which was then incubated at 4°C for 120 minutes. At the end of incubation, the protein G agarose was

pelleted by centrifugation and the supernatant removed. Pellets were washed twice with GF buffer and 1xSDS-PAGE buffer was added to the samples. Samples were analysed by SDS-PAGE, Coomassie staining and western analysis. The western analysis utilised a monoclonal anti-His antibody (TetraHis, Qiagen), to avoid cross-reactivity with the immunoprecipitation antibody which was raised in a rabbit.

### **2.9.1.2 Immunoprecipitation using protein A microbeads**

Reactions containing proteins of interest were incubated in GF buffer at 25°C for 15 minutes to enable any protein/protein association to take place. Samples were placed on ice and 2 µg of affinity purified anti-M antibody was then added to each reaction which was incubated for 20 minutes on ice. 50 µl of Protein A microbeads (Miltenyi Biotec) was added to each reaction to magnetically label the immune complex. Reactions were incubated on ice for 45 minutes. Bead-protein complexes were recovered by magnetic separation according to the manufacturer's protocol. For washing steps GF buffer was used.

### **2.9.2 Nucleotide exchange assays**

M protein (1 µM) was incubated with 5 µM AvrM in the presence of 0.5 mM AMPPNP for 15 minutes at 25°C. The protein samples were then put through an ice cold 1 ml NAP-10 (GE healthcare) desalting column to remove any excess unbound nucleotides. In short, generally 300 µl assays were applied to NAP-10 columns that had been pre-equilibrated with ice-cold GF buffer. When the entire sample had run onto the column, 700 µl of ice cold GF buffer was applied. After the GF buffer was completely loaded, 300 µl of GF buffer was added and simultaneously the 300 µl elution was collected. The sample was then processed for ATP/ADP quantification and protein quantification.

### **2.10 Multiple Sequence Alignment of the NB-ARC Region of M, other R proteins, Apaf-1 and CED-4**

The multiple sequence alignment was performed with the alignment program ClustalW (Thompson et al., 1994) using the BioManager by ANGIS database (<http://www.angis.org.au>). The protein sequences included in the alignment were as follows (NB: the residues and UniProtKB accession numbers are detailed); human Apaf-1 (residues 108-405, O14727), nematode CED-4 (residues 113-496, P30429); R proteins, flax M (residues 236-566, P93244), flax L6 (residues 221-552, Q40253), tobacco N (residues 171-500, Q40392), *A. thaliana* SSi4 (residues 168-496, Q8GUQ4), RPS4 (191-520, Q9XGM3), RPS2 (139-484, Q42484), RPM1 (154-516, Q39214), potato Rx (124-471, Q9XGF5) and tomato I-2 (153-506, Q9XET3), Mi-1

(505-852, O81137). The alignment was annotated further using the GeneDoc program (Nicholas et al., 1997).

### **2.11 Structure Model of the NB-ARC Domain of M**

The three dimensional structure model of the NB-ARC region of M was generated by Pradeep Sornaraj. To build a molecular model of the NB-ARC region of M, the structure of the NB-ARC region of Apaf-1 was used as a template. A paired sequence alignment was generated between the NB-ARC region of M and Apaf-1, in a similar fashion to that used for the multiple sequence alignment (2.10). The model of the NB-ARC domain of M was generated by the following steps. The GeneDoc alignment file (saved as a \*.pir file) was converted into an \*.ali file and \*.top file using the perl script converter pir.pl. These converted files were then processed through the modelling program Modeller9V2 (<http://www.salilab.org/modeller/>) to generate 50 different possible models using the default parameters. The PROCHECK module of the CCP4 package (Bailey, 1994) was used for the stereochemistry analysis of the generated models. The PROCHECK analysis gave detailed information on residue stereochemistry. Models with bad contacts in conserved regions were eliminated and the model with minimum bad contacts in non-conserved regions was chosen. PYMOL (<http://www.pymol.org>) was used to visualise the modelled structure and a PYMOL image of the structure was captured and is presented in Chapter 4 (Figure 4.11).

# **Chapter 3: Purification of the M and L6 Flax Rust Resistance Proteins Expressed in *Pichia pastoris*.**

This results chapter represents data, part of which, has been published in the following publication:

Schmidt, S. A., Williams, S. J., Wang, C. I. 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 Journal*, **50**, 1107-1117.

### 3.1 Introduction

A well recognised difficulty in obtaining soluble full-length recombinant R proteins has reduced the output of *in vitro* R protein functional studies. In one published biochemical study, attempts were made to produce full-length I-2 protein in *Saccharomyces cerevisiae* and *E. coli* expression systems, however, these attempts were unsuccessful (Tameling et al., 2002). In this case, a truncated version of the I-2 protein without the LRR domain could be produced in *E. coli* and the subsequent published reports have provided tremendous insight into the understanding of R protein activation, and the role of the NB-ARC domain (see 1.6.2/1.8) (Tameling et al., 2002, Tameling et al., 2006). However, it is generally accepted that to gain further insight into R protein function, biochemical and structural studies of full-length recombinant R proteins are required. Therefore, the development of a method to facilitate production of full-length R proteins is imperative.

It has been the long-term goal of the Anderson laboratory to express and purify the flax rust R proteins, M and L6, in near full-length form, to facilitate biochemical and structural studies. Previous work within our group of collaborators had demonstrated that the production of soluble full-length L6 protein could not be achieved using an *E. coli* expression system, although expression of the L6 TIR domain alone in *E. coli* has yielded soluble protein (J. Ellis and B. Kobe, personal communication). In order to try alternatives to the *E. coli* expression system, the *Pichia pastoris*-based protein expression system (Invitrogen) was chosen by the Anderson laboratory in the year 2000. A number of potential *P. pastoris* expression strategies, including; intracellular and secreted expression systems, changes to induction temperatures and variation in lysis conditions were investigated and reported by (Schmidt, 2002). This work culminated in the report of an expression system that enabled the production and purification of soluble M protein (Schmidt et al., 2007b).

This chapter will detail the purification of soluble recombinant M protein from the *P. pastoris* expression system. The same method of expression and purification for M can be used to produce soluble recombinant L6 protein. These results, combined with the expression studies, have been published (Schmidt et al., 2007b). Since this report, a number of changes have been made to the purification strategy as a result of work carried out for this thesis. These changes, in particular the reduction of imidazole concentration during the process of protein concentration, have subsequently made the system more amenable for use in biochemical studies. This chapter concludes with a system that enables the expression and purification of near full-length, truncated and mutated variants of the flax rust R proteins, M and L6, which facilitated the *in vitro* biochemical studies described in Chapters 4 and 5 of this thesis.

## 3.2 Results

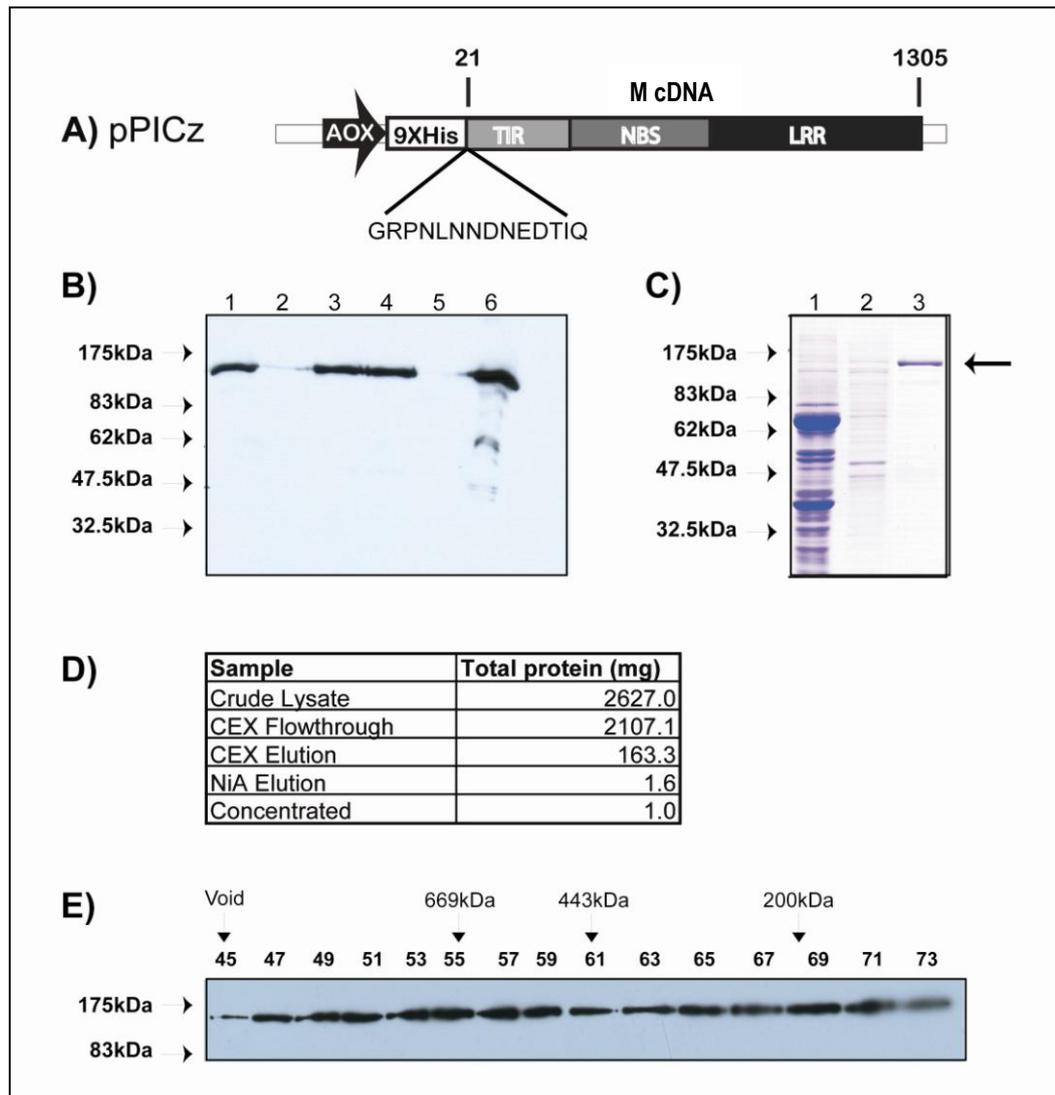
### 3.2.1 Three-step purification of M and L6

For the heterologous expression of M and L6 their cDNAs (minus the regions that encode the first 21 and 28 amino acids, respectively) were cloned into the pPICz expression vector (*P. pastoris* expression manual, Invitrogen) and stably integrated into the genome of *P. pastoris* using electroporation (see 2.3.4). For M, all cloning and transformation steps were performed by Dr. Simon Schmidt; for L6, the DNA construct was prepared by Dr. Peter Dodds and I performed the transformation into *P. pastoris*. In this system expression of the flax R proteins is under the control of the alcohol oxidase (AOX) promoter and protein induction is achieved with the addition of methanol during culturing (see 2.4). A 9x histidine tag, engineered at the N-terminus, facilitated metal affinity purification and also western detection using an anti-His antibody (Figure 3.1.A and Figure 3.2.A). An anti-M antibody, raised against an N-terminal epitope of M, was also used in western blot experiments for the detection of M protein (Figure 3.1.A). Positive transformants were screened (see 2.3.5) by western blot analysis and those that expressed protein at the highest levels were chosen for large scale expression and purification studies.

A three-step purification method, involving cation exchange (CEX), nickel affinity (NiA) and gel filtration (GF) chromatography, was used to purify near full-length M and L6 proteins. In order to be concise, in remaining chapters, near full-length protein will be referred to as M or L6 and truncations or mutations will be designated with a prefix or suffix. For example, the removal of the TIR domain, or mutation of the conserved lysine in the P-loop motif in M, will be indicated as M $\Delta$ TIR and M<sup>K286L</sup>, respectively.

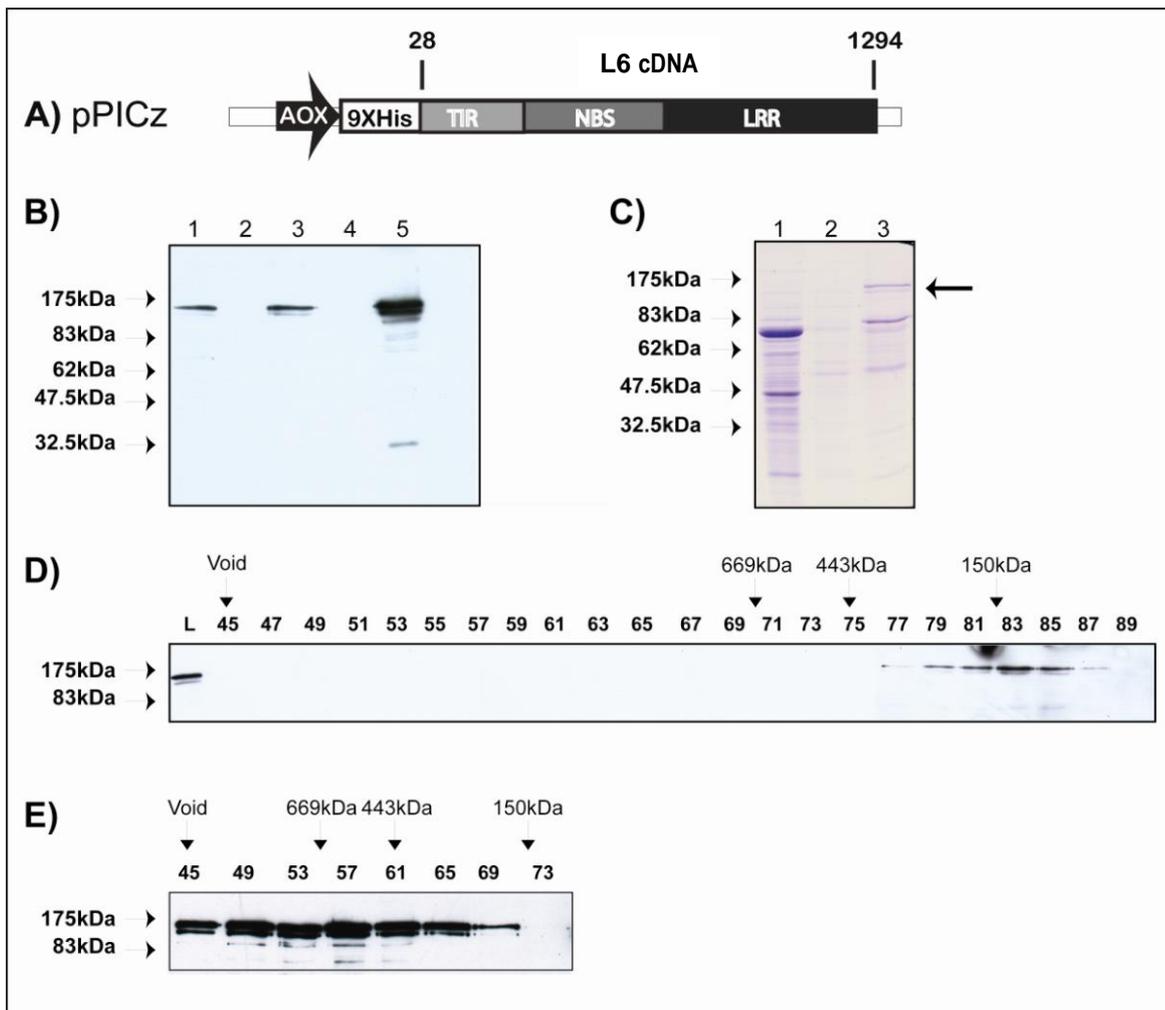
Western blot analysis demonstrates good recoveries of both M (147 kDa) and L6 (145 kDa) protein from CEX and NiA chromatography (Figure 3.1.B/3.2.B). At the time of these experiments, and even now, there is no clear activity assay for flax R proteins. This makes estimates of protein purity extremely difficult. For this reason it can only be concluded from visual inspection of Coomassie stained gels, that M and L6 protein purified by this three-step purification procedure are approaching homogeneity (Figure 3.1.C/3.2.C). Lysis and purification from a 500 g wet cell weight (wcw) pellet, obtained from a 4L fermentation culture (see 2.4.2), yields approximately 1 mg of M protein (Figure 3.1.D) and a similar quantity can be obtained for L6 (data not shown). GF analysis of M protein after elution from CEX demonstrated that the M protein was monomeric at this stage of the purification (Schmidt et al., 2007b), however, the NiA purified and concentrated M protein had a much broader profile, indicating that the protein is likely to be self-associating when highly enriched (Figure 3.1.E). The same phenomenon was observed for L6. Analysis of *P. pastoris* crude lysates containing the L6 proteins, demonstrated

that the L6 protein had an elution point when separated over GF that corresponded to an estimated molecular weight of approximately 100 kDa (Figure 3.2D) (estimated from a standard curve generated from known proteins (Appendix 2)). While this is lower than its theoretical molecular weight of 145 kDa, the standard curve is generated from globular proteins and does not account for protein shape or any interaction the protein may have with the resin. This elution profile therefore suggests that at lysis the L6 protein is monomeric, however, following purification and concentration L6 appears to be self-associating and has a broad GF profile (Figure 3.2.E).



**Figure 3.1: Expression and purification of recombinant M**

A) A schematic of the cloning vector (pPICz) containing the M cDNA (minus the first 21 amino acids) with an N-terminal 9x histidine tag and the N-terminal anti-M epitope highlighted. Expression in *P. pastoris* is under the control of the alcohol oxidase (AOX) promoter and the domains of the protein are designated TIR, NBS and LRR as defined in the text. B) Western blot analysis using an anti-M antibody shows the presence or absence of the M protein during purification. Lane 1, crude lysate; lane 2, CEX flowthrough; lane 3, CEX pH 9 elution; lane 4, CEX pH 10.5 elution (NiA pre-column sample was pH 9 and pH 10.5 CEX elution sample); lane 5, NiA flowthrough; lane 6, NiA elution. C) Coomassie stained gel of protein separated by SDS-PAGE showing enrichment of M protein during purification. Lane 1, crude lysate; lane 2, NiA pre-column sample; lane 3, NiA elution, arrow indicates the M protein. D) Calculation of total protein during each step of M purification. E) Following NiA chromatography, concentrated protein was further separated over a Superdex 200 PG GF column. The 2 ml fractions were analysed by western blot with an anti-M antibody. Molecular weights above the gel are positioned where proteins of known molecular weight elute from the same column.



**Figure 3.2: Expression and purification of recombinant L6**

A) A schematic of the cloning vector (pPICz) containing the L6 cDNA (minus the first 28 amino acids) with an N-terminal 9x histidine tag. Expression in *P. pastoris* is under the control of the alcohol oxidase promoter and the domains of the protein are designated TIR, NBS and LRR as defined in the text. B) Western blot analysis using an anti-His antibody shows the presence or absence of the L6 protein. Lane 1, crude lysate; lane 2, CEX flowthrough; lane 3, CEX pH 10.5 elution (nickel affinity pre-column sample); lane 4, NiA flowthrough; lane 5, NiA elution. C) Coomassie stained gel of protein separated by SDS-PAGE showing enrichment of L6 protein during purification. Lane 1, crude lysate; lane 2, NiA pre-column sample; lane 3, NiA concentrated elution, arrow indicates the L6 protein. A lower molecular weight polypeptide of 75kDa is likely to represent a co-purifying contaminant protein from the expression host. D) Crude lysate containing L6 was separated over a Sephacyl 400 GF column. The 2 ml fractions were analysed by western blot with an anti-His antibody. Molecular weights above the gels are positioned where proteins of known molecular weight elute from the same column. E) Following NiA chromatography concentrated protein was further separated over a Superdex 200 PG GF column. The 2 ml fractions were analysed by western blot with an anti-His antibody.

### 3.2.2 Small scale expression and one-step purification

A major detraction of the purification method outlined above was the large scale required, and the length of time needed, for the return of a relatively small amount of purified protein. A typical expression required one week for fermentation, and the purification could span over 16-24 hours. Large media and buffer volumes were required for both expression and purification and there was a long equipment usage time. Furthermore the purification was both expensive and time consuming, which did not lend itself well to *in vitro* analysis of the flax R proteins, where simultaneous expression and purification of proteins would be desirable. In an effort to reduce time and input, a smaller scale expression system and a one-step purification technique using NiA was developed (Figure 3.3). Previous attempts at purifying M protein using NiA as the first step had failed (Schmidt et al., 2007b), however, access to new generation fast flow affinity resins now made this a possibility (Figure 3.3). Rather than using 500 g wcv pellets from fermentor grown cultures, 3 or 6 g wcv pellets from shaker flask culture of 50 to 100 mls were used, substantially reducing the starting volumes.

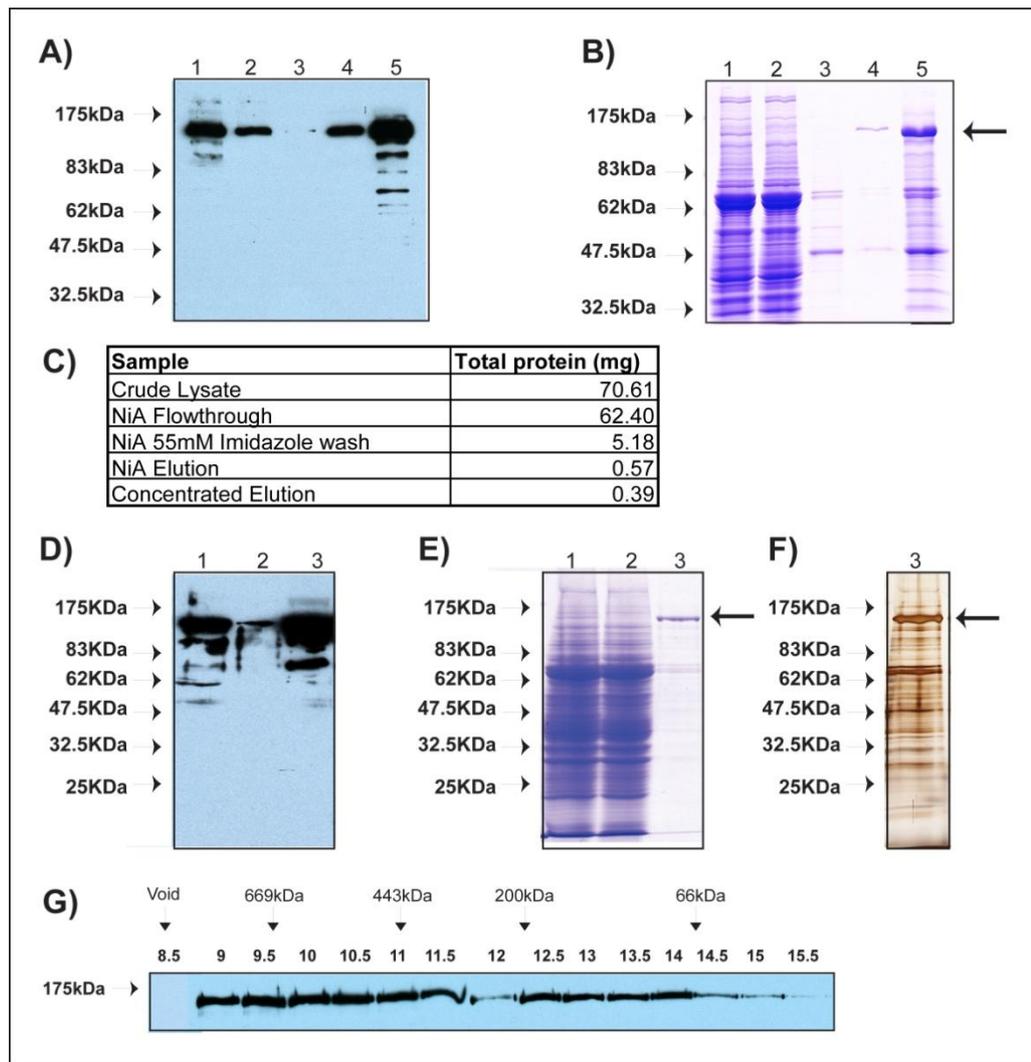
With these adjustments to the expression system, M protein could be purified directly from *P. pastoris* cleared cell lysates using NiA chromatography. Utilising a 10 column volume (c/v) imidazole wash, purities estimated between 50-70% (from visual inspection of Coomassie stained SDS-PAGE gels) could be achieved with a yield of 20-40 $\mu$ g/g (M protein/wcv pellet) (Figure 3.3.A-C). A 100c/v wash strategy increased the purity by removing more of the 65kDa and 48kDa major contaminants, however, a yield penalty was incurred, dropping the total yield to 5-10 $\mu$ g/g (M protein/wcv pellet) (Figure 3.3D-E). The 100c/v strategy did not remove all contaminant proteins as shown by an overexposed silver stained gel of concentrated protein separated by SDS-PAGE. This shows both contaminants are still present (Figure 3.3.F). GF profiles of the NiA purified M protein (Figure 3.3.G) resembled the profile that was observed from the three-step purification strategy (Figure 3.1.E). At this point, due to a broad elution profile of purified M protein, it was considered that the use of GF could not increase purity (see 2.8.3).

While the one-step method produced a less pure form of protein, it significantly reduced the manipulation of the protein and the different buffers that the protein was exposed to during purification. It was anticipated that this shorter protocol would be more suited to biochemical studies. It was also considered that any biochemical study conducted on less than 100% pure protein could be normalised with an empty vector control. Also the use of an in-gel method to determine protein concentration would enable reliable estimations of recombinant R protein concentrations.

Common to both the small and large scale methods of purification was the self-association phenomenon that occurred with the more enriched forms of both M and L6 proteins.

This was of interest, as elicitor-mediated oligomerisation of the tobacco R protein, N, has been reported (Mestre and Baulcombe, 2006). In addition numerous studies of STAND proteins implicate oligomerisation as a phenomenon associated with the activated state of the protein. The literature also suggests that nucleotide binding, most likely ATP, is required for oligomerisation.

Although the self-association phenomenon associated with the enriched and concentrated M and L6 protein was likely to be a physical one and potentially associated with protein concentration, it was decided to further investigate the formation of the higher order structure of the R proteins and determine if this was related to nucleotide binding. Numerous GF experiments of purified M protein were performed in the presence of ATP and  $Mg^{2+}$ , or in the presence of EDTA, which would theoretically strip the  $Mg^{2+}$  coordinating ion preventing nucleotide binding. However, these buffer changes had little effect on the size of the protein over GF (data not shown). The oligomerisation studies on tobacco N had implicated the TIR in coordinating the formation of the oligomer (Mestre and Baulcombe, 2006). This is not surprising given the role of the TIR domain in mediating homo-typic interactions in TLR proteins (see 1.6.3). To determine if the TIR was causing self-association of the enriched M and L6 proteins, TIR truncated forms of M and L6 proteins were generated (for M see 2.3.1.1, the L6 construct was generated by Dr. Peter Dodds) and transformed into *P. pastoris* for analysis. The truncations were designed at the exon/intron boarder of M and L6, however, later alignments revealed that these truncations are likely to still contain a small C-terminal element of the TIR domain (Appendix 3). Regardless, the results of expression and purification of  $M\Delta TIR$  are shown here. These results provided important clues about phenomena that effect protein size and thus were critical in subsequent biochemical assays performed on the proteins.



**Figure 3.3: Small Scale expression and purification of M for biochemical analysis**

A) Western blot analysis using an anti-M antibody showing the presence and absence of M protein during purification. Lane 1, crude lysate; lane 2, NiA flowthrough; lane 3, NiA 55mM imidazole wash (10c/v); lane 4, NiA elution; lane 5, concentrated elution. B) Coomassie stained gel of protein separated by SDS-PAGE with the same loading arrangement as in the western blot analysis, arrow indicates M protein. C) Calculation of total protein during M purification. A 100c/v imidazole wash strategy was also used during NiA chromatography. Subsequently panels D, E and F show the results of this purification protocol. D) Western blot analysis using an anti-M antibody showing the presence and absence of M protein during purification. Lane 1, crude lysate; lane 2, NiA flowthrough; lane 3, concentrated elution. E) Coomassie stained gel of protein separated by SDS-PAGE with the same loading arrangement as in the western blot analysis, arrow indicates M protein. F) Silver stained analysis of the concentrated sample (E, lane 3) separated by SDS-PAGE, arrow indicates M protein. G) Following NiA chromatography, concentrated protein was further separated over a Superdex 200 HR GF column. The 500  $\mu$ l fractions were analysed by western blot with an anti-M antibody. Molecular weights above the gel are positioned where proteins of known molecular weight elute from the same column.

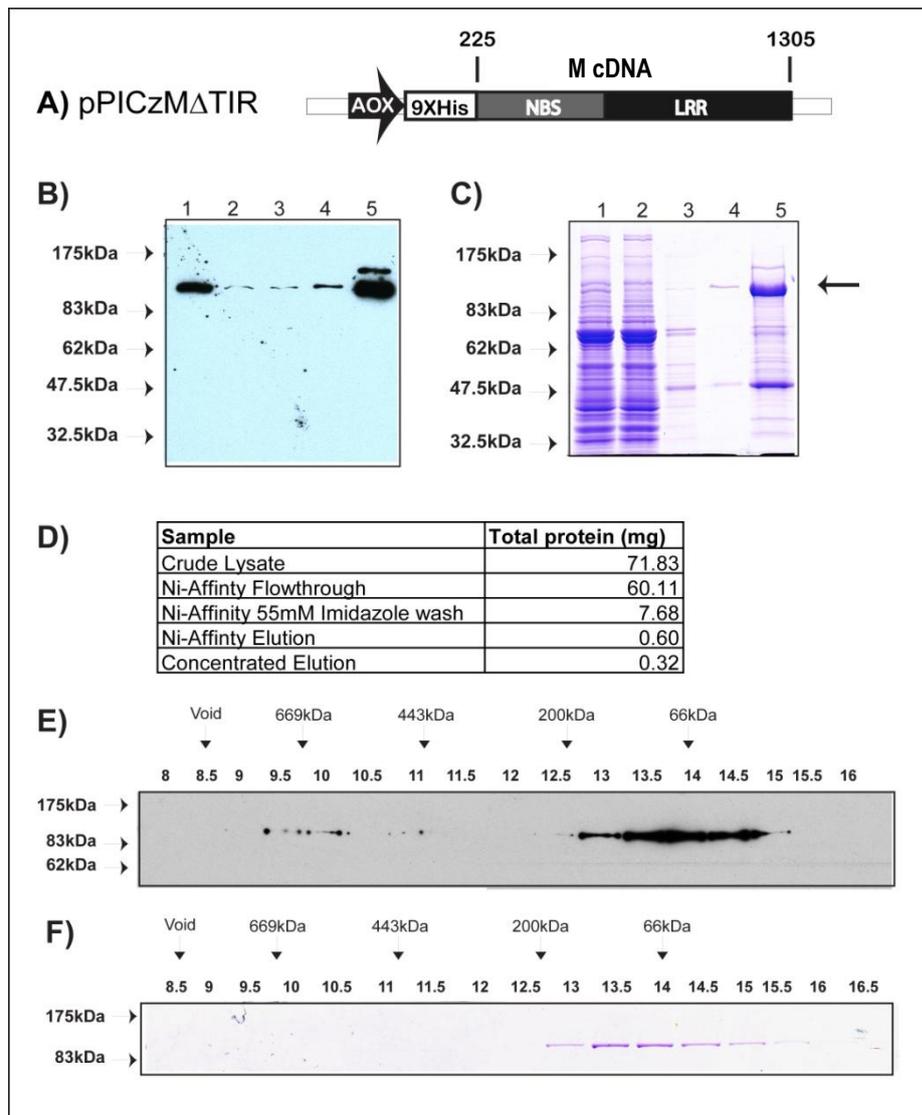
### 3.2.3 Potential issues in the functionality of the purified proteins

The M $\Delta$ TIR protein could be expressed and purified from *P. pastoris* whole cell lysates in the same way as the M protein (Figure 3.4A-D). Following NiA chromatography, purified M $\Delta$ TIR (124 kDa) was separated by GF and found to elute at an estimated molecular weight of approximately 110 kDa (Figure 3.4E/F). It was therefore concluded that this represented the monomeric form of the M $\Delta$ TIR protein. This result suggested that the cause of the spread GF profile of the concentrated full-length M protein was self-association that was likely to be coordinated by the TIR domain (Figure 3.4E-F). Repeated purification of M $\Delta$ TIR, however, demonstrated that the native size of this protein, assayed by GF, varied between preparations. Much of this work is not included in this thesis, however in brief, after repeated purification and GF analysis it was found that a combination of protein concentration in a 250 mM imidazole buffer and length of exposure to 250 mM imidazole caused the M $\Delta$ TIR protein to self-associate. In some cases, prolonged exposure (24 hours at 4°C) of concentrated protein to 250 mM imidazole caused the M $\Delta$ TIR protein to precipitate.

Through mainly trial and error, it was found that to maintain M $\Delta$ TIR protein as a monomer, the NiA elution needed to be diluted immediately in GF buffer to ensure that the concentration of imidazole did not exceed 75 mM (Figure 3.5). This proved to be an effective and repeatable method in maintaining monomeric protein. Importantly, when M $\Delta$ TIR was buffer exchanged into the GF buffer it could be highly concentrated (up to ~8 mg/ml) without affecting its mono-dispersed nature. This demonstrated that the imidazole effect was the underlying issue in most cases of self-association and in fact the protein was forming soluble aggregates that were presumably non-functional and non-biological.

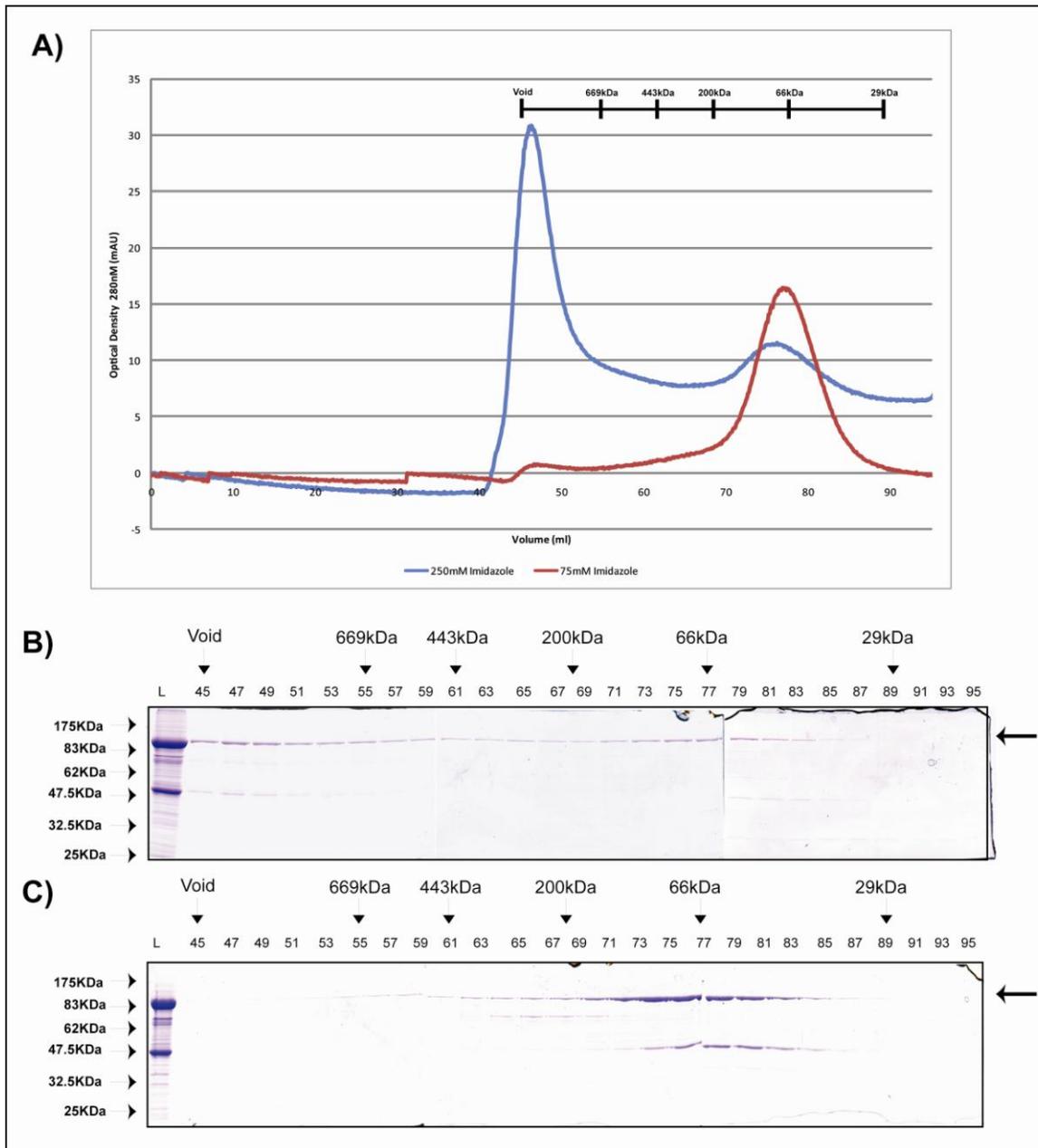
To determine if imidazole was having the same effect on the M and L6 proteins, the same dilution step was added prior to concentration and GF. M, L6 and L6 $\Delta$ TIR protein purified with these amendments eluted as a monomer on GF (Figure 3.6), indicating that the cause of the broad elution profile for these proteins on GF was also an effect of imidazole concentration. While some higher molecular weight forms of the full-length proteins could still be observed, the majority was monomeric, with a slight skew towards a higher molecular weight form.

The aggregation effect of imidazole on the purified R proteins was a significant finding. The aim of this thesis was to perform functional analyses on the purified proteins, and therefore any potential adverse effects to protein structure would more than likely compromise its ability to function. All future purifications therefore involved a dilution of the total imidazole concentration following NiA elution and prior to protein concentration.



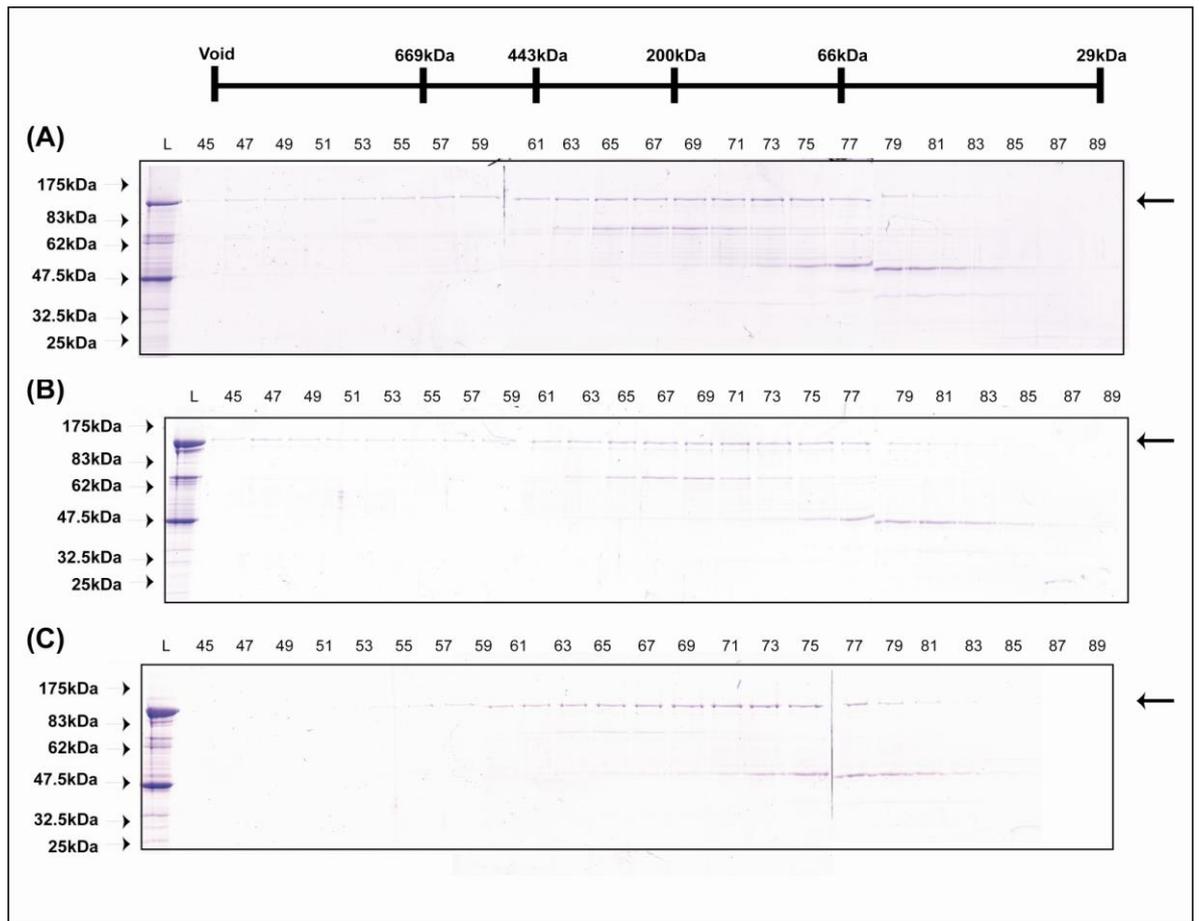
**Figure 3.4: Purification of M $\Delta$ TIR and gel filtration analysis**

A) A schematic of the cloning vector (pPICz) containing the M cDNA (minus the first 225 amino acids) with an N-terminal 9x histidine tag. Expression in *P. pastoris* is under the control of the alcohol oxidase promoter and the domains of the protein are designated NBS and LRR as defined in the text. B) Western blot analysis using an anti-His antibody showing the presence or absence of the M $\Delta$ TIR protein. Lane 1, crude lysate; lane 2, NiA flowthrough; lane 3, NiA 55 mM imidazole wash; lane 4, NiA elution; lane 5, concentrated elution. C) Coomassie stained gel with the same loading arrangement as in the western blot analysis, arrow indicates M $\Delta$ TIR protein. D) Calculation of total protein during each step of M $\Delta$ TIR purification. Following NiA chromatography, concentrated protein was further separated over a Superdex 200 HR GF column. The 500 $\mu$ l fractions were analysed by western blot with an anti-His antibody (E) or protein separated by SDS-PAGE were Coomassie stained (F). Molecular weights above the gel are positioned where proteins of known molecular weight elute from the same column.



**Figure 3.5: Imidazole affects M $\Delta$ TIR monomeric state**

Following NiA purification, M $\Delta$ TIR protein was concentrated directly from elution buffer (containing 250 mM imidazole) or diluted 3.5 fold into GF buffer (diluted to below 75 mM imidazole). A) A comparison of the resulting 280nm UV trace profile of the different protein preparations when separated over the S200PG GF column (Blue trace line represents the profile of M $\Delta$ TIR protein concentrated in 250 mM imidazole; red trace line represents the profile of M $\Delta$ TIR protein concentrated in 75 mM imidazole). The resulting 2 ml fractions of protein concentrated in 250 mM imidazole (B) and 75 mM imidazole (C) were analysed by SDS-PAGE and Coomassie stained. Molecular weights above gels are positioned where proteins of known molecular weight elute from that same column.



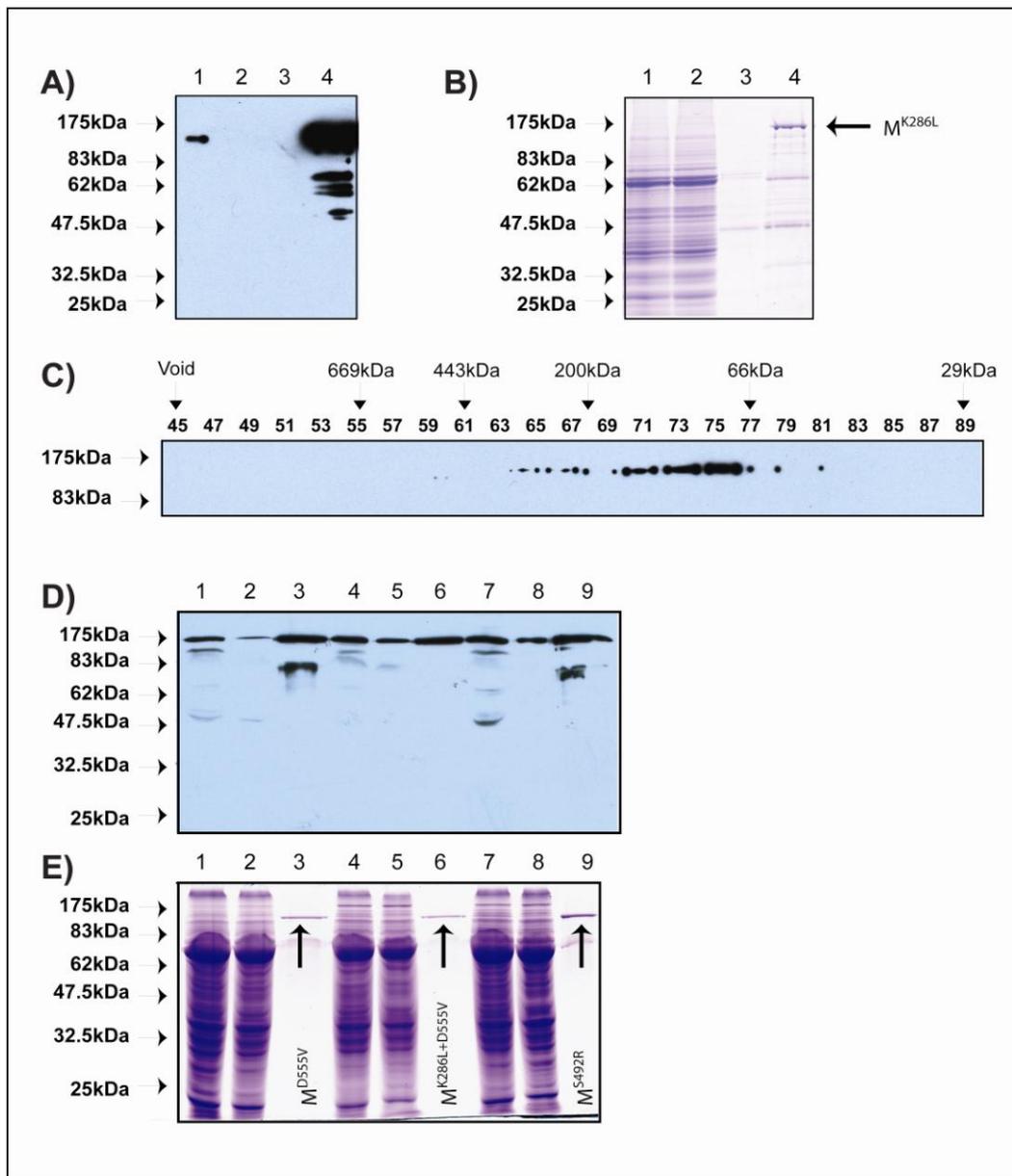
**Figure 3.6: Gel filtration comparison of M, L6, and L6 $\Delta$ TIR**

NiA purified M, L6 and L6 $\Delta$ TIR proteins were concentrated in buffer containing 75 mM imidazole and separated over the S200PG GF column. The 2 ml fractions were separated by SDS-PAGE and Coomassie stained. Analysis of M (A), L6 (B) and L6 $\Delta$ TIR (C) protein, arrows indicate the respective R proteins. Molecular weights at the top of gels are positioned where proteins of known molecular weight elute from the same column.

### 3.2.4 Full-length M mutants can be purified from *P. pastoris*

For the functional analysis of the NB-ARC domain, a series of mutations were introduced by site-directed mutagenesis (see 2.3.2) into the M cDNA at key nucleotides that encode conserved residues predicted to affect function. These mutations were predicted to affect R protein function, based on known *in planta* phenotypes in other R proteins (see Table 4.1 for explanation). Using the established purification protocol, mutants of M could be purified in the same way as M (Figure 3.7). GF analysis of the M<sup>K286L</sup> mutant protein demonstrated that the majority of the purified protein was monomeric, similar to the profile of M. The ability to purify these proteins would facilitate future biochemical studies, and this is covered in Chapters 4 and 5.

The most comprehensive biochemical analysis of an R protein has come from studies of the tomato R proteins I-2 and Mi-1 (Tameling *et al.*, 2002, Tameling *et al.*, 2006). Both studies were performed on proteins truncated for the LRR. In an effort to replicate some of this work a TIR-NB-ARC protein was generated. The truncation of the LRR was made at the site predicted to be the start of the first repeat of the LRR domain (see 2.3.1.2) (Appendix 3). This protein was designated M $\Delta$ LRR. Unfortunately, M $\Delta$ LRR could not be purified in the same way as M or M $\Delta$ TIR. There was poor binding of M $\Delta$ LRR to the NiA column and only small quantities of the protein could be obtained (Appendix 4). At this stage we are unsure why this protein could not be purified, however, no further attempts to purify this protein were pursued.



**Figure 3.7: Purified M full-length mutants**

A) Western blot analysis using an anti-M antibody showing the presence or absence of the  $M^{K286L}$  protein. Lane 1, crude lysate; lane 2, NiA flowthrough; lane 3, NiA 55 mM imidazole wash (10c/v); lane 4, concentrated NiA elution. B) Coomassie stained gel of protein separated by SDS-PAGE, with the same loading arrangement as in the western blot analysis. C) Concentrated protein was further separated over a Superdex 200 PG GF column and the 2 ml fractions were separated by SDS-PAGE and analysed by western blot with an anti-M antibody. Molecular weights above gels are positioned where proteins of known molecular weight elute from the same column. D) Western blot analysis using anti-M antibody showing the presence or absence of mutated M proteins. Lane 1-3, purification of  $M^{D555V}$ ; Lane 4-6, purification of  $M^{D555V+K286L}$ ; Lane 7-9, purification of  $M^{S492R}$ . Lane 1, 4 and 7, crude lysate; lane 2, 5 and 8, NiA flowthrough; lane 3, 6, and 9, concentrated NiA elution. E) Coomassie stained gel of proteins separated by SDS-PAGE with the same loading arrangement as in the western blot analysis, arrows indicate respective NiA purified and concentrated M proteins.

### 3.3 Discussion

The results outlined in Chapter 3 demonstrate that soluble recombinant flax rust R proteins, M and L6, can be produced in, and purified from, the *P. pastoris* expression system (Schmidt et al., 2007b). This technique was the first published report of an expression and purification method that facilitated the production of tri-domain, near full-length, recombinant R protein and has thus opened the way for biochemical studies of the M and L6 proteins.

#### 3.3.1 Expression and purification of flax R proteins facilitates *in vitro* studies

Purification of proteins can be achieved using a combination of chromatography techniques including CEX, NiA and GF, however, for small-scale expression and purification, a one-step NiA technique is possible. The large-scale expression and three-step purification strategy (as described by Schmidt et al., (2007b) had a number of features that made it less amenable to protein production for biochemical studies of the purified proteins. Many of these problems were alleviated using a small-scale expression and one-step purification strategy. One draw back, using this strategy, was that the purity of the recombinant R protein was reduced with the removal of the CEX chromatography step. The one-step purification method, however, had the advantage of decreased scale, time and cost over the previous published three-step method. Using the small-scale method, a washing strategy of 100 c/v gave higher purity, 80-90%, as estimated from Coomassie staining of SDS-PAGE gels. GF as a second step did not substantially improve purity, as major contaminants (65 kDa and 48 kDa) elute from GF in fractions that overlap with the fractions containing the recombinant proteins. The major contaminants of 65 kDa and 48 kDa are likely a result of an ionic association between *P. pastoris* proteins and the NiA column. Both contaminants can be removed from the column with prolonged washing, however, they are still detected by silver stain analysis after a 100c/v washing step. The use of higher ionic strength NiA buffers to prevent the binding of these proteins was not considered. Schmidt et al., (2007b) demonstrated that buffers of high ionic strength (>300 mM) cause aggregation of M and L6 protein. Despite the less than pure R protein obtained by this one-step procedure it was anticipated that a purification of protein from *P. pastoris* cells transformed with an empty pPICz vector would act as a suitable background control for subsequent biochemical studies.

This modified purification strategy facilitated the purification of full-length, TIR truncated M and L6 proteins, and mutated variants of M. The M $\Delta$ TIR protein was routinely purified with greater yields and stability in comparison to the M protein. Conversely, the yield of the M $\Delta$ LRR protein from the NiA was extremely low due to poor binding. The exact reason for this behaviour during purification is unknown and consequently the M $\Delta$ LRR protein is excluded from any *in vitro* analysis outlined in Chapters 4 and 5.

### 3.3.2 Physical properties of the flax R proteins cause self-association and aggregation

Early studies carried out on the M protein had demonstrated that it was prone to severe aggregation as a result of exposure to high salt concentrations (ionic strength > 300 mM) (Schmidt et al., 2007b). Studies had also demonstrated that after lysis and elution from cation exchange, the M protein was monomeric. This was also observed from analysis of L6 protein in whole cell crude lysates (Figure 3.2D). However, the enriched NiA purified forms of both M and L6 proteins, when loaded onto GF, eluted over a broad range indicative of soluble protein aggregates. It was predicted that protein aggregation was most likely the result of concentration-dependent physical interactions between hydrophobic regions of the proteins, although some speculation regarding the role of a bound nucleotide in the NB-ARC domain was also suggested as a possible cause of this phenomenon (Schmidt et al., 2007b).

During the course of this study a number of potential causes of protein aggregation and instability in enriched samples of M and L6 were investigated. Earlier studies had demonstrated the importance of glycerol in maintaining the native state of the purified R proteins (Schmidt, 2002). Glycerol is a protective osmolyte, a small organic compound known to protect the native state of a protein. Glycerol, therefore, is often added to purification buffers to help protect against protein denaturation (Meng et al., 2004, Street et al., 2006). Schmidt et al., (2007b) demonstrated that highly degraded forms of M protein were observed in crude lysates when glycerol was left out of the cell lysis buffer. Furthermore, on addition of glycerol in GF buffers the protein is protected from interaction with the GF column stationary phases (data not shown). However, even with the addition of glycerol as a protective factor, protein instability and aggregation in the enriched forms of M and L6 is still observed. After a process of elimination, it was found that protein concentration in high imidazole concentrations in the NiA elution buffer (250 mM), caused varying amounts of protein aggregation. The exact effect that imidazole is exerting on the protein is unknown, and literature searches found little information regarding this type of phenomenon. It is not simply an ionic effect, as the total ionic strength of the NiA elution buffers are slightly less than the ionic strength of the GF buffer (now the preferred concentration buffer). For this reason, NiA elutions were immediately diluted into GF buffer such that the final imidazole concentration did not exceed 75 mM.

Full-length M appears to be slightly more prone to forming larger protein complexes when compared to M $\Delta$ TIR. It is tempting to speculate that this is a feature of the TIR domain, given its suspected role in coordinating the oligomerisation of the N protein (Mestre and Baulcombe, 2006). Indeed GF analysis of the TIR domain of L6 expressed and purified as a separate domain has shown that it elutes as both a monomer and a dimer (T. Ve, personal communication), thus

suggesting a potential role of this domain in coordinating protein self-association. Research with this idea in mind is ongoing in the Anderson laboratory and our collaborating laboratory led by Professor Bostjan Kobe.

### **3.3.3 Concluding remarks**

During the course of this study, continued development and refinement of the purification procedure introduced by Schmidt et al., (2007b) have been undertaken. A purification strategy was finally arrived at that enabled the production of predominantly monomeric M and L6 proteins (see 2.6.3). During the refinement of this purification strategy a number of assays designed to test the potential function of purified proteins were concurrently performed. These assays included, ATP hydrolysis and ATP/ADP quantification analysis, which in related R proteins and STAND proteins had been used to demonstrate function (Marquenet and Richet, 2007, Tameling et al., 2002, Ueda et al., 2006). The results of much of this concurrent work have not been included in this thesis because the experiments were performed on proteins purified and concentrated in 250 mM imidazole. Despite some initial results that were consistent with the mutations that were introduced in the M protein, no repeatable data was obtained. These results add little to the overall message of this thesis and have therefore been excluded. It is my belief that the most likely cause of this inconsistency in biochemical analysis is that the protein function was compromised by imidazole-induced aggregation. Although this idea needs to be investigated further, the biochemical studies presented in Chapters 4 and 5 were performed on protein purified according to the methods described in section 2.6.3.

# **Chapter 4: Biochemical Investigation of the NB-ARC Domain of Purified Flax Rust R Proteins**

## 4.1 Introduction

Armed with an expression and purification strategy for the flax R proteins, M and L6, this study now re-focussed towards assessing the biochemical properties of the R proteins. As highlighted throughout this thesis, only a small proportion of biochemical studies have been published in the plant disease resistance field that add to our overall understanding of the function of an R protein. It is these reports, and the observed gaps in our overall understanding, that have inspired this work. The particular focus of this study, was to use recombinant protein, in *in vitro* assays, to decipher what role the NB-ARC domain plays in R protein activation and regulation. To do this we used mutant forms of the M protein predicted to affect activation of the resistance response.

A host of mutations in the NB-ARC domain of R proteins have been demonstrated to cause either loss- or gain-of-function *in planta* (reviewed by (Takken et al., 2006)). This mutational analysis alone has long implicated the NB-ARC domain in the activation and associated regulation of an R protein (see 1.6.2/1.8 and Figure 1.3). To investigate this further, three mutations within the NB-ARC domain of M were generated with predicted loss- and gain-of-function consequences (Table 4.1). To justify the selection of these mutations, an explanation for each mutation is detailed below.

In the NB subdomain, the invariant lysine within the P-loop was mutated to a leucine. A mutation of this equivalent lysine to a leucine in the R protein RPS2 prevented its function in *A. thaliana* (Tao et al., 2000). Furthermore, a mutation of the same lysine to a methionine in the flax rust R protein L6, also prevented its *in planta* function (Howles et al., 2005). It was therefore anticipated that a mutation of the P-loop lysine to a leucine at residue 286 in M, would inactivate the protein. In other R proteins this equivalent lysine residue has previously been investigated for its impact on the biochemical properties of the protein. *In vitro* studies of I-2, Mi-1, and N showed this lysine is crucial for the coordination of ATP binding (Tameling et al., 2002, Ueda et al., 2006). Indeed in the crystal structures of Apaf-1 and CED-4, the P-loop lysine coordinates hydrogen bonding with the bound nucleotides ADP and ATP, respectively (Riedl et al., 2005, Yan et al., 2005). It was therefore anticipated that this mutation in M would prevent any associated nucleotide binding and/or hydrolysis activity.

The remaining two mutations were putative autoactive (gain-of-function) mutations within the ARC2 subdomain (Table 4.1). The highly conserved aspartate within the MHD motif, D555, was mutated to a valine, while the serine within the motif VIII, S492 was mutated to an arginine. An aspartate to valine mutation within the MHD motif, has been demonstrated to causes an autoactive phenotype in a number of R proteins including, Rx, I-2, L6, Mi-1 and Rpi-blb1 (Bendahmane et al., 2002, de la Fuente van Bentem et al., 2005, Howles et al., 2005, van Ooijen

et al., 2008b). However, an autoactive phenotype is not always associated with this mutation in an R protein. For example, the MHV mutation when introduced into the *A. thaliana* R protein RPS4 was instead shown to cause loss-of-function (Jane Parker, conference presentation, Plant Innate Immunity, Keystone conference, 2008). In the NB-ARC-LRR class of R proteins, the histidine within this motif is invariant, and the aspartate is highly conserved. Manipulation of the histidine in the MHD motif of I-2, demonstrated that 15 out of 19 different amino acids substitutions resulted in an autoactive phenotype, highlighting the regulatory power that this motif has over the protein's ability to induce a HR (van Ooijen et al., 2008b). Multiple-sequence alignments of R proteins with Apaf-1 reveal an equivalent LHD motif in Apaf-1 that aligns with the MHD motif of R proteins (Figure 1.3). This motif, however, is absent in CED-4. In the Apaf-1 structure, the histidine of the LHD motif forms a hydrogen bond with the  $\beta$ -phosphate of the bound ADP (Riedl *et al.*, 2005). In both Rx and L6, mutation of the conserved lysine within the P-loop motif in an MHV background, has been shown to override the autoactive phenotype and cause inactivity (Bendahmane et al., 2002, Howles et al., 2005). This suggests that nucleotide binding is required for an MHV mutated protein to be autoactive. For this reason a P-loop/MHV double mutant was generated in M and included in the presented *in vitro* study. While mutations at the MHD motif that cause autoactivity have been well defined, the molecular consequences on the protein that lead to autoactivity are yet to be determined.

The concept of the putative autoactive mutation in the motif VIII (Table 1.2/Figure 1.3) is less well documented. The idea to mutate this motif originated from a study of the *A. thaliana* protein, SSi4, whereby a glycine to arginine mutation, at the equivalent position of a serine in M, resulted in autoactivity (Shirano et al., 2002). In the NB-ARC-LRR R proteins, serine is highly conserved at this position. In a multiple-sequence alignment with other R proteins, the motif aligns only loosely to Apaf-1 and CED-4 (Figure 1.3). A serine in Apaf-1 (S371) and a cysteine in CED-4 (C381) align to the serine of the R proteins, or glycine in the case of SSi4 (Figure 1.3) (van Ooijen et al., 2008b). In the crystal structure of Apaf-1 and CED-4 these residues are not involved in coordinating the binding of the nucleotide in either structure (Riedl et al., 2005, Yan et al., 2005), however, C381 of CED-4 has been proposed to form part of the protein's active site (Takken et al., 2006). The motif VIII is far less well defined than many other motifs in the NB-ARC domain of R proteins, with no further functional or biochemical data since that reported by Shirano et al., (2002) on SSi4.

It was anticipated that an *in vitro* study of mutants predicted to cause loss- and gain-of-function phenotypes would further expose the role that nucleotide binding and hydrolysis plays in R protein activation. The mutated proteins were expressed and purified using the same protocol developed for M (Chapter 3). This chapter describes a biochemical study of these mutated M

proteins. The *in planta* analysis of these mutants at the time of this thesis submission was ongoing (E. deCourcy-Ireland, personal communication). The *in planta* phenotypes generated by *Agrobacterium* infiltration of constructs carrying the mutations therefore have not been confirmed. However, as described above, in the closely related flax R protein L6, three mutations equivalent to those made to M, have been functionally defined *in planta*. As M shares 78% overall amino acid identity with L6 (Anderson et al., 1997) and 82% identity (89% similarity) in the NB-ARC region we predicted that the equivalent mutations in M will cause the same functional phenotypes as observed in L6. As, however, the phenotypes in M are yet to be defined they are consequently described in the text as having putative phenotypes.

| M mutants                | Motif affected | Domain  | Predicted <i>in planta</i> function | Reference  | Predicted <i>in vitro</i> analysis       | Reference                                  |
|--------------------------|----------------|---------|-------------------------------------|--|--|--|
| M <sup>K286L</sup>       | P-loop         | NB      | Loss-of -function                   | (Tao et al., 2000)   | Unable to bind ATP                       | (Tameling et al., 2006, Ueda et al., 2006) |
| M <sup>S492R</sup>       | Motif VIII     | ARC2    | Autoactive (gain-of-function)       | (Shirano et al., 2002)   | Unknown                                  | n/a  |
| M <sup>D555V</sup>       | MHD            | ARC2    | Autoactive (gain-of-function)       | (Bendahmane et al., 2002, de la Fuente van Bentem et al., 2005, Howles et al., 2005, van Ooijen et al., 2008b) | Unknown                                  | n/a  |
| M <sup>K286L+D555V</sup> | P-loop /MHD    | NB/ARC2 | Loss-of -function                   | (Howles et al., 2005, van Ooijen et al., 2008b)  | Unknown                                  | n/a  |
| M $\Delta$ TIR           | n/a            | TIR     | n/a                                 | n/a  | NB-ARC-LRR of N binds and hydrolysis ATP | Ueda et al, 2006                           |

**Table 4.1: The M-mutants analysed *in-vitro***

Predictions are made based on previously published functional analyses of other R proteins.

## 4.2 Results

### 4.2.1 M and L6 proteins are purified with ADP bound

To determine if the M and L6 proteins purified from the *P. pastoris* expression system are bound with endogenously derived adenine nucleotides, a luciferase-based ATP/ADP identification and quantification assay was employed (see 2.8.2). To release any bound nucleotides from the proteins, an acid-based nucleotide extraction method, and a boiling precipitation method, were trialled (I. Menz, personal communication). It was clear that the acid extraction method greatly underestimated (5.5 fold) the amount of nucleotide present, when compared to the boiling precipitation method (Appendix 5). The boiling method was shown by Bradford assay (see 2.2.9) to precipitate all of the concentrated protein and was therefore used for all subsequent experiments (data not shown). The ATP/ADP quantification assay was performed as described for the nucleotide quantification analysis of the STAND protein, MalT (Marquetet and Richet, 2007)

Early studies were performed on re-concentrated monomeric fractions of both M, and M $\Delta$ TIR protein, that underwent the following preparation; NiA purification with 10c/v imidazole wash, concentration of NiA elution to 2 ml, S200 PG GF separation and re-concentration of the monomeric fraction to 1 ml. Adenine nucleotides were not detected in the reconcentrated monomeric fractions containing the M proteins. It is probable that nucleotide disassociation during the NiA, GF and two concentration steps may explain such a result. Therefore in repeat experiments, NiA purified proteins were analysed immediately after concentration. In both M and M $\Delta$ TIR protein samples residual levels of ATP were observed (not graphed), however, ADP concentrations were significantly greater than the buffer alone, suggesting that ADP was co-purifying with both proteins (Figure 4.1F). At this point, the M proteins were not homogenous (Figure 4.1A/B) and therefore the accurate determination of M protein concentration using a chemical-based assay was not possible. We could however approximate, that the ADP concentration was about 10-20% that of the total protein concentration. If, therefore, all of the R protein present was capable of binding ADP, 10-20% of that protein would be ADP bound. To ensure that the presence of ADP was not due to the co-purification of a yeast protein, a *P. pastoris* strain transformed with an empty pPICzB vector was induced, and protein was purified in the same way as for the R proteins (Figure 4.1E). For clarity this will be designated as an empty vector control in the subsequent text. The ADP present in the protein purified from the empty vector control was substantially lower than that observed for the purifications of M and M $\Delta$ TIR

(Figure 4.1F). L6 and L6 $\Delta$ TIR proteins were purified in the same way and both of these protein preparations were also found to contain ADP (Figure 4.1C/D/F).

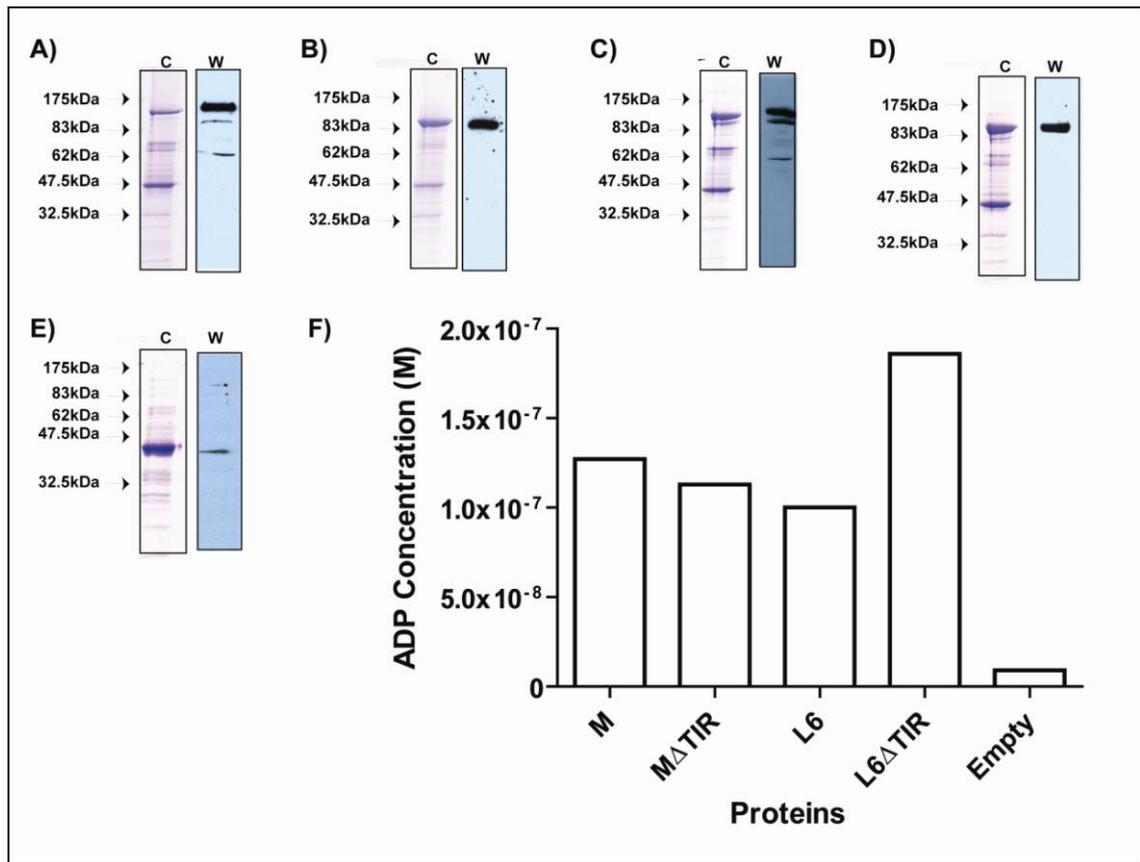
To provide further evidence of an association between ADP and the M protein, M and M $\Delta$ TIR proteins were subjected to GF and the resulting fractions were analysed for the presence of ADP. M $\Delta$ TIR elutes from S200 PG GF at the predicted size of monomeric protein, peaking at approximately 77 ml. The peak ADP concentration was observed at 75 ml, which is one fraction before the protein peak, however, the UV 280 nm trace and the ADP concentration correlated closely (Figure 4.2A). Interestingly, the ADP concentrations increased again after the protein peak, between fractions 93-107 ml. This is likely to be a result of ADP disassociating from the proteins during their migration through the GF column. Upon disassociation, ADP would migrate much more slowly through the column and subsequently elute in later fractions. The native ADP (ATP) peak on S200PG GF is at approximately 116-117 ml (data not shown). The results outlined in Figure 4.2A were repeated in subsequent experiments (data not shown) and strongly suggested that ADP bound to the M $\Delta$ TIR protein is disassociating during migration through the GF column. Analysis of the M protein demonstrated a similar phenomenon to that shown for M $\Delta$ TIR. The monomeric M protein peak between 73-75 ml, as shown by UV 280 nm trace (Figure 4.2B) and western analysis of fractions (Figure 4.2C), was slightly earlier than for M $\Delta$ TIR, which was expected due to the differences in monomeric molecular weight. Importantly, the measurement of ADP in the eluted fractions also peaked earlier at 73 ml. The ADP profile correlated strongly to the elution profile of the monomeric forms of both M and M $\Delta$ TIR, supporting the conclusion that ADP is associated with monomeric forms of the M protein.

To determine the percentage of M and M $\Delta$ TIR protein molecules that contained a bound ADP molecule, repeat NiA purifications were subjected to ATP/ADP quantification assays, however this time protein concentration was measured more accurately. Proteins were separated by SDS-PAGE, stained with Sypro Ruby, and quantified against a standard curve, generated by known protein quantities of BSA run on the same gel (see 2.8.3). The percentage of protein molecules that are bound (occupied) with an ATP or ADP molecule is presented here as a percentage of ATP or ADP occupancy. As an NB-ARC-LRR R protein is predicted to contain one nucleotide binding site, the assumption that one protein molecule is capable of binding one ATP or ADP molecule is acceptable. Furthermore to calculate the percentage of R protein molecules that are either ATP or ADP bound, the values, in terms of molar amounts, of ATP and ADP obtained for the empty vector control were subtracted from the values obtained for all protein samples.

The percentage of M and M $\Delta$ TIR protein that was occupied with ADP was indistinguishable at  $35\% \pm 2.95\%$  and  $41\% \pm 6.81\%$  respectively. In both protein preparations

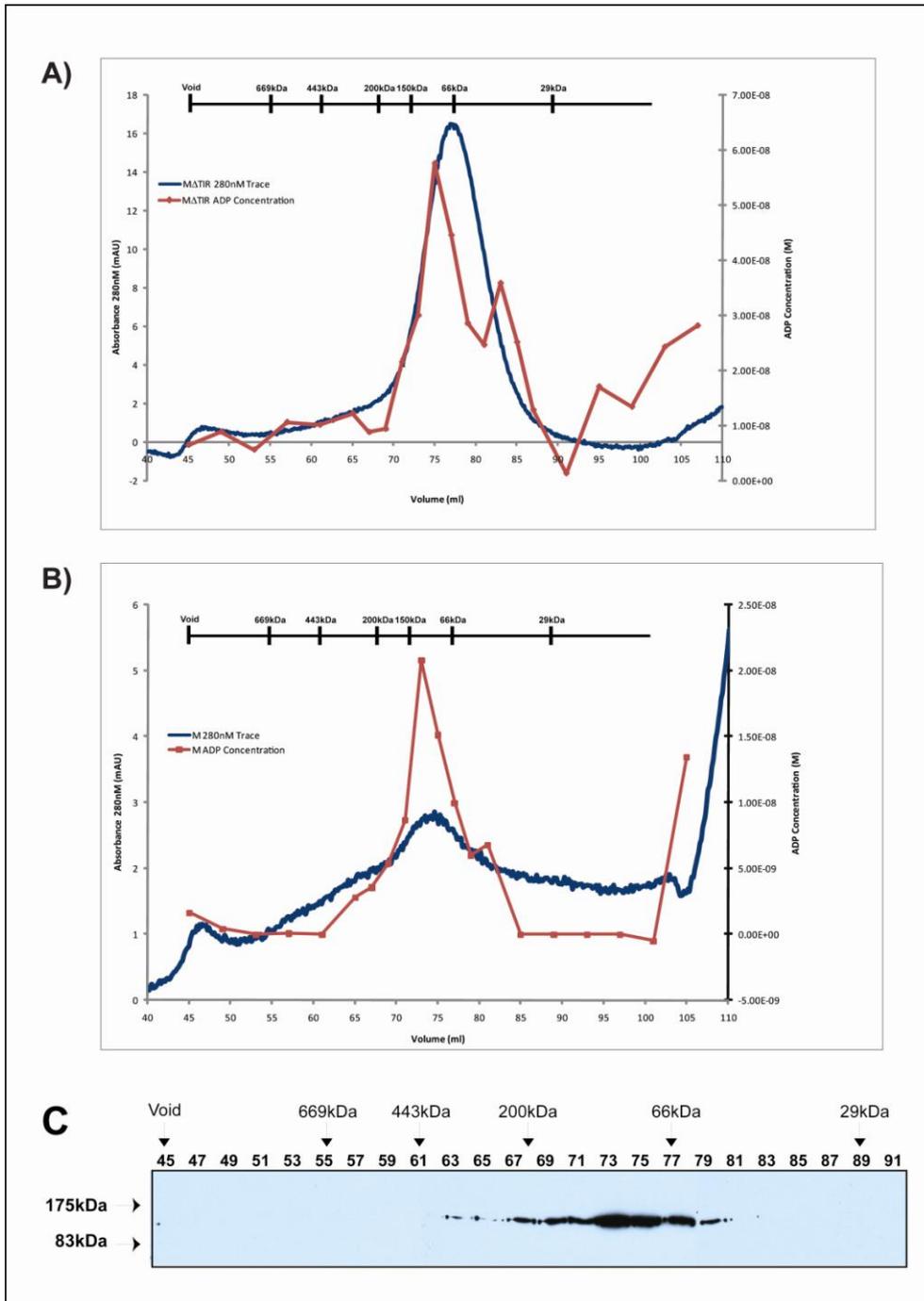
only residual levels of ATP were found (Appendix 5) (Figure 4.3). The same protein samples (stored at 4°C) contained virtually no ADP when re-analysed 24 hours later. Efforts to rebind nucleotides to unoccupied M and M $\Delta$ TIR protein were made by incubating proteins with 100x molar excess amounts of ATP or ADP (and Mg<sup>2+</sup>). Before ATP/ADP quantification analysis was performed, excess unbound nucleotides were separated from the protein samples by running the samples through NAP10 desalting columns (see 2.9.2, use of NAP10 columns). Under these conditions both proteins were unable to reform an association with the supplied nucleotides (data not shown).

From the results of the experiments outlined above, there can be no doubt that endogenous ADP is associated with M and L6 proteins, and that the removal of the TIR does not affect this association. In order to identify specific motifs within the M protein that were influencing nucleotide binding, mutant proteins were analysed in the same way as for the M and M $\Delta$ TIR. The obvious candidates were mutations within the NB-ARC domain that are predicted to affect protein function (see 4.1).



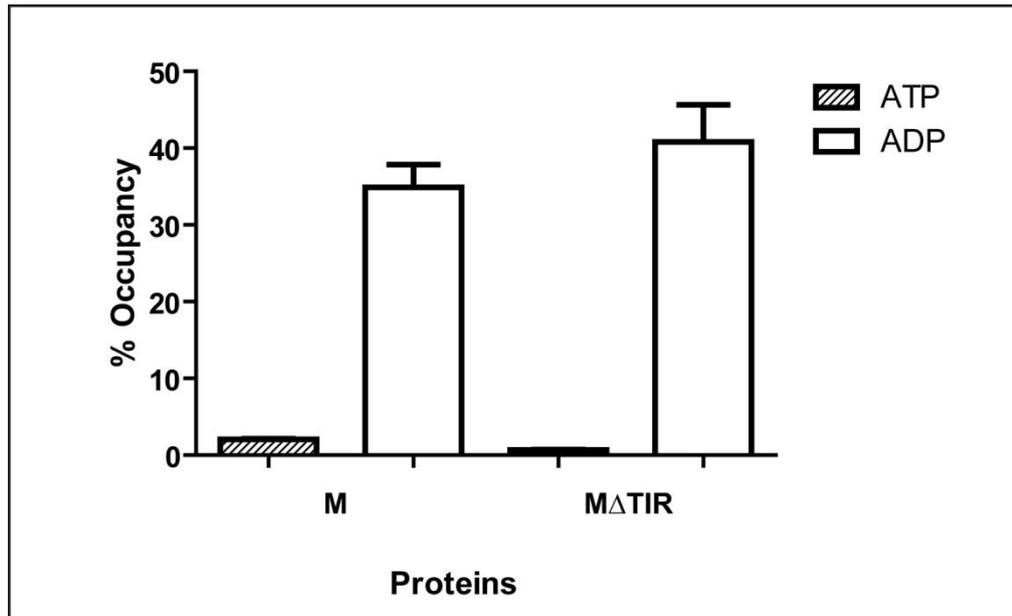
**Figure 4.1: M and L6 proteins after NiA purification are associated with ADP**

Proteins purified using NiA, with a 10c/v wash, were concentrated and a portion of that protein was separated by SDS-PAGE and analysed by Coomassie staining (c) or western blot using an anti-His antibody (w); M (A), M $\Delta$ TIR (B), L6 (C), L6 $\Delta$ TIR (D) and Empty vector (E). An ATP/ADP quantification assay was performed to detect any endogenous ATP and ADP. F) ATP levels were negligible and are not shown, while the concentration of ADP in the respective protein samples is illustrated. No attempt was made here to determine the relationship between ADP and protein concentration (ie occupancy of nucleotide).



**Figure 4.2: Gel filtration fractions containing M and M $\Delta$ TIR also contain ADP**

A) Following NiA purification, M $\Delta$ TIR protein was further separated over a S200PG GF column and the UV 280 nm trace (blue) shows its elution profile. Fractions from GF were processed to determine the ADP concentration and the results were graphed (red). B) As for (A), however, in this case the M protein was analysed. C) The GF fractions from M protein GF analysis were subjected to SDS-PAGE and western analysis using the anti-M antibody. Molecular weights above the gel are positioned where proteins of known molecular weight elute from the same column.



**Figure 4.3: Percentage of M and M $\Delta$ TIR proteins occupied with ATP or ADP**

M and M $\Delta$ TIR proteins purified from NiA with a 10 c/v wash were concentrated and the endogenous ATP and ADP concentrations were determined. The percentage of M and M $\Delta$ TIR protein occupied with ATP and ADP was calculated by comparing nucleotide concentration to protein concentration, the latter determined by the Sypro Ruby staining technique. The percentage occupancy values for M and M $\Delta$ TIR proteins represent the mean which was calculated from 4 independent purifications originating from 2 independently grown cell culture expressions. Error bars represent standard error.

#### 4.2.2 Mutations within the NB-ARC domain that affect protein function have different nucleotide binding capacities

As demonstrated in Chapter 3, M proteins containing mutations within the NB-ARC domain could be purified in the same fashion as M, thereby facilitating their biochemical analysis. The mutations generated were predicted to affect aspects of M protein function, a process that is likely to require nucleotide binding. In light of this, and the observed association between M and ADP, the ATP/ADP quantification assay was used to determine if endogenous ATP and ADP association was present in preparations of the mutated proteins. To enable direct comparison between different mutants, R protein concentrations were determined by Sypro Ruby staining (see 2.8.3) and an example of such a gel is shown in Figure 4.4. An on-gel BSA standard curve (Figure 4.4, inset) was used to quantify the concentration of full-length M proteins. Quantification of endogenous ATP and ADP in the various mutants was performed and this information is tabulated in Appendix 5.

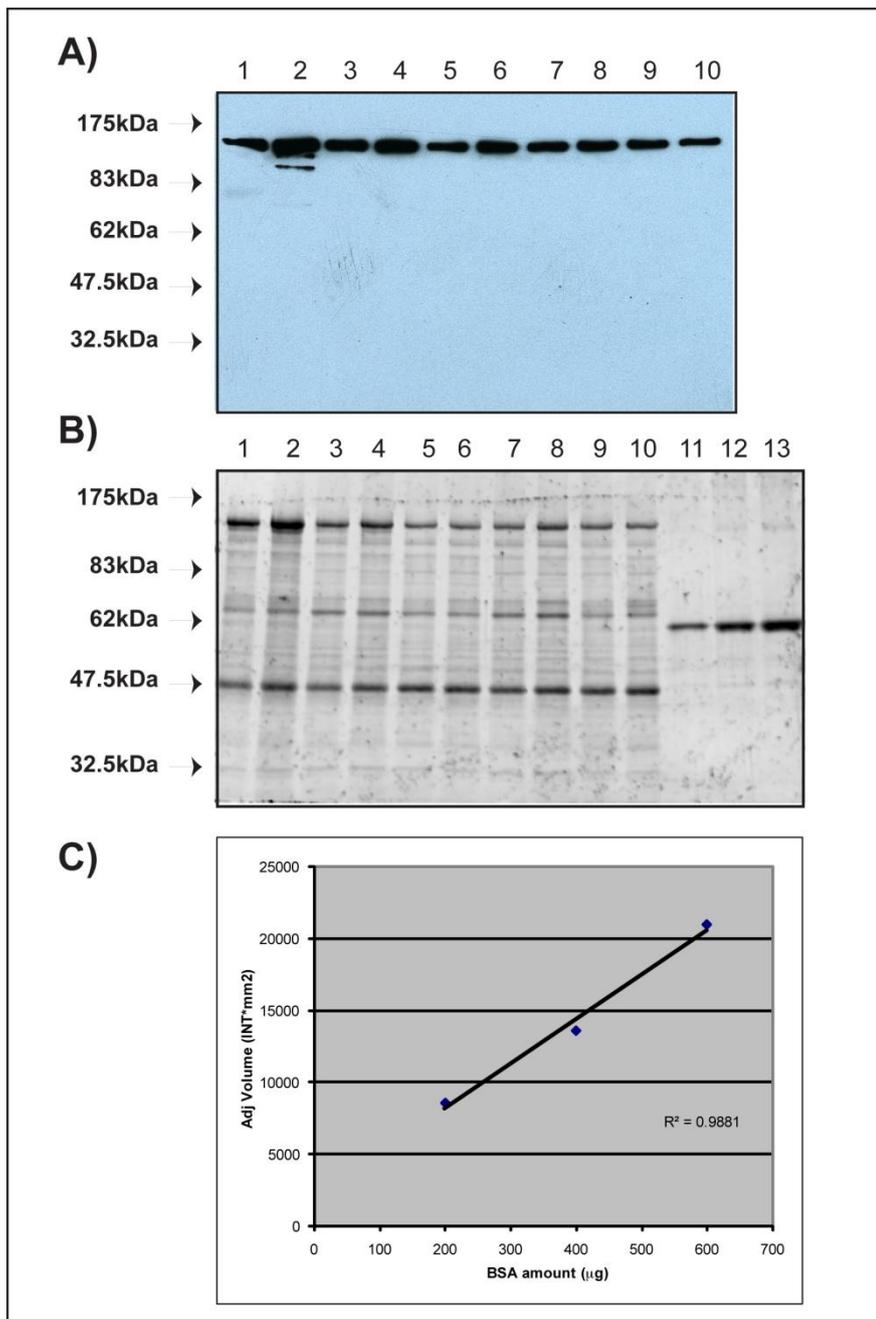
Mutation of the invariant lysine to a leucine within the P-loop motif ( $M^{K286L}$ ) is anticipated to cause a loss of function in M. The M protein had a percentage ADP occupancy of  $35 \pm 2.95\%$ , while in the  $M^{K286L}$  protein, the percentage ADP occupancy was  $2.7 \pm 1.30\%$  (Figure 4.5). This is approximately a 13-fold reduction in the relative amount of bound ADP and indicates that the P-loop lysine in M is critical for ADP binding. This result also confirmed that the NB-ARC domain of M is responsible for the binding of ADP.

In contrast to wildtype M and  $M^{K286L}$ , a mutant protein containing an aspartate to valine change in the MHD motif ( $M^{D555V}$ ), was associated with higher percentage ATP occupancies ( $18 \pm 4.15\%$ ) than ADP ( $5.40 \pm 2.03\%$ ) (Figure 4.6A). As discussed above, an aspartate to valine change within this motif results in the autoactivation of the HR in a number of other R proteins, including L6. It was of great interest to analyse further putative autoactive mutants within the ARC2 subdomain. For this reason, and with reference to the work carried out on SSi4 (Shirano et al., 2002), the serine in the lesser characterised motif VIII was mutated to an arginine ( $M^{S492R}$ ).  $M^{S492R}$  protein was also found to be associated with ATP; however, to a lesser extent than the  $M^{D555V}$  protein. The percentage ATP occupancy of ( $8.49 \pm 1.57\%$ ) and ADP occupancy of ( $4.69 \pm 2.50\%$ ) are statistically indistinguishable (Figure 4.6B).

Of particular interest was the finding that endogenous levels of ATP in both  $M^{D555V}$  and  $M^{S492R}$  are considerably greater than M (Figure 4.6C). The percentage ATP occupancies in M was  $2 \pm 0.08\%$ , while for  $M^{D555V}$  and  $M^{S492R}$  the percentage ATP occupancies was  $18 \pm 4.15\%$  and  $8.5\% \pm 1.57\%$ , respectively. In the case of the  $M^{D555V}$ , a mutation in this protein at the conserved lysine in the P-loop motif ( $M^{K286L+D555V}$ ) abolished any ATP binding, with measured

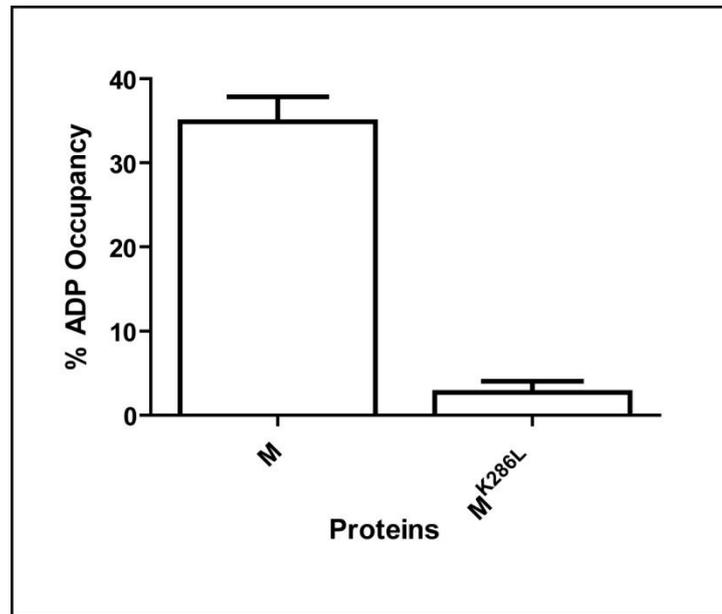
percentage ATP occupancies of  $0.8 \pm 0.08\%$ . This demonstrated that a functional P-loop is required not only for the binding of endogenous ADP in M, but also for the binding of endogenous ATP in the M<sup>D55V</sup> protein.

These results provide further *in vitro* evidence towards the model of R protein activation, first suggested by the Takken group (Takken *et al.*, 2006, Tameling *et al.*, 2006). Within this model, the role of hydrolysis is suggested to enable protein recycling/re-signalling. To our knowledge, hydrolysis has yet to be demonstrated in any full-length, or near full-length, R proteins. Given that flax R proteins could be purified with endogenously derived nucleotides, we now wanted to determine if they were capable of hydrolysing this nucleotide. ATP hydrolysis therefore became the focus of investigation.



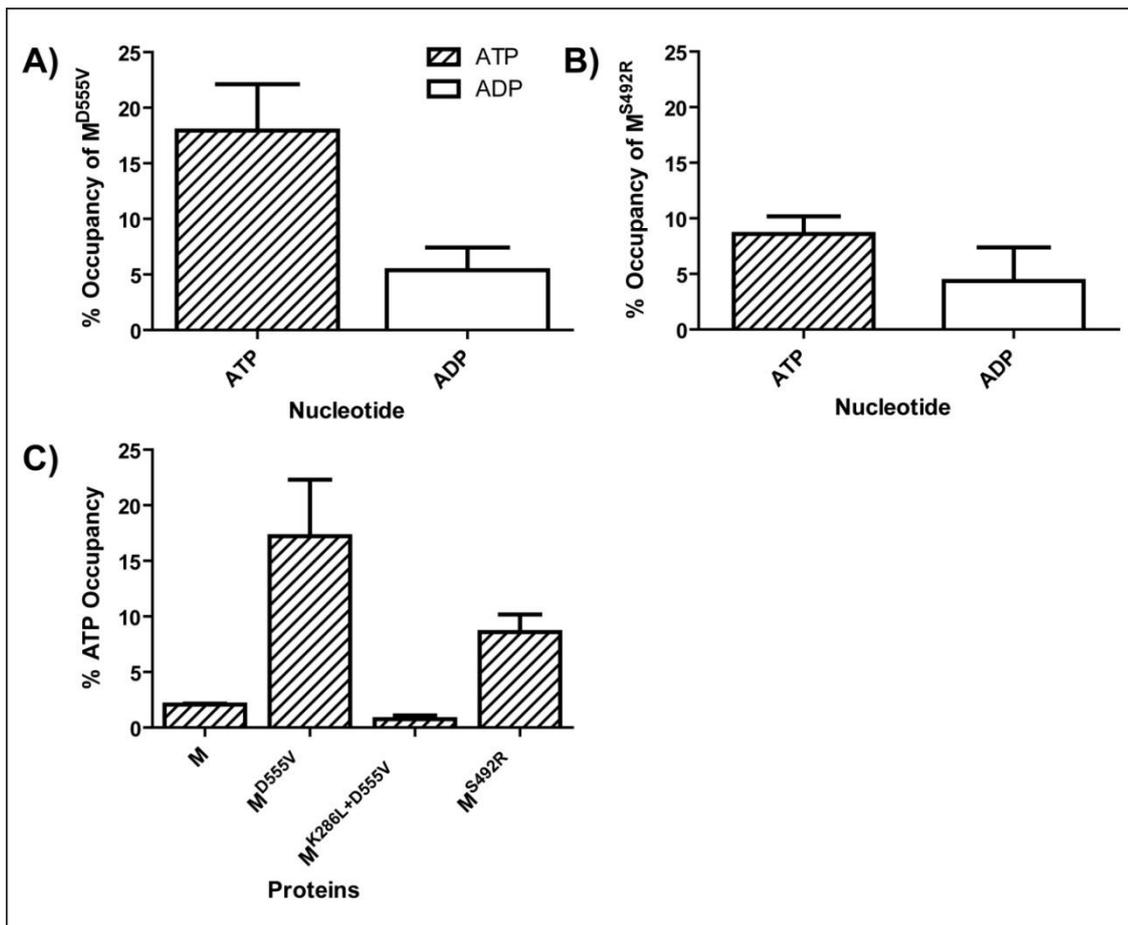
**Figure 4.4: Purification of M and M mutants for ATP and ADP quantification**

Independent purifications of M, and mutant M, proteins purified from NiA (10 c/v imidazole wash) and concentrated following elution. 5 µl of each purified protein sample was loaded and analysed by SDS-PAGE and western blot using an anti-M antibody (A) and Sypro Ruby stained for quantification (B). Lane arrangement of (A): lanes 1 and 2, M purifications 1 and 2; lanes 3 and 4, M<sup>K286L</sup> purifications 1 and 2; lanes 5 and 6, M<sup>D555V</sup> purifications 1 and 2; lanes 7 and 8, M<sup>K286L+D555V</sup> purifications 1 and 2; lane 9 and 10, M<sup>S492R</sup> purifications 1 and 2. Lane arrangement of (B): same as (A) except lanes 11-13, are loaded with BSA at 200 µg, 400 µg and 600 µg, respectively. (C) Graph represents the BSA generated standard curve used to determine M protein concentration. Adjusted volume was calculated as a measure of amount of Sypro Ruby staining and used for protein quantification, for details see 2.8.3.



**Figure 4.5** ADP binding is dependent on a functional P-loop

M and M<sup>K286L</sup> proteins purified using NiA with a 10 c/v wash were concentrated post elution and their endogenous ADP concentrations were determined. Percent ADP occupancy was calculated by comparing nucleotide concentration to protein concentration (as determined by the Sypro Ruby staining technique). The percentage ADP occupancy values for M and M<sup>K286L</sup> represent the mean, which was calculated from 4 independent purifications originating from 2 independently grown cell culture expressions. Error bars represent standard error.



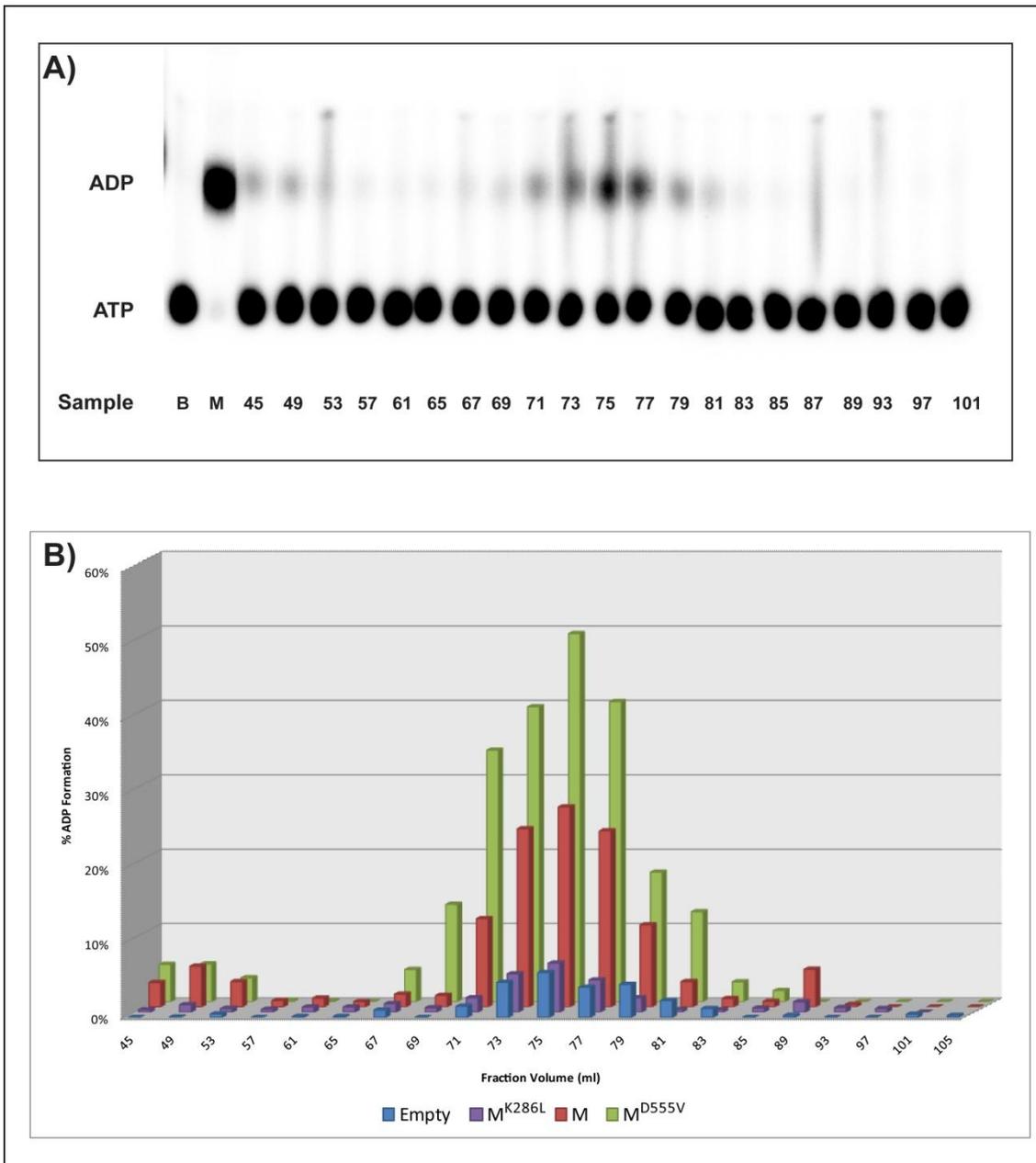
**Figure 4.6: Putative autoactive mutants are purified with ATP bound**

Mutant M proteins purified using NiA with a 10c/v wash were concentrated post elution and their endogenous ATP and ADP concentrations were determined. The percent ATP and ADP occupancy was calculated by comparing nucleotide concentration to protein concentration (as determined by the Sypro Ruby staining technique). A) The percentage of M<sup>D555V</sup> protein occupied with ATP and ADP, (B) The percentage of M<sup>S492R</sup> protein occupied with ATP and ADP (C) A comparison of the percentage of M, M<sup>D555V</sup>, M<sup>K286L+D555V</sup>, M<sup>S492R</sup> proteins occupied with ATP. The percentage ATP and ADP occupancy values graphed represent the mean, which was calculated from 4 independent purifications originating from 2 independently grown cell culture expressions. Error bars represent standard error.

### 4.2.3 ATP hydrolysis analysis of M proteins

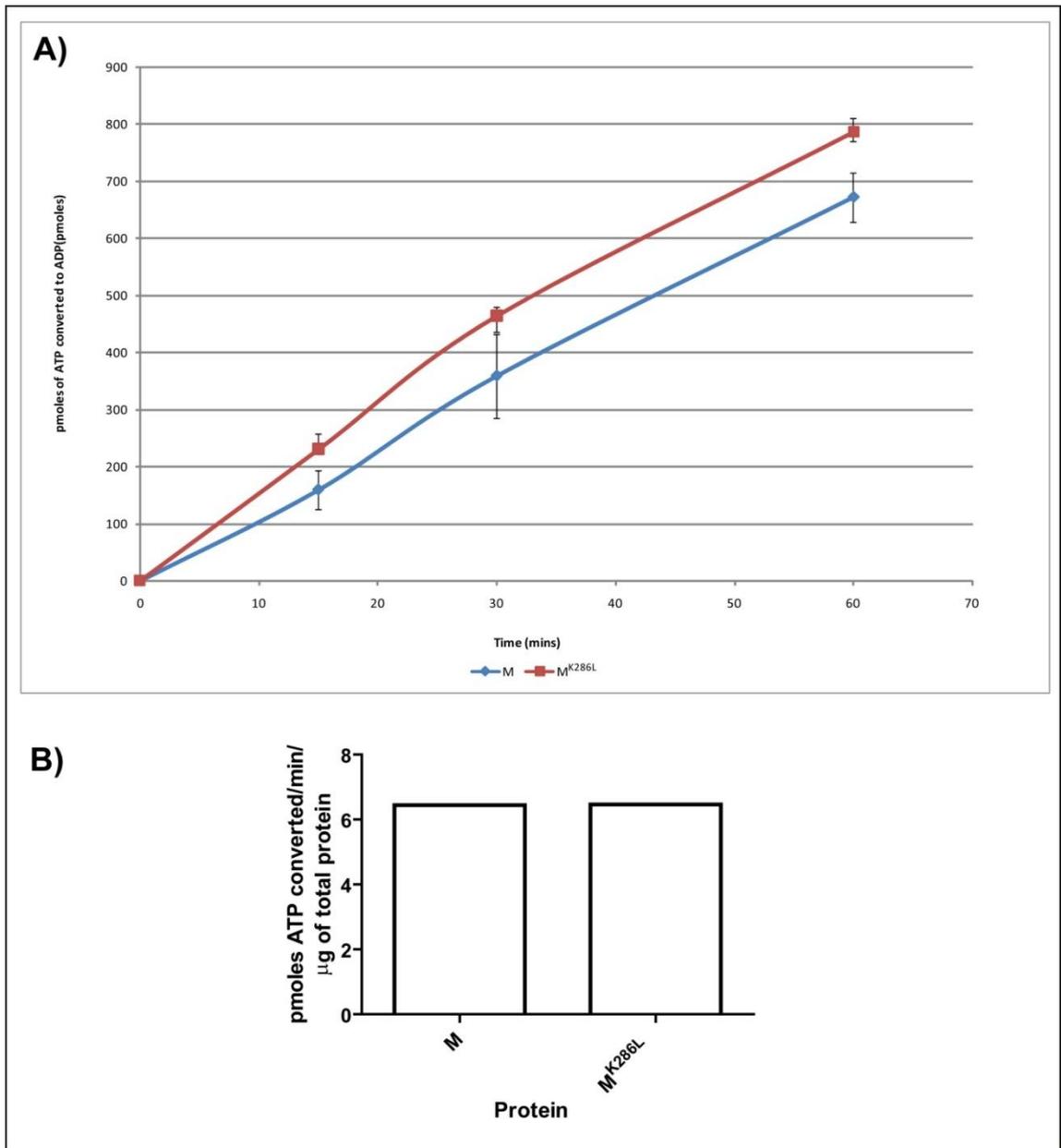
Given that M contains all the catalytic hallmarks of a protein capable of hydrolysing ATP, there was considerable interest in determining if M protein, purified from the *P. pastoris* expression system, had such activity. To investigate this, M, M<sup>K286L</sup>, M<sup>D555V</sup> and protein from an empty vector control were purified by NiA chromatography using a 10 c/v wash. Protein was further separated over GF and the resulting GF fractions were analysed for ATPase activity (Figure 4.7A). Hydrolysis activity was observed in fractions that contained M proteins (69-79 mls, Figure 4.2C) and the peak hydrolysis conversion was observed close to the protein peak of monomeric M (73-75 mls). The same general profile of ATPase activity was, however, observed in the empty vector control, although it was less pronounced. The ATP hydrolysis rates were higher for M<sup>D555V</sup> and M in comparison to M<sup>K286L</sup> and empty vector (Figure 4.7B), although specific rates of conversion could not be calculated as the protein concentrations of the GF fractions were too low to be accurately determined.

To interpret this further, M and M<sup>K286L</sup> proteins were purified using the same NiA chromatography strategy, and the monomeric fractions from GF were reconcentrated to 500  $\mu$ l and analysed for ATPase activity. Surprisingly, very similar conversion rates were observed for M and the M<sup>K286L</sup> proteins, and the specific activity of both proteins was indistinguishable (Figure 4.8). Our previous results had demonstrated that a functional P-loop was critical for both ATP and ADP binding (see 4.2.2). It therefore appeared that the ATPase activity observed in preparations of the M protein is most likely a contaminant ATPase that coincidentally elutes close to the monomeric elution range of M. In an attempt to remove such contaminant proteins from preparations of M a 100 c/v imidazole wash of M bound to the NiA column was incorporated into the purification protocol. In this experiment, M, M<sup>D555V</sup>, M<sup>K286L+D555V</sup> and protein from the empty vector control were included and protein assays were performed on duplicate protein purifications. The conversion rates observed for the empty vector were not notably different from M<sup>D555V</sup>, M<sup>K286L+D555V</sup> proteins while variation was observed in M, below and above conversion rates for protein from the empty vector control (Figure 4.9). It was found that after several repeat experiments, the extraction of consistent and meaningful data from our ATP hydrolysis experiments was not possible. Unfortunately, under these purification conditions, any ATPase activity, above that of the co-purifying contaminating ATPases, could not be attributed to the M or the putative autoactive M<sup>D555V</sup> protein. Similar background levels of ATPase activity also confounded the analysis of the M $\Delta$ TIR protein (data not shown).



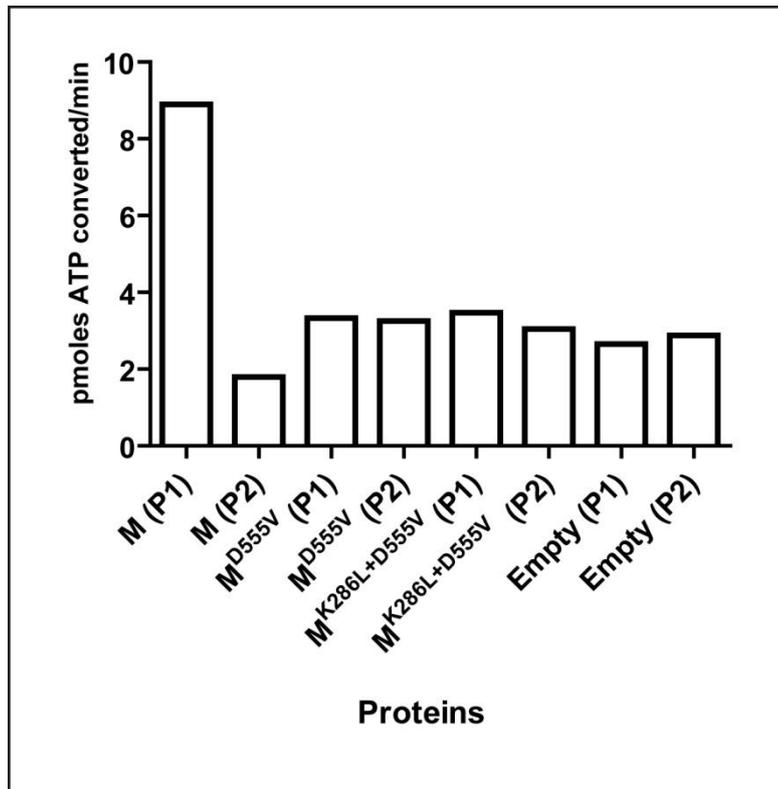
**Figure 4.7: ATPase activity in GF fractions of M and M-mutants**

Fractions from the GF of concentrated M, M<sup>K286L</sup>, M<sup>D555V</sup> and empty vector were analysed for hydrolysis activity in an ATPase assays. A) Thin layer chromatography was used to separate the ATP and ADP resulting from the ATPase reactions. Sample B represents buffer only control (negative control), sample M represents a myosin positive ATPase control, and samples 45-101 represent the fraction volumes analysed from the M purification. B) The conversion rates in the GF fractions were tabulated as a percentage of ADP formed. The graph represents M (red), M<sup>K286L</sup> (purple), M<sup>D555V</sup> (green) and empty vector (blue).



**Figure 4.8: M and M<sup>K286L</sup> proteins have indistinguishable hydrolysis activities**

A) The conversion of ATP to ADP in a 60 minute reaction, observed from fractions containing concentrated monomeric M and M<sup>K286L</sup> proteins. B) The specific hydrolysis activity of both protein preparations.



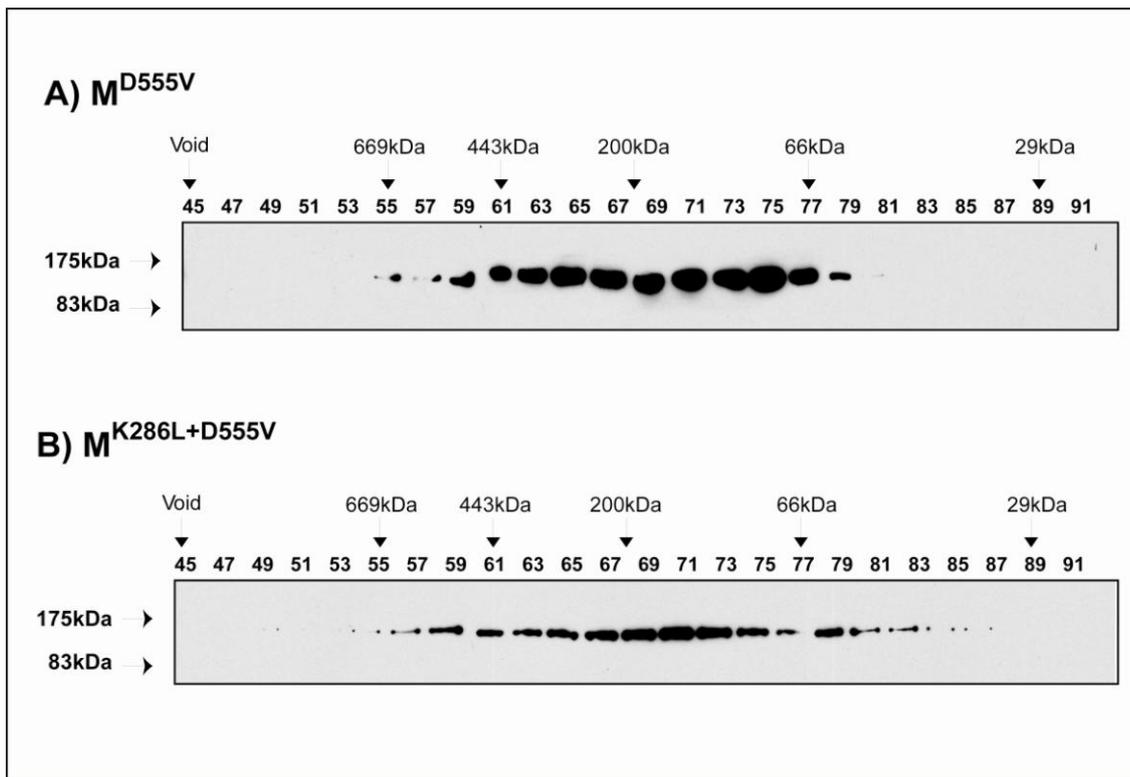
**Figure 4.9: Long wash NiA purification does not remove background ATPase activity**

ATPase activity of NiA purified M and M mutants utilising a 100 c/v (55 mM imidazole) wash (P1: purification 1, P2: purification 2). The ATPase activity was calculated as the picomoles of ATP converted per min for each protein preparation.

#### 4.2.4 Do autoactive M proteins form oligomers?

By drawing analogy to their counterparts in mammalian innate immunity, oligomerisation of R proteins has been speculated upon, although only one clear indication of this exists in the literature. *In planta* studies of the tobacco N protein, using immunoprecipitation techniques, demonstrated that N oligomerisation is mediated by its viral elicitor, p50, and is dependent on a functional lysine within the P-loop (Mestre and Baulcombe, 2006). This suggests, that oligomerisation is involved in N protein activation, and is dependent on the protein's ability to bind a nucleotide. Currently, no *in vitro* investigation of R protein oligomerisation has been published.

The results of this thesis demonstrate that the M protein is purified predominantly in the monomeric form, and a proportion of this monomeric protein has ADP bound within its NB pocket (Figure 4.2B/C). To determine if a putative activated form of M has an intrinsic capacity to oligomerise the  $M^{D555V}$  purified protein was analysed by GF. To do this,  $M^{D555V}$  protein was purified in the presence of excess ATP (which the protein is known to bind) and magnesium acetate. The purified protein was then separated over GF with the same nucleotide components in the mobile phase, as that in the purification buffers (Figure 4.10A). The results suggest that two potential  $M^{D555V}$  protein forms exist. It is important to note that this profile is different to any salt or imidazole induced R protein aggregation that had been observed earlier. The monomeric form of  $M^{D555V}$ , also observed for M, had a peak fraction at 73 mls, while the other peak fraction at 65 mls, was indicative of a protein size of 275 kDa, the approximate dimeric form of the protein. The predicted non-functional form of this ATP-bound mutant,  $M^{K286L+D555V}$ , was purified and analysed in the same way as  $M^{D555V}$  (Figure 4.10B). ATP/ADP quantification studies had demonstrated that this protein is unable to bind ATP, while a study of tobacco N with a single P-loop (lysine) mutation show such a protein does not oligomerise (Mestre and Baulcombe, 2006). The results of our GF analysis demonstrate that the  $M^{K286L+D555V}$  protein is observed in all the same fractions as the  $M^{D555V}$  protein (Figure 4.10A/B). This makes it difficult to determine if a difference in oligomeric state actually exists between the two proteins. A clear difference does exist, however, between the GF profile of M (Figure 4.2C) and  $M^{D555V}$  (Figure 4.10A). It should be noted that the former experiment was performed without ATP and magnesium in the mobile phase of GF. Although there is a hint of a dimeric form of  $M^{D555V}$  protein in the GF profile, little can be concluded from this result. The potential oligomeric properties of the  $M^{D555V}$  proteins are currently under further investigation and the outcomes from some of these additional experiments are discussed below (P. Somaraj, personal communication).



**Figure 4.10: GF analysis of M<sup>D555V</sup> and M<sup>K286L+D555V</sup>**

NiA concentrated M<sup>D555V</sup> and M<sup>K286L+D555V</sup> protein, purified in the presence of ATP and Mg<sup>2+</sup>, were further separated over a Superdex 200 PG GF column. The GF buffer was also supplemented with ATP. The 2 ml fractions were analysed by western blot with an anti-M antibody. Molecular weights above gels are positioned where proteins of known molecular weight elute from the same column.

### 4.3 Discussion

The results presented in this chapter clearly indicate that *P. pastoris* is capable of manufacturing recombinant flax R proteins that possess, at least some of, the functional capabilities predicted of the NB-ARC-LRR class of R proteins. It is demonstrated here, that M and L6 have a capacity to bind endogenous ADP, while mutants of M, predicted to display an autoactive phenotype, have considerably greater levels of bound endogenous ATP. Importantly, the binding of both endogenous ADP and ATP to the M and M<sup>D555V</sup> protein, respectively, is dependent on a functional P-loop motif. Whilst an ATP hydrolysis study was undertaken, the role of ATP hydrolysis in M protein function is yet to be resolved. Data is also presented that raise questions regarding the potential for R protein oligomerisation in the ATP binding mutant, M<sup>D555V</sup>.

Collectively, the information ascertained from this biochemical study generally supports the conclusions made in published biochemical studies performed on other R proteins. Our results provide further evidence towards a model of R protein activation that involves an “off”, ADP bound state, and an “on”, ATP bound state; however, the *in planta* phenotypes of the M mutants studied need to be established before this can be confirmed.

#### 4.3.1 ADP bound is the “OFF” state

The ADP measured in preparations of M, L6 and the TIR truncations was endogenous, demonstrating that M and L6 are capable of binding ADP within *P. pastoris* cells. While the kinetics of the ADP interaction remain unknown, a high affinity between M and ADP is likely, given that the nucleotides remain associated with a portion of the protein post-purification. However, loss of the bound nucleotide from the M protein is observed during GF and both extended purification, and storage of the protein following purification, results in the eventual complete loss of nucleotide from the NB-ARC domain.

It is tempting to speculate that 100% of M protein produced in *P. pastoris* has ADP in its NB pocket at the point of cell lysis, and nucleotide loss during purification is the reason that only 35% of the purified protein contains ADP at the point of nucleotide assay. However, it is possible that only a portion of the recombinant proteins are produced with nucleotide bound. Experiments concurrently run in the laboratory with those presented here demonstrated that M protein purified in the presence of nucleotides (ATP and ADP) and Mg<sup>2+</sup> have indistinguishable nucleotide occupancies compared with those purified without (P. Sornaraj, personal communication). Also attempts to reconstitute an ADP bound form of the M protein that has lost its nucleotide, by bathing the protein in exogenous nucleotide, have so far proved unsuccessful. Taken together,

these results suggest that the M protein is not capable of rebinding ADP; however, further analysis using alternative assays is required before this can be indisputably concluded.

The demonstration of an M·ADP and L6·ADP endogenous association is a significant finding in the context of biochemical studies of both R proteins and STAND proteins. In studies of I-2 truncated for the LRR (I-2N), Tameling et al., (2006) present data that support the idea that the I-2N protein forms a stable complex with ADP, while the complex of I-2N·ATP is unstable and could not be captured experimentally. Their findings suggested that I-2N has a higher affinity for ADP, rather than ATP, and a low ADP disassociation rate, which is a likely result of conformational changes induced through ADP binding (Tameling et al., 2006). It was these findings that stimulated the presentation of a model that predicts R proteins are ADP bound in a preactivated/off/autoinhibited state (Tameling et al., 2006). Studies of STAND proteins have also supported the idea that an ADP molecule is bound to the protein in the autoinhibitory state. In particular, the STAND protein, MalT, is purified as a monomer with endogenous ADP (Marquetet and Richet, 2007). In the case of Apaf-1, also a monomer in its autoinhibited form, there is some conjecture regarding the identity of the bound nucleotide. The crystal structure of Apaf-1, truncated for its C terminal WD-40 repeat domain, was solved with ADP bound in its NB pocket. Biochemical studies have demonstrated that pre-activated Apaf-1 is either dADP/ADP bound (Bao et al., 2007, Riedl et al., 2005), however, other biochemical evidence suggest that Apaf-1 is dATP bound, and that the hydrolysis of dATP to dADP precedes its activation (Kim et al., 2005). The results presented in this study demonstrate that M protein can be purified from *P. pastoris* as a monomer, and the purified monomer is ADP bound. These observations strongly support the concept that an R protein resides in an ADP bound form in its autoinhibitory state.

### 4.3.2 ATP bound is the “ON” state

If the autoinhibitory, or “off”, state of the R protein is ADP bound it implores the question; how would a mutation that converts the protein into an autoactive state affect the preference and/or interaction of a bound nucleotide? Indeed Tameling et al., (2006) have gone some way to answering this question. In studies of I-2, they demonstrated that mutations in the NB subdomain that caused autoactivity *in planta*, had a reduced capacity to hydrolyse ATP (Tameling et al., 2006). These findings inferred that hydrolysis was not a requirement for R protein activation, and suggested that the mutated proteins were held, or stuck, in an ATP bound state. Their result led to the concept that an R protein is active in an ATP bound state (Tameling et al., 2006). As yet, no other biochemical studies of R proteins have been published to support or refute this theory, however, support for an activated ATP bound state has come from work with the STAND proteins, Apaf-1 and MalT (Bao et al., 2005, Kim et al., 2005, Marquetet and Richet, 2007). The

*in vitro* studies presented here, also add support to this theory by demonstrating that purified M proteins with mutations in the NB-ARC domain, which are predicted to generate an autoactive phenotype *in planta*, are bound with considerably higher amounts of endogenous ATP, compared to wildtype M. The two mutations analysed here, are located in the motif VIII and MHD motif, which are at the N- and C-terminal ends of the ARC2 subdomain, respectively. To explore the potential reasons for the observed differences in nucleotide preference in comparison to wildtype M, and the predicted associated changes to the *in planta* phenotype, the ARC2 domain and the motifs MHD/motif VIII require further definition.

Growing evidence suggests that the ARC2 subdomain is critical for the regulation and activation of an NB-ARC domain. The Apaf-1 NB-ARC three-dimensional structure revealed that residues in the ARC2 subdomain (winged-helix domain) were involved in the coordination of ADP binding. This was a unique feature not found in the structurally similar AAA+ ATPase proteins (Riedl et al., 2005). Reidl et al., (2005) also suggested that the ARC2 subdomain may be prone to conformation changes, and predicted that the domain would require significant reorientation in order to form the heptamer structure of the pro-caspase 9 activating apoptosome. More recent biochemical and structure studies of STAND proteins have supported this conclusion. In a structure-based review, Danot et al (2009) suggest that the ARC2 subdomain of STANDs undergoes major restructuring to transform the protein from a preactivated to an activated state. Not surprisingly, this domain is thought to be critical in the regulation of STAND protein activation (Danot et al., 2009). A host of evidence also exists that implicates the ARC2 subdomain as a critical component in the regulation and activation of an R protein. Studies of Rx showed that autoactive mutations tend to map to the ARC2 subdomain (Bendahmane *et al.*, 2002), and that this domain is involved in the maintenance of the autoinhibitory state in the absence of the effector, and the activation state when the effector is present (Rairdan and Moffett, 2006).

In R proteins, three conserved motifs are located in the ARC2 subdomain, namely motif VIII, the RNBS-D and MHD motif. Autoactivating mutations have been localised to all three motifs (Figure 1.3), however, in this study motif VIII and the MHD motif were investigated. The RNBS-D motif was not included as the consensus sequence of this motif varies in TIR and non-TIR proteins (Table 1.2), and the original mutational work was performed in the non-TIR protein, Rx (Bendahmane et al., 2002).

The importance of the MHD motif in R protein function has been well characterised in a number of R proteins including Rx, I-2, L6, Mi-1 and Rpi-blb1 (Bendahmane et al., 2002, de la Fuente van Bentem et al., 2005, Howles et al., 2005, van Ooijen et al., 2008b). Recent exhaustive structure/function studies of the MHD motif in I-2 and Mi-1 suggest that it acts as a sensor II motif (van Ooijen et al., 2008b). The concept of the sensor motif originates from studies

of the AAA+ ATPases, whereby key residues (predominantly arginines) act as sensors to transduce the chemical event of ATP hydrolysis and/or binding into a mechanistic outcome for the protein (Ogura et al., 2004). The sensor II in AAA+ proteins has been implicated in functional roles including hydrolysis, ATP binding and the coordination of interactions between subdomains (Ogura et al., 2004). R proteins and Apaf-1 have no recognisable sensor II motif that is comparable to those found in the AAA+ ATPase protein family. Reidl et al., (2005) predict from the three dimensional structure of Apaf-1, that the equivalent LHD motif can replace the sensor II found in the AAA+ ATPase proteins. The results of the R protein study of I-2 and Mi-1 add further support to this concept suggesting that the MHD motif may provide the function of the sensor II. In this role, the MHD is predicted to coordinate subdomain interactions that are dependent on the bound nucleotide state (van Ooijen et al., 2008b).

As mentioned above, structure studies of Apaf-1 demonstrate that the histidine of the equivalent LHD motif helps to coordinate the binding of the  $\beta$ -phosphate of ADP, and links the ARC2 domain to the nucleotide binding pocket. The same functional role has been predicted in I-2 from a model of the I-2 NB-ARC regions that used the Apaf-1 structure as the template. A number of predictions regarding the cause of autoactivity in an MHD to MHV mutation in I-2 have been made using this model (van Ooijen et al., 2008b). The aspartate residue of the MHD motif is situated at the amino-terminal end of an alpha-helix. Generally, negatively charged residues are positioned at the amino-terminal end of a helix to help stabilise the positively charged helix-dipole. Substitution of the negatively charged aspartate, with an uncharged valine, has the potential to destabilise the helix and through repositioning, subsequently weaken or disrupt the interaction of the preceding histidine with ADP (van Ooijen et al., 2008b). These authors also suggest that a change in the aspartate could directly prevent a predicted salt bridge interaction between the aspartate and an arginine, from the so-called sensor I motif (in I-2 corresponds to R313). This could potentially create a more open protein conformation (van Ooijen et al., 2008b).

A structural model of M, using Apaf-1 as a template, was generated in the laboratory and is shown in Figure 4.11 (P. Sornaraj). In this model, the histidine within the MHD motif of M forms the same functional role in its coordination of the  $\beta$ -phosphate of ADP, as observed in the original Apaf-1 structure and the modelled I-2 structure (Riedl et al., 2005, van Ooijen et al., 2008b). Also, the following aspartate occupies the amino-terminal end of a helix, and therefore, the predicted structural changes for I-2 induced by an aspartate to valine change are likely to hold true for M. It could be extrapolated, therefore, that the reduced capacity to bind ADP and the more open conformation of the NB pocket provide the characteristics conducive for the mutated protein to preferentially bind ATP, consistent with the observations of this study. However, in the

absence of a relevant three dimensional structure, the cause of ATP binding can only be speculated upon.

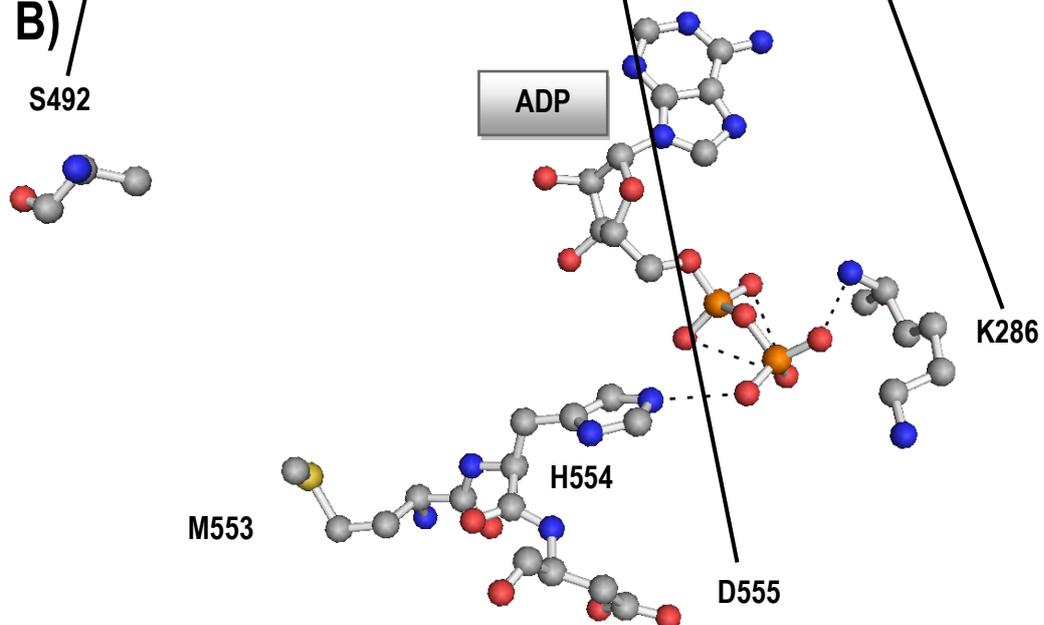
The so-called motif VIII is substantially understudied in comparison to the MHD motif. It was hypothesised that at the position of the serine in this motif, a small amino acid was required to maintain the native protein structure. Glycine is present at this position in the SSi4 protein, and its mutation to the large bulky amino acid arginine disrupts the autoinhibited state of SSi4 enough to lead to autoactivation of the HR (Shirano et al., 2002). In our *in vitro* investigation, the serine to arginine mutation within motif VIII had substantially greater amounts of bound ATP compared to wildtype M protein, however, it was less pronounced than the MHV mutation. It is extremely difficult to speculate why this amino acid substitution would enable the protein to preferentially bind ATP. Unlike for the case of the MHD motif, using structural models to assist in interpreting the effect of the serine to arginine substitution is less appropriate, as the motif aligns poorly to the Apaf-1 and CED-4 sequences in this region (Figure 1.3). Making any further predictions beyond that of a spatial change that is induced by the substitution of a bulky amino acid, such as arginine, at a position that requires a smaller residue, is not possible. The data presented here does, however, suggest a critical role for this motif in protein regulation in the NB-ARC domain of R proteins. Its exact role will remain elusive in the absence of a crystal structure of this region.

The *in vitro* analyses presented for M require the analogous mutations to be tested *in planta* studies. Considering that the L6 MHV mutant has been shown to be autoactive (Howles et al., 2005), it is anticipated that these findings will generally support the conclusions that ATP binding is associated with an activated state of an R protein, and that the ARC2 subdomain has important regulatory roles in R protein function.

A)



B)



**Figure 4.11: A structural model of the NB-ARC domain of M**

The M NB-ARC domain structure model was generated using the ADP bound Apaf-1 structure (PDB code 1z6t, chain A) as a template. A) The structure is presented here in cartoon representation. The subdomains of the NB-ARC are colour coded; NB-red, ARC1-green, ARC2-blue. The residues K286, S492, M553, H554, D555 and the ADP molecule are depicted in ball-and-stick representation. B) Represents the ball-and-stick represented residues (A) with the rest of the structure hidden. Arrows protruding from (B) highlight the positions of the K286, S492 and D555 residues in (A), which were targeted for mutation in this study. Black broken lines represent hydrogen bonds between the residues and bound ADP molecule. Atom colour coding: oxygen-red, nitrogen-blue, phosphorus-orange, sulphur-yellow and hydrogen-grey.

### 4.3.3 Comparisons with published *in vitro* studies of R proteins

From the data presented here conclusions regarding ATP hydrolysis in M and M mutants could not be made. The background ATPase levels observed in our empty vector control prevented the accurate interpretation of any hydrolysis activity derived from purified M protein.

Purified I-2N protein was capable of rebinding both ATP and ADP (Tameling et al., 2006). This does not appear to be a feature of the purified M protein studied here. The most likely reason for this observed disparity is the differences in the two proteins analysed. The most distinct difference between the I-2N study and the M protein analysed, is that M is a near full-length protein, while the I-2N protein was truncated for the LRR domain. The presence of the LRR domain could potentially have a dual effect on the purified M proteins. The LRR region may help to maintain the association with a bound nucleotide during purification; however, once the NB pocket becomes un-occupied it may prevent nucleotide rebinding. The idea that the LRR provides a conformation that enables the protein to maintain a nucleotide association is somewhat supported in the literature. Studies of Rx have proposed that the LRR helps to regulate and maintain the protein in its autoinhibitory state (Moffett et al., 2002, Rairdan and Moffett, 2006), which is potentially dependent on nucleotide binding.

Kinetic studies of the I-2N protein and relevant mutants, demonstrated that it was difficult to capture the protein in its ATP bound state, instead a highly stable ADP bound form of the protein existed (Tameling et al., 2006). The results presented here indicate that the  $M^{D555V} \cdot \text{ATP}$  complex was less stable than the  $M \cdot \text{ADP}$ , as lower amounts of ATP were found associated with  $M^{D555V}$  compared with ADP for M. Nevertheless, the  $M^{D555V} \cdot \text{ATP}$  bound state was maintained by a portion of the purified protein throughout the purification process before quantification analysis. It is therefore possible, that the presence of the LRR is important in maintaining the ATP bound state. Given the results observed for I-2N, it also appears possible that conformational changes induced by the  $M^{D555V}$  mutation enable this protein to form a stable association with ATP.

An inability to achieve ATP or ADP rebinding to wildtype M protein will undoubtedly have a number of effects on future experiments. For example, if M cannot rebind a nucleotide, experiments to determine ATPase activity and the kinetics of nucleotide binding are not possible. However, the fact that recombinant M cannot rebind ADP once disassociated should not necessarily be interpreted as representative of what occurs *in planta*. It is possible that structural changes in the recombinant M protein upon ADP release, in combination with or separate from the LRR domain affect, prevent its rebinding. Also, rebinding may require the support of other proteins that are not present in the purified protein samples. Indeed genetic and

immunoprecipitation studies have demonstrated that a number of R proteins associate with chaperone and co-chaperone-like proteins (reviewed by (Lukasik and Takken, 2009)). The association of these proteins appears to be critical for the folding and stability of the preactivated form of an R protein (reviewed by (Shen and Schulze-Lefert, 2007)). These interacting proteins may potentially play a role in maintaining the association between ADP and the preactivated form of the R protein. Additionally, they may also help facilitate the nucleotide binding processes and/or nucleotide exchange. To date, no experimental evidence has been published to support this idea, nor is there any data to show that flax R proteins, including M, interact with chaperones. Consequently, at this stage such a statement is highly speculative.

Unfortunately, purification of M $\Delta$ LRR proteins, although attempted (Appendix 4), could not be achieved using the protocol established for M and M $\Delta$ TIR. This, of course, did not enable the identity of the endogenous bound nucleotide to be determined and no conclusion regarding its capabilities in nucleotide rebinding can be made. In light of the results observed for I-2 and Mi-1, one could speculate that the LRR truncated protein would be purified without any bound nucleotide, and that the purified protein would be capable of binding and potentially hydrolysing ATP. Such experiments are ongoing in the laboratory.

In the only other published biochemical study of an NB-ARC-LRR R protein, the TIR truncated N protein was shown to bind ATP, and this binding was required for interaction with its viral effector, p50 (Ueda *et al.*, 2006). The analysis of the TIR truncation of M (M $\Delta$ TIR) presented here, suggests that this protein is ADP bound and in this study evidence for ATP binding could not be demonstrated. The explanation of such discrepancies between our data and that of Ueda *et al.* (2007) are difficult. Lacking from the study of N was any analysis of the integrity of the purified N proteins used in the study. It is possible that a correctly folded LRR was not produced in the *E. coli* expression system, which could explain its inability to negatively regulate the NB-ARC domain. However, without repeating the work of N in the *P. pastoris* system, the reasons for the observed differences can only be speculated upon.

#### 4.3.4 Oligomerisation in the ATP bound state?

Oligomerisation has been demonstrated as a key feature in the activation of a number of STAND proteins, including NALP1, MalT and Apaf-1 (Faustin *et al.*, 2007, Marquet and Richet, 2007, Riedl *et al.*, 2005). Oligomerisation also appears to be important in the resistance of tobacco to TMV, mediated by the tobacco N protein (Mestre and Baulcombe, 2006). In this study, the oligomeric state of the M<sup>D555V</sup> protein was investigated, and whilst giving some hints, failed to provide any definitive data. While potential dimeric forms of the putative autoactive protein were observed when it was purified with ATP and Mg<sup>2+</sup> in the purification buffer, analysis of the double

mutant  $M^{K286L+D555V}$ , known to reduce the capacity of  $M^{D555V}$  to bind ATP, did not drastically change the elution profile over GF. Understandably, analysis has continued along this theme within the laboratory, and recent  $M^{D555V}$  protein purified in the presence of excess ATP, followed by GF in the absence of ATP, has given a profile of more protein at the monomeric peaks and less at the dimeric peak. We are hypothesising that for the  $M^{D555V}$  protein to maintain a dimeric form, it requires ATP binding. To test this, fractions from the GF analysis of  $M^{D555V}$  were quantified for ATP and ADP. These initial investigations have indicated that  $M^{D555V}$  proteins that elute earlier, and are therefore larger than the monomer, have ATP bound, while ADP is observed in the monomeric fractions (P. Sornaraj, personal communication). These results provide initial hints that the ATP bound form of the  $M^{D555V}$  protein enables the formation of higher order structure, more than likely dimers. At present, however, this study is in its preliminary stages and more work is required before one could conclude that oligomerisation is a feature of the activated form of the M protein. Given the literature, it is tempting to draw such conclusions. In the STAND proteins NALP1, MalT and Apaf-1, ATP (or dATP) is necessary for oligomerisation, while a mutation that prevents ATP hydrolysis, and in the case of MalT, trapped it in an ATP bound state, was oligomeric (Faustin et al., 2007, Marquet and Richet, 2007, Riedl et al., 2005).

#### **4.3.5 Conclusion**

The results presented in this chapter are very exciting. They provide insight into the mechanisms of R protein activation. Our data are consistent with one model of R protein activation, although we cannot at this stage, completely test this model without resolving the issues of nucleotide binding and hydrolysis. Attempts to do so are ongoing in the laboratory.

**Chapter 5: A Study of M and AvrM  
Interaction *In Vitro***

## 5.1 Introduction

The cloning of plant R, and pathogen effector, genes has enabled researchers to focus on the important interplay between the proteins these genes encode. Research reported so far has revealed that effector recognition is facilitated through both direct, and indirect, interactions with a corresponding R protein (see 1.7). The current challenge for researchers is to further define the R/effector interaction and determine the functional consequence that such an interaction has on the R protein.

Examples of direct R/effector interactions have been demonstrated using Y2H assays for a number of pathosystems, including the flax-flax rust system. In flax, perhaps the most well characterised interaction is that between the AvrL567 effectors and the L5, L6 and L7 R proteins. Using Y2H assays, in combination with *in planta* assays, the specificity of the gene-for-gene interaction was demonstrated to be a result of the direct interaction between an effector and its corresponding R protein (Dodds et al., 2006). The nature of this direct interaction was further deciphered by modelling of the LRR structure of L5, and using the crystal structure of the AvrL567A effector protein, to position the docking site of the two proteins. This model hypothesised that the interaction and detection of AvrL567A was coordinated by the LRR region of L5 (Wang et al., 2007). Y2H analysis has also shown that the NB-ARC-LRR region of L6 is capable of interacting with AvrL567, but the LRR and NB-ARC regions alone are not (P. Dodds, personal communication). These results suggest that the structural coordination between the NB-ARC and LRR regions is important in generating a protein that can perceive the effector protein. At present, however, there are no published examples of *in vitro* protein-protein interactions involving full-length NB-ARC-LRR R proteins and their corresponding effectors. Given the recombinant R protein system introduced here investigating direct interaction *in vitro* is therefore of particular interest.

Importantly, perception of the effector leads to the activation of the resistance protein and the induction of signalling for the hypersensitive response. This likely requires significant changes to the R protein that enable it to leave its autoinhibitory state and take-on an activated state. In the model of R protein activation the direct or indirect interaction with an effector is predicted to facilitate nucleotide exchange (ADP to ATP) resulting in an activated state (Lukasik and Takken, 2009). Whilst no experimental evidence has been published for effector-mediated R protein nucleotide exchange, *in vitro* experiments involving the STAND proteins MalT and Apaf-1, have demonstrated increased ATPase activity, and nucleotide exchange, in the presence of the elicitors maltotriose and cytochrome c, respectively (Bao et al., 2007, Kim et al., 2005, Marquenet and Richet, 2007, Riedl et al., 2005). Although these experimental systems are far removed from

that of the plant R protein, the growing body of evidence linking R proteins to this broad class of ATP activating (STAND) proteins gives some precedent to models of R protein activation and the design of experiments.

The flax *M* gene provides resistance to strains of flax rust that carry the *AvrM* gene in a gene-for-gene dependent manner (Catanzariti et al., 2006). Direct interaction appears to underlie the interaction between M and AvrM as demonstrated by Y2H experiments (P. Dodds, personal communication). In this chapter, using purified M and AvrM proteins, direct interaction between M and AvrM was tested *in vitro* and the consequence of that interaction on the state of the M protein was investigated.

## 5.2 Results

### 5.2.1 Purification of AvrM, avrM and avrMC\*

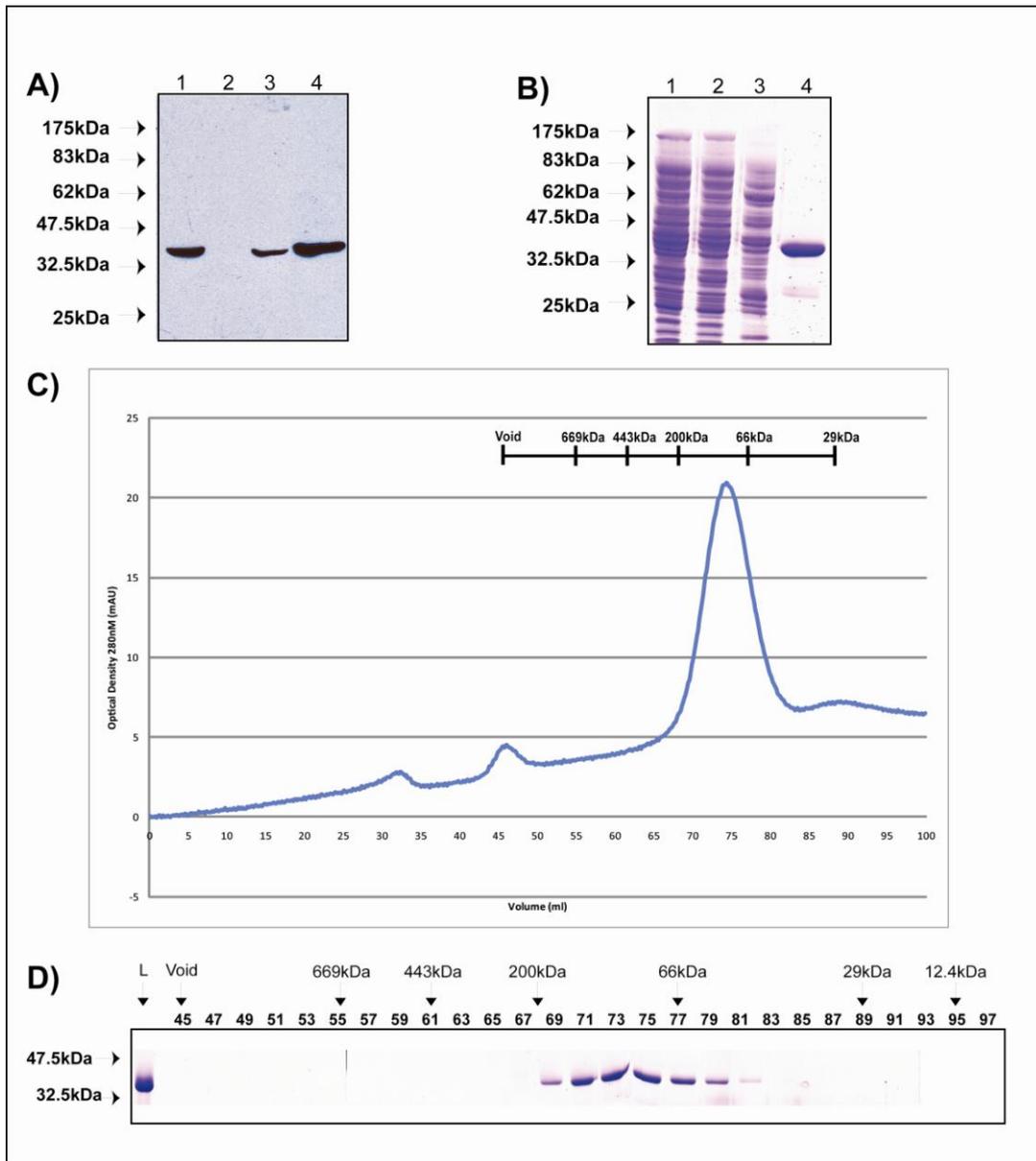
The AvrM locus, described by Catanzariti et al., (2006) contains six genes, AvrMA, AvrMB, AvrMC, AvrMD, AvrME and avrM. AvrMA-D are all recognised by M and elicit a HR, with AvrMA eliciting the strongest response in *Agrobacterium* infiltrated plant tissue. AvrME and avrM, do not elicit a response and are therefore presumed to be undetected by M. A large C-terminal truncation in AvrME precluded its use as a negative control in this study and instead avrM was selected. In order to further examine the interaction between M and AvrM-A, and the non-detected avrM, the products of these genes need to be expressed and purified. In the interests of clarity AvrMA will be defined from this point as AvrM.

For the heterologous expression of AvrM and avrM, their cDNA (minus the first 28 residues, which encode a predicted signal peptide (Appendix 6)) were cloned into the pET expression vector. This vector contains an N-terminal 6x histidine tag, and a thrombin digest site (constructs were obtained from Dr. Peter Dodds, CSIRO Canberra). The vectors containing inserts were transformed into the *E. coli* expression strain, BL21 pLysS. The presence of a histidine tag enabled the purification of the expressed proteins from whole cell lysates using NiA purification (see 2.7.5).

A 38.7 kDa (including histidine tag) AvrM protein could be purified to near homogeneity using NiA purification, with a total yield of approximately 4 mg/L (protein/volume of expressing cells) (Figure 5.1). NiA purified AvrM gave a peak elution point over GF that corresponded to a molecular weight of approximately 120 kDa, suggesting that the protein was purified in an oligomeric state (Figure 5.1C/D). In contrast, only residual amounts of avrM could be produced using the same strategy (Figure 5.2C). This appeared to be the result of poor binding of the avrM protein to the NiA column (Figure 5.2A/B). Protein that was purified was shown by SDS-PAGE and western analysis to be approximately 33 kDa in size (Figure 5.2A/B), which is slightly smaller than the predicted molecular weight (36.2 kDa) of the histidine tagged avrM protein. Analysis of the concentrated protein also suggested that a C-terminal breakdown product was being produced. GF analysis of the concentrated sample of purified protein demonstrated that the larger of the two proteins was aggregated and eluted close to the void volume of the GF column, while the smaller protein could be resolved using GF (Figure 5.2D). The primary amino acid sequence of AvrM and avrM indicate that the avrM gene encodes a C-terminal extension not possessed by AvrM (Appendix 6). Furthermore, the GF analysis of avrM suggested that the C-terminally degraded protein had increased solubility. It was therefore predicted that the C-terminal extension may be the cause of full-length avrM insolubility. Consequently, the construct

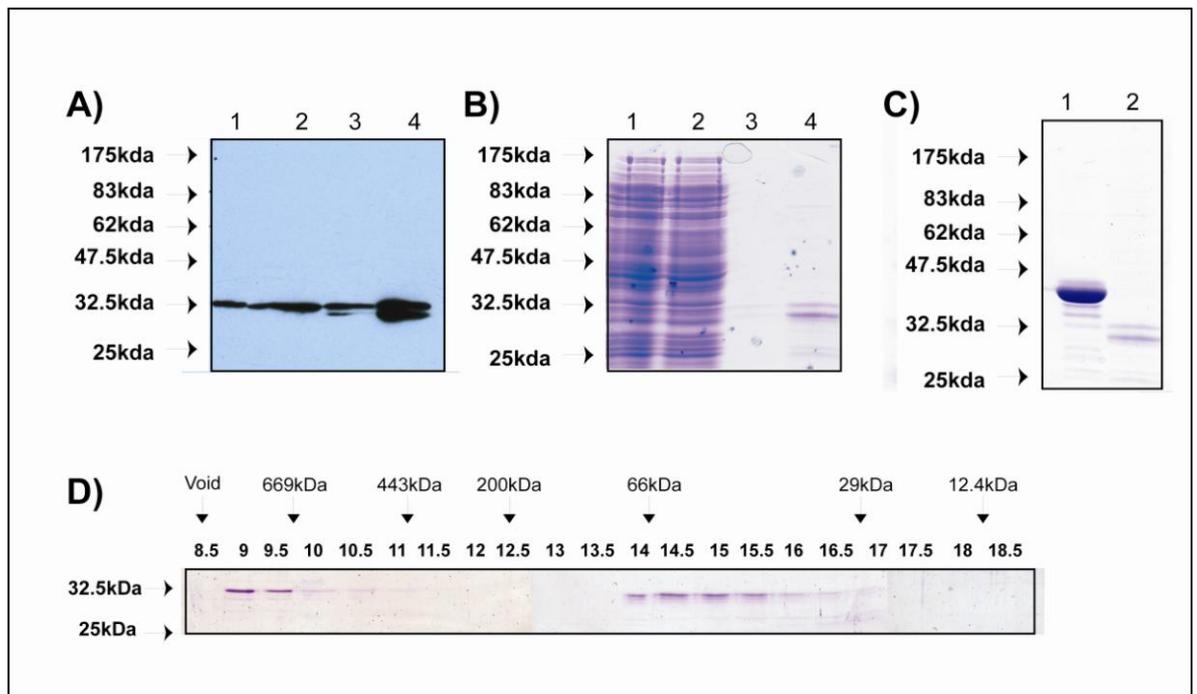
containing *avrM* was reengineered to delete the sequence at the 3' end of the gene that encoded this region (see 2.7.1). The new form of *avrM* was designated *avrMC\**.

The expression and purification of *avrMC\** (32.2 kDa) was performed in the same manner as *AvrM* and *avrM*. Large amounts of soluble *avrMC\** could be purified using this strategy, with post-NiA purification quantities determined to be approximately 20 mg/L (protein/volume of expressed cells) (Figure 5.3). GF analysis of *avrMC\** gave a peak elution that corresponded to a native molecular weight of approximately 72 kDa (Figure 5.3C/D), which as with *AvrM*, suggests this protein exist as a stable oligomer. It was predicted that the *avrMC\** protein would provide a suitable negative control in M and *AvrM* interaction experiments, as this protein has been shown not to elicit a HR when expressed in transgenic tobacco carrying the M gene, nor does it interact with M in Y2H experiments (A. Catanzariti, personal communication). The *avrMC\** protein was, however, not included in all interaction experiments as some of those presented were performed before this protein was expressed and purified.



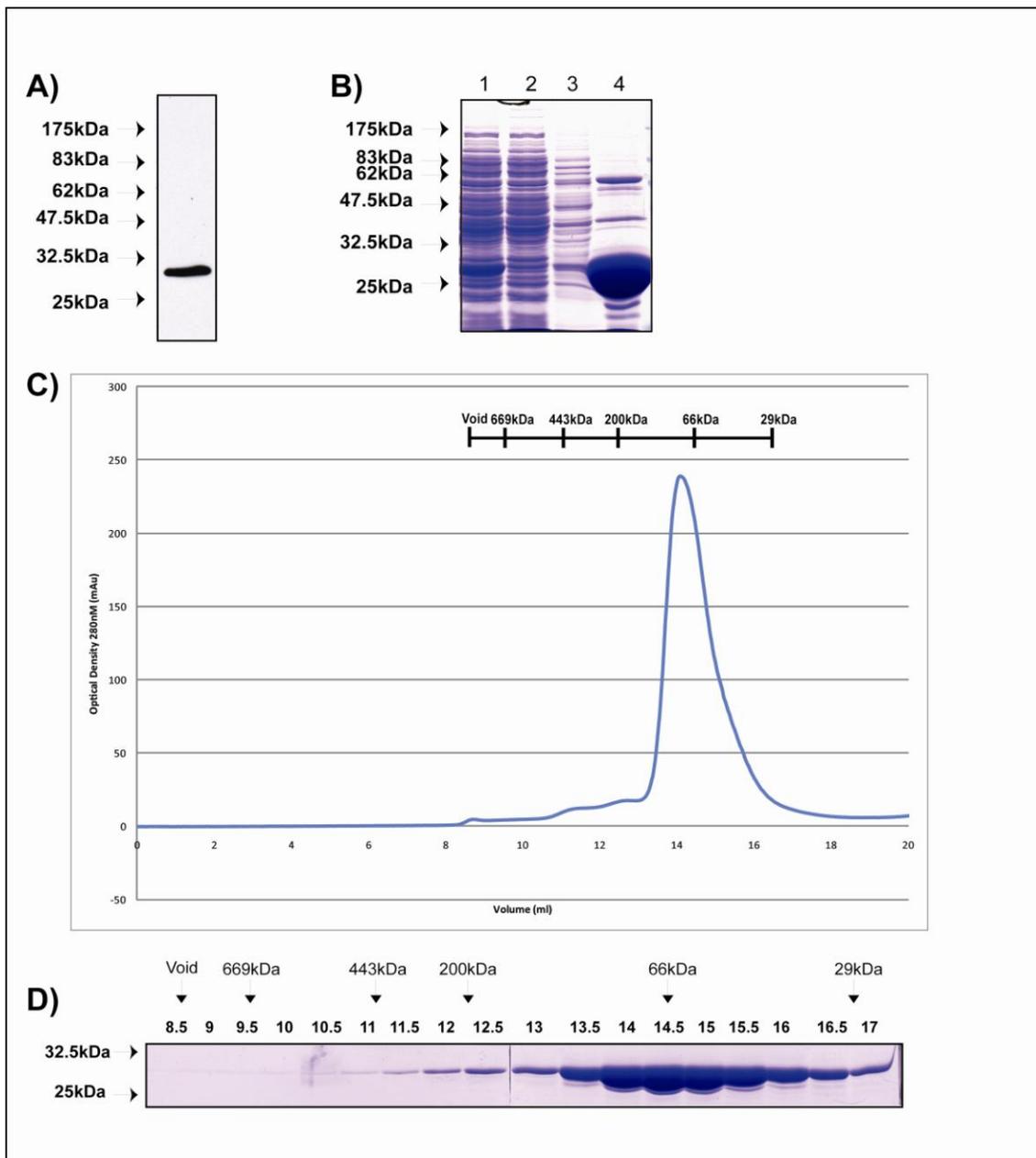
**Figure 5.1: Purification of AvrM**

A) Western blot analysis using an anti-His antibody showing the presence or absence of the AvrM protein during NiA purification. Lane 1, crude lysate; lane 2, NiA flowthrough; lane 3, NiA 65mM imidazole wash; lane 4, concentrated elution. B) Protein separated by SDS-PAGE and Coomassie stained with the same loading as for the western blot analysis. Following NiA chromatography, concentrated protein was further separated over a Superdex 200 PG GF column. C) The resulting UV 280 nm profile was graphed against volume. D) The 2 ml fractions from GF were separated by SDS-PAGE and Coomassie stained. Molecular weights above the gel are positioned where proteins of known molecular weight elute from the same column.



**Figure 5.2: Purification of avrM**

A) Western blot analysis using an anti-His antibody showing the presence or absence of the avrM protein during NiA purification. Lane 1, crude lysate; lane 2, NiA flowthrough; lane 3, NiA 65mM imidazole wash; lane 4, concentrated elution. B) Protein separated by SDS-PAGE and Coomassie stained with the same loading as for the western blot analysis. C) Comparison of purified and concentrated AvrM and avrM protein separated by SDS-PAGE and Coomassie stained. Following NiA chromatography, concentrated avrM protein was further separated over a Superdex 200 HR GF column. D) The 500 µl fractions from GF were separated by SDS-PAGE and Coomassie stained. Molecular weights above the gel are positioned where proteins of known molecular weight elute from the same column.



**Figure 5.3: Purification of avrMC\***

A) Western blot analysis using an anti-His antibody showing avrMC\* in crude lysate. B) Protein separated by SDS-PAGE and Coomassie stained shows the enrichment of avrMC\* protein from NiA purification. Lane 1, crude lysate; lane 2, NiA flowthrough; lane 3, NiA 65mM imidazole wash; lane 4, concentrated elution. Following NiA chromatography, concentrated avrMC\* protein was further separated over a Superdex 200 HR GF column. C) The resulting UV 280 nm profile was graphed against volume. D) The 500  $\mu$ l fractions from GF were separated by SDS-PAGE and Coomassie stained. Molecular weights above the gel are positioned where proteins of known molecular weight elute from the same column.

## 5.2.2 Co-immunoprecipitation experiments using purified recombinant proteins support the hypothesis of a direct protein-protein interaction between M and AvrM

The expression and purification of recombinant M and AvrM protein enabled the investigation of a direct protein-protein interaction *in vitro*. Co-immunoprecipitation (Co-IP) is a well recognised technique for identifying protein-protein interactions, and was used here to investigate if such a direct interaction between M and AvrM proteins could be captured. For interaction experiments, the effector proteins (AvrM and avrMC\*) were added in 2-5x molar excess to the M protein. M protein was targeted in the immunoprecipitation experiments using the anti-M antibody, which is a peptide raised antibody generated in a rabbit and is directed towards the N-terminus of the M protein (Appendix 3, Figure 3.1) (Schmidt et al., 2007b). If M and AvrM form a direct interaction, they would therefore generate a complex, and, if stable enough, the AvrM protein should only be detected in the bound IP fraction when the M protein is present in the reaction.

Initial Co-IP experiments were performed using an agarose-based protein G resin (see 2.9.1.1) which was used to bind the anti-M antibody in order to capture any complexes. Fractions bound to the antiM antibody were analysed by SDS-PAGE, followed by Coomassie or western analysis. An anti-His mouse monoclonal antibody was used as the probe for western analysis and could detect both M and AvrM histidine tagged proteins (Figure 5.4A, lane 1 and 2). M protein could be successfully immunoprecipitated using the anti-M antibody (Figure 5.4A/B, lane 4 and 5) and an immuno-reactive band believed to be the AvrM protein was also observed in the same bound fractions. However, in the negative control involving the incubation of the anti-M antibody with AvrM alone, the same size immuno-reactive band was observed in bound fractions (Figure 5.4A/B, lane 3). This demonstrated either cross-reactivity between anti-M and AvrM, or AvrM is interacting directly with the protein G resin. Regardless of the reason, the presence of AvrM in these fractions did not allow it to be concluded that its presence was dependent on the M protein. It is possible that higher washing stringencies were required to appropriately analyse the potential interaction between M and AvrM using this method. Unfortunately, the protein G agarose system being used relied on a batch binding/washing strategy, which is, at times, not conducive to stringent washing.

To further investigate protein-protein interaction using Co-IP a new generation system was employed that utilised metallic-labelled beads, and a column based system, as opposed to a batch-based approach described above (see 2.9.1.2). In the repeat experiments with the new system, avrMC\*, which is not detected by the M protein in Y2H analysis, was added to the experiment as an additional negative control. Given the results presented in Chapter 4, ATP and Mg<sup>2+</sup> were supplemented into all binding and wash buffers in subsequent experiments. Using this method, in the AvrM only control an anti-His immuno-reactive band that corresponds to AvrM was

detected in the unbound but not the bound fraction (Figure 5.5A, lanes 1 and 2), however, a faint Coomassie stainable band of this size was observed (Figure 5.5B, lane 2). A similar result was observed for *avrMC\** (Figure 5.5A/B, lane 3 and 4), suggesting that some of the effector proteins may be associating with the resin/columns, however, in the western blot they were not detected.

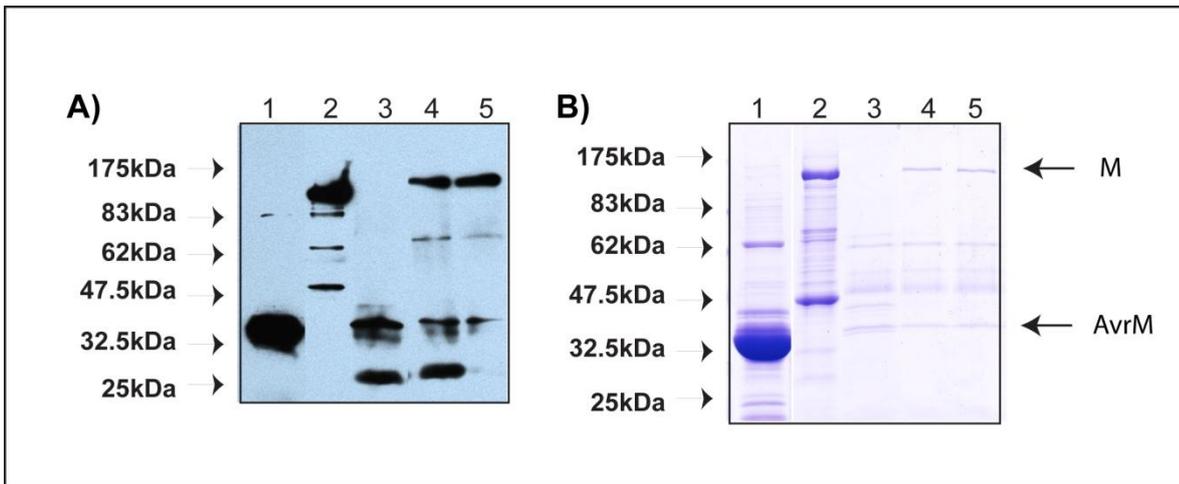
In the Co-IP experiments with purified M protein, a clearly distinguishable Coomassie stained band at the expected size of AvrM protein was observed in the bound fraction (Figure 5.5A/B, lane 6). Western blot analysis with an anti-His antibody confirmed that this protein has a histidine tag. In the Co-IP of M with *avrMC\**, no protein of the *avrMC\** size was observed in the bound fraction (Figure 5.5A/B, lane 7). These results indicate that AvrM, but not *avrMC\**, co-purified with the M protein in the bound Co-IP fraction, suggesting that AvrM and not *avrMC\** is interacting with the M protein.

This experiment was repeated with an additional control, the mutant M protein,  $M^{K286L}$ . A lysine to leucine mutation within the P-loop of M prevents the interaction between M and AvrM protein in Y2H experiments (P. Dodds, personal communication). In a repeat experiment the same Co-IP method was used, with the addition of single IP's of M and  $M^{K286L}$  (Figure 5.6). The results indicate that both M and  $M^{K286L}$  protein can be immunoprecipitated with the anti-M antibody (Figure 5.6A/B, lane 3 and 6, respectively). A number of immuno-reactive bands between 47.5kDa and 83kDa were observed in these fractions, suggesting the IP of some C-terminal breakdown products of the full-length M protein. Importantly, the control AvrM-IP experiment did not detect an immuno-reactive band or a Coomassie stainable band in the bound fraction at the size of AvrM (38.7kDa). An extra wash was performed in this experiment compared to the experiment outline in Figure 5.5 and it is likely that this removed any non-specific binding between AvrM and the antibody or resin. Western and Coomassie analysis of the Co-IP experiments (Figure 5.6C/D) demonstrated that in the M and AvrM experiment, an immuno-reactive and Coomassie stained band of the appropriate size of AvrM was detected (Figure 5.6C/D, lane 3). As demonstrated previously, the *avrMC\** protein was not observed in the bound fraction when Co-immunoprecipitated with M (Figure 5.6C/D, lane 6). In the Co-IP experiment of  $M^{K286L}$  and AvrM, a band corresponding to the AvrM protein was not detectable by Coomassie or western analysis (Figure 5.6C/D, lane 9). Taken together the experiments outlined in Figures 5.5 and 5.6, suggest that the AvrM protein is interacting with the M protein, and that this association can be captured using a Co-IP approach. Importantly, the results obtained thus far are entirely consistent with those gained from Y2H analysis, which demonstrate that for interaction with AvrM, the M protein requires a functional NB pocket and we would speculate further that for interaction, the M protein needs to have an ADP molecule bound with its NB pocket.

Western blot analysis was also performed with anti-AvrM-specific antibodies (supplied by Pam Gan, Australian National University (ANU), Canberra), however, these antibodies were raised in a rabbit and the use of rabbit IgG specific secondary antibodies caused cross-reactivity with the anti-M antibodies used in the initial immunoprecipitation. Consequently, these western results were uninterpretable (data not shown).

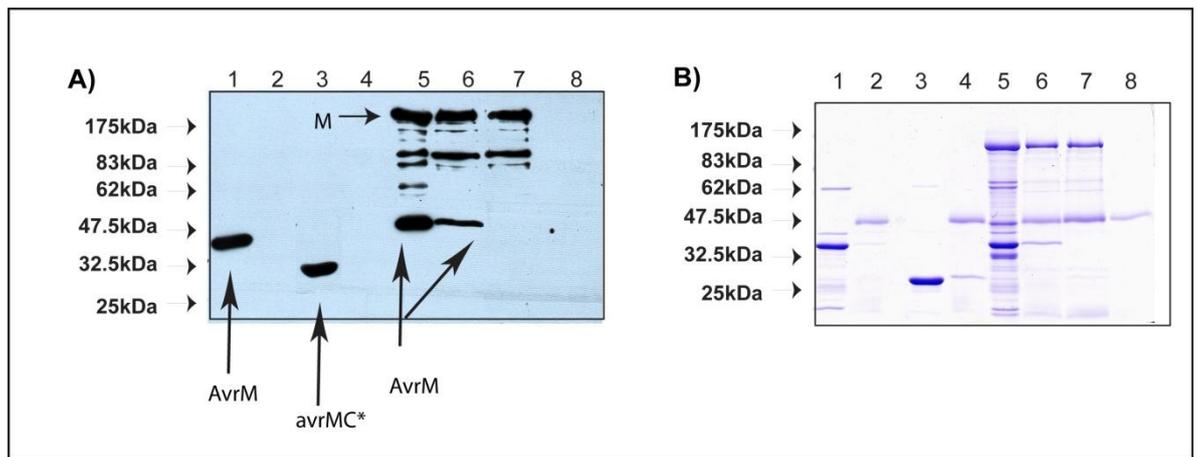
A portion of the M and AvrM proteins incubated for Co-IP experiments were also analysed by GF. The elution profiles of the proteins were not notably different from their individual profiles, suggesting that the interaction is not stable enough to be captured using the technique of GF (data not shown).

Consistent with the working model of R protein activation, the outcome of the M and AvrM interaction presumably converts M from an inactive to an active form, which then relays the signal(s) for cell death within the plant cell. Results outlined in Chapter 4 suggest that the inactive form of M is ADP bound, while a mutant predicted to cause autoactivity, has a preference for binding ATP within its NB-ARC domain. The literature speculates that in the presence of an effector, R proteins can exchange a nucleotide in their NB pocket and/or increase their rate of ATP hydrolysis. In light of these results, and hypotheses, the focus of this interaction study turned towards attempting to show if any such changes in ADP or ATP binding could be detected in the M protein following M/AvrM interaction.



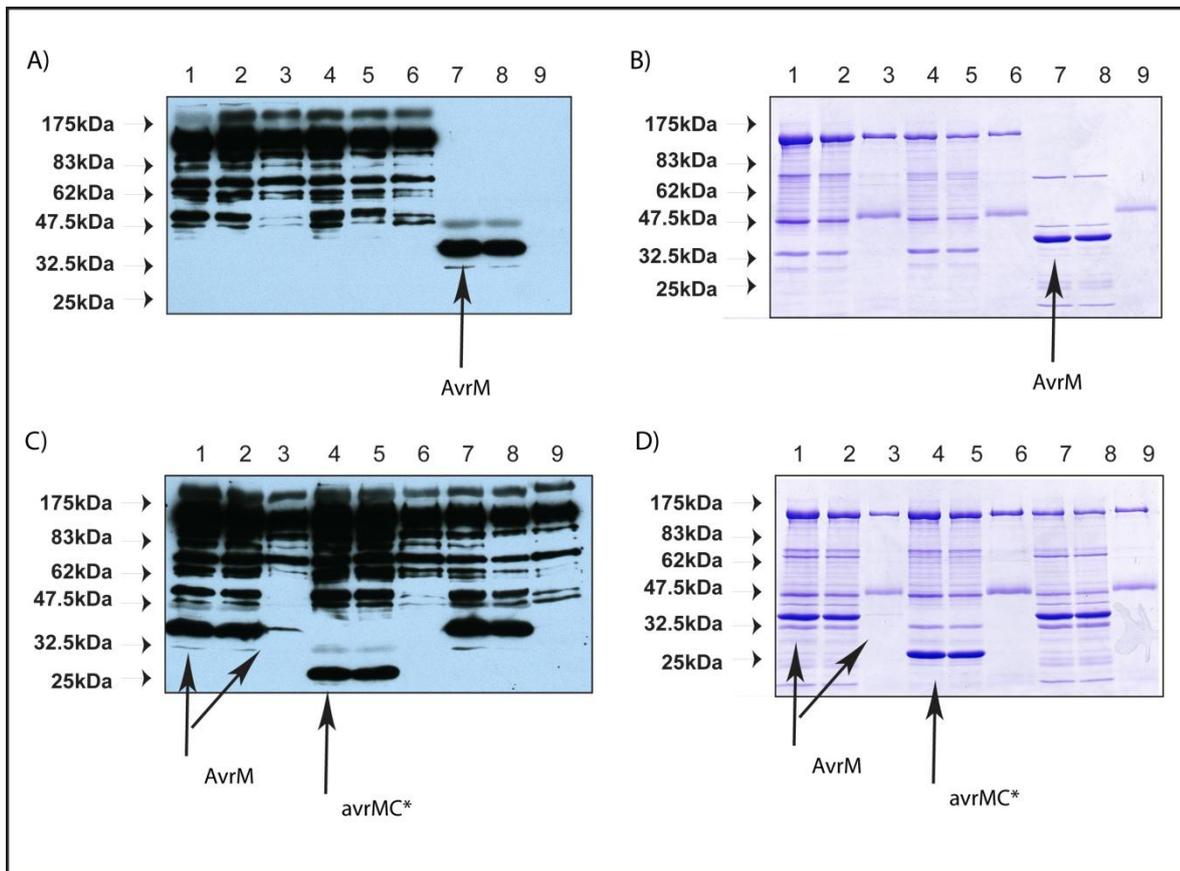
**Figure 5.4: Co-immunoprecipitation of M and AvrM**

Co-immunoprecipitation (Co-IP) was performed with anti-M antibodies. A) Western blot analysis using a mouse monoclonal anti-His antibody. Lane 1: AvrM protein control; lane 2: M protein control; lane 3: AvrM IP control, bound fraction; lane 4: M and AvrM Co-IP, bound fraction; lane 5: M and AvrM Co-IP with ATP/ADP in buffer, bound fraction. B) Same as (A) but proteins separated by SDS-PAGE were Coomassie stained. Arrows identify bands that correspond to M and AvrM proteins



**Figure 5.5: Co-immunoprecipitation of M, AvrM and avrMC\***

Immunoprecipitation (IP) and Co-IP was performed with anti-M antibodies. A) Western blot analysis of different IP and Co-IP combinations using a mouse monoclonal anti-His antibody. Lane 1-2, AvrM IP: unbound and bound respectively; lane 3-4, avrMC\* IP: unbound and bound; lane 5-6, M and AvrM Co-IP: unbound and bound; lane 7, M and avrMC\* Co-IP: bound; lane 8, anti-M antibody only: bound. B) Same as (A) but proteins separated by SDS-PAGE were Coomassie stained.



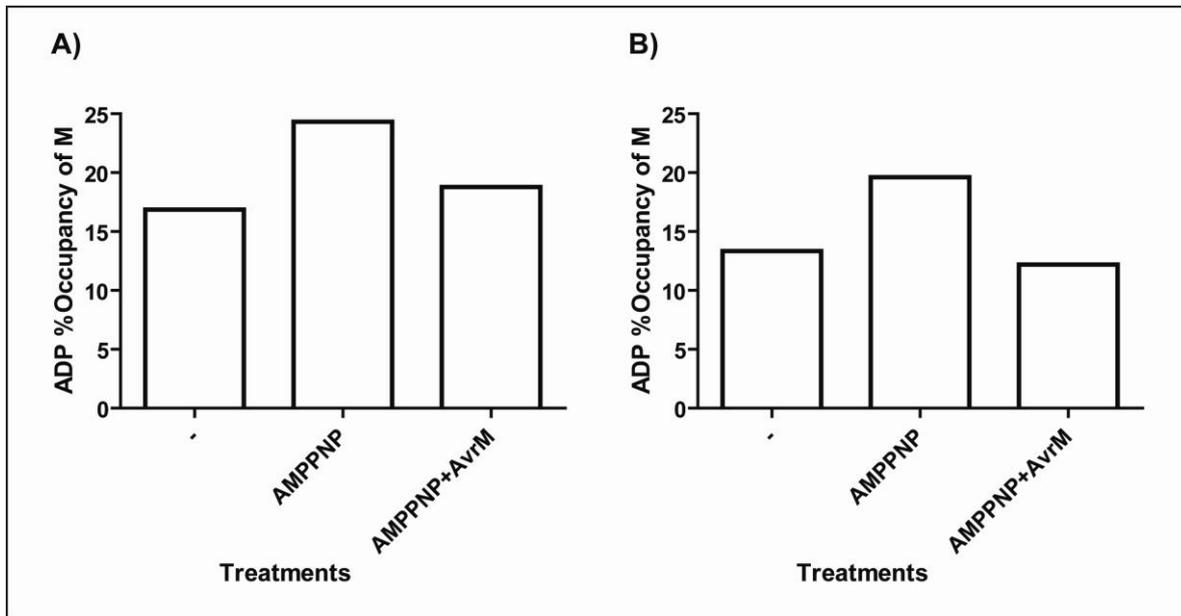
**Figure 5.6: Co-immunoprecipitation of M, MK<sup>286L</sup>, AvrM and avrMC\***

Immunoprecipitation (IP) and Co-IP was performed with anti-M antibodies. A) Western blot analysis of different IP and Co-IP combinations using a mouse monoclonal anti-His antibody. Lane 1-3, M IP: loaded, unbound and bound; lane 4-6, MK<sup>296L</sup> IP: loaded, unbound, bound; lane 7-9, AvrM IP: loaded, unbound, bound. B) Same as (A) but proteins separated by SDS-PAGE were Coomassie stained. C) Western blot analysis as for (A). Lane 1-3, M and AvrM Co-IP: loaded, unbound and bound, respectively; lane 4-6, M and avrMC\* Co-IP: loaded, unbound, bound; lane 7-9, MK<sup>296L</sup> and AvrM Co-IP: loaded, unbound, bound. (D) Same as (C) but proteins separated by SDS-PAGE were Coomassie stained.

### 5.2.3 *In vitro* experiments fail to identify nucleotide exchange in M induced by AvrM interaction

Given the availability of purified recombinant forms of AvrM and M, and that these proteins interact *in vitro*, experiments now focussed on whether AvrM could stimulate a nucleotide exchange event in the M protein *in vitro*. Using purified forms of MalT, and the luciferase ADP/ATP quantification assay, investigators successfully demonstrated that the purified ADP-bound form of MalT could exchange ADP for the non-hydrolysable form of ATP (AMPPNP) when incubated with its elicitor, maltotriose (Marquet and Richet, 2007). In light of this work, and the results of Chapter 4 which had demonstrated an M·ADP association, a similar experiment to that designed for MalT, was designed for M. The experiment involved AvrM as the potential inducer of nucleotide exchange, and the non-hydrolysable analogue of ATP, AMPPNP, as a potential exchange factor. AMPPNP was included in the experiment as it does not participate in the luciferase quantification assay. If a reduction in ADP levels in M were observed when M, AvrM and AMPPNP were incubated together, but not in the appropriate controls, this would signify that exchange of ADP for AMPPNP in the M protein was induced by AvrM.

The results of these experiments are shown in Figure 5.7A/B. It was found that ADP occupancy levels (average 35%) dropped independently of the addition of other compounds when the M protein was incubated alone for 15 minutes at 25°C. This dissociation of ADP from the M protein was expected given the results presented in Chapter 4, and these values were used as a baseline for comparison with the other treatments. What was observed when a non-hydrolysable analogue of ATP, AMPPNP, was incubated with M, was in direct contrast to what was expected. The amount of ADP associated with M protein incubated with AMPPNP was higher compared to the untreated sample, suggesting that the presence of AMPPNP reduces ADP loss from the M protein. The amount of ADP present in the M protein, however, remained unchanged when M, AvrM and AMPPNP were incubated. These same experiments were also performed with the addition of another non-hydrolysable ATP analogue, ATP- $\gamma$ -S, instead of AMPPNP. ATP- $\gamma$ -S was still capable of participating in the luciferase quantification assay, so in these experiments ATP and ADP levels were quantified. The results obtained from these experiments, however, failed to demonstrate any consistent, or repeatable, data that would suggest that AvrM is promoting nucleotide exchange in the NB pocket of the M protein (data not shown).



**Figure 5.7: Nucleotide exchange assay**

NiA purified M protein underwent a series of treatments and was incubated at 25°C for 15 minutes; no treatment (-), 0.5 $\mu$ M AMPPNP added (AMPPNP), 0.5 $\mu$ M AMPPNP and AvrM added (AMPPNP and AvrM). After the removal of unbound nucleotide, ADP quantification was performed. A) and (B) represent the experimental outcomes of two independent M protein purifications.

### 5.3 Discussion

The results presented in this chapter detail the expression and purification of the flax effector protein, AvrM, and a truncated mutation of the virulence protein *avrM*, designated *avrMC\**. Access to purified AvrM effector proteins, combined with the ability to produce functional recombinant M protein (Chapter 3), enabled *in vitro* protein-protein interaction studies using Co-IP techniques to be conducted. The results of the Co-IP study support the theory first proposed by Y2H analysis, that M and AvrM proteins interact directly (Dodds, personal communication). While a number of biochemical assays were used in an attempt to identify the effect that this interaction may have on the M protein, the results of these investigations remain unclear. Nevertheless, this study has stimulated research to further investigate the interaction between a rust effector, and a flax R protein, and the molecular consequence(s) of this interaction.

#### 5.3.1 Purification of flax rust effector proteins

The expression and purification of the AvrM protein in the heterologous *E. coli* expression system, achieved high yields of relatively pure protein after NiA chromatography. The ability to purify soluble AvrM protein stimulated a number of other ongoing research studies outside the laboratory, but within the collaborative group. Polyclonal antibodies against AvrM protein were generated and are currently being used to assist in AvrM localisation studies (P. Gan, personal communication). AvrM is also the current focus of crystallisation trials aimed at determining its three dimensional structure (T. Ve, personal communication).

Interestingly, the elution profile of AvrM from GF suggests it may form an oligomeric complex that resembles a homotrimer or homodimer. Further investigation of this phenomenon in the crystallisation project has shown that purified AvrM protein does indeed form a homodimer (T. Ve, personal communication). This was demonstrated using more accurate size estimation technology that included multi-angle light laser scattering techniques. The importance of this dimer formation in its activity and/or interaction with the M protein is currently under investigation.

The C-terminal extension region possessed by *avrM* makes the protein insoluble when expressed in *E. coli*. Interestingly, this protein can not be detected in yeast in Y2H experiments (P. Dodds, personal communication), suggesting that it may be unstable. Removal of this extension yielded highly expressed and soluble protein in *E. coli*. The *avrMC\** protein not only provides a useful negative control in interaction experiments, it is now a future target of crystallisation studies, and will assist in defining the areas of the AvrM effector protein that are recognised and interact with the M protein.

### 5.3.2 Co-IP studies support a direct interaction between M and AvrM

The Co-IP experiments suggest that an interaction between AvrM and M can be captured using this approach. The antibody used in the western blot analysis was specific against histidine tagged proteins, and therefore detected both M and AvrM/avrMC\* specific bands on the same gel. An immuno-reactive band present at the expected molecular weight of AvrM, in the M and AvrM Co-IP bound lanes (Figures 5.5 and 5.6) indicate that AvrM is interacting with M. As both the proteins have been purified, the interaction between the two is therefore most likely a direct one. Importantly, an immuno-reactive band at the molecular weight of avrMC\* is not present in the Co-IP experiment of M and avrMC\*. These results are in agreement with Y2H analysis which demonstrated that avrMC\* does not interact with M (A. Catanzariti, personal communication). Also, an immuno-reactive band at the molecular weight of AvrM is not present in the M<sup>K286L</sup> and AvrM Co-IP experiment. The M<sup>K286L</sup> protein, therefore, failed to form an association with AvrM that could be captured by Co-IP, which is also in agreement with Y2H analysis of this mutant (P. Dodds, personal communication). The inability to observe an interaction between AvrM and M<sup>K286L</sup> protein suggests that a functional P-loop is required to facilitate the M/AvrM interaction. *In vitro* studies in Chapter 4 demonstrated that the integrity of the P-loop lysine is critical in maintaining both ADP and ATP association in M proteins. It is therefore probable that either an ADP bound state and/or ATP bound state are required for interaction.

It appeared in the Co-IP experiments that lower quantities of AvrM protein compared to M protein were captured on the gels, possibly indicating that during washing the interaction between the two proteins was disrupted. This could suggest that the interaction between M and AvrM is transient and rather unstable in nature. This is consistent with the fact that M and AvrM interaction could not be captured using GF. Another possible explanation for this result is that only a percentage of the M protein is in a form, possibly due to its nucleotide bound state, that is capable of interacting with AvrM. We know from data presented in Chapter 4 that following purification only ~38% of M protein is bound with ADP. It is conceivable from the Co-IP data, that it is only this ADP bound form of M that can directly interact with AvrM. Undoubtedly, more work is required to decipher the requirements for M/AvrM interaction and the nature of the interaction.

### 5.3.3 The effect of the effector

Protein oligomerisation, increased levels of hydrolysis activity, and the induction of nucleotide exchange, are all events that have all been demonstrated experimentally to occur as a consequence of an interaction between a STAND protein and its corresponding elicitor protein (see 1.8). In experiments aimed at deciphering the potential effect that the AvrM interaction may have on the M protein, nucleotide exchange experiments were conducted. The results of these

experiments were largely, however, inconclusive. As discussed in Chapter 4, the inability of purified protein to reform an association with a nucleotide, may be an issue in the design and outcome of these experiments. It is possible that other proteins are necessary to facilitate the *in planta* consequence of an interaction between M protein and its effector. Although the *in vitro* study of flax R and flax effector proteins remains a primary objective for this laboratory we are not excluding plans to investigate the roles of other plant proteins in the biochemical study of these proteins. Some of these future experiments will be outlined in Chapter 6.



## **Chapter 6: Conclusions and Future Work**

## 6.1 Overview

Protein biochemistry is essential for the interpretation of a protein's molecular function. In a biochemical study that utilises recombinant protein derived from a heterologous host, the expressed and purified protein should ideally resemble the native protein's conformation and structure. Only when these requirements are achieved can the subsequent results obtained from *in vitro* experiments be meaningfully interpreted as indicative of the protein's native function. Unfortunately, adequately addressing this requirement can often be difficult when little is known about the molecular mechanisms by which the protein of interest operates.

When we began our biochemical investigation of the flax R proteins, M and L6, we were confronted with this somewhat circular argument. The studies of I-2, Mi-1 and N (Tameling et al., 2002, Tameling et al., 2006, Ueda et al., 2006), and the assignment of R proteins into the STAND classification (Leipe et al., 2004), assisted in the experimental design of the research presented here. The results demonstrate the ability to produce and purify recombinant M and L6 proteins, and that these proteins demonstrate at least some of the functions predicted to be involved in the initiation of a plant's immune response.

To summarise the presented chapters: Chapter 3 demonstrates that near full-length versions of the flax R proteins, M and L6, can be expressed and purified from *P. pastoris* as soluble, predominantly monomeric proteins. These proteins remained soluble during purification, however, exposure to high concentrations of imidazole, cause the protein to aggregate. Functional analysis of protein, prepared before the negative implications of imidazole were identified, generated no repeatable/interpretable data. This certainly highlights the importance of obtaining protein that is stable and soluble prior to biochemical analysis. With the necessary changes made to the purification procedure, functional study of the M and L6 proteins was possible. The results obtained in Chapters 4 and 5 provide further evidence supporting a model of R protein activation involving the exchange of ADP for ATP and the direct interaction between M and AvrM. However, at present, the necessary *in planta* experiments for the mutation analysis (Chapter 4), and the full repertoire of control experiments for the Co-IP experiments (Chapter 5) are yet to be completed.

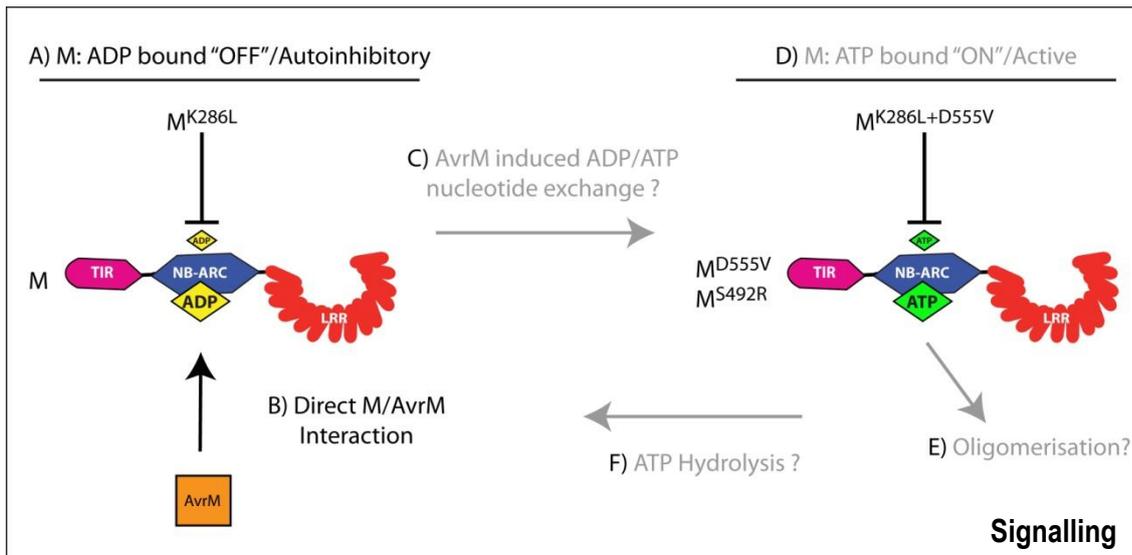
The methods established during this thesis and outcomes achieved signify a major advance in our understanding of R protein biochemistry. Continued research utilising the systems developed here will therefore enable greater insight into the molecular function of an R protein. This will undoubtedly be of interest and importance to the plant disease resistance protein field, and to the study of innate immunity in general.

## 6.2 Model of M Protein Activation

The data obtained during the course of this thesis has been summarised in a model of M protein activation (Figure 6.1). This model supports the results of the biochemical studies involving the tomato R proteins I-2 and Mi-1 (Takken et al., 2006, Tameling et al., 2002, Tameling et al., 2006). The current gaps in our understanding are highlighted and additional experiments that could be used to further elucidate and improve this working model are detailed below (see 6.3).

The model of M protein activation, Figure 6.1: M exists in an autoinhibited state as a monomer with ADP bound. In the presence of AvrM, a direct interaction between M and AvrM takes place, as demonstrated by Y2H (P. Dodds, personal communication) and Co-IP protein-protein interaction experiments presented here (Chapter 5). Upon this interaction we speculate that the M protein becomes activated. The activation of M stimulated by AvrM perception presumably induces conformation changes in M that facilitate the initiation of immune signalling pathways in flax. We predict that M is ATP bound in its activated state, and arrive at this conclusion from the analysis of the putative autoactive proteins M<sup>D555V</sup> and M<sup>S492R</sup>. Quantification analysis demonstrated that these proteins bind considerably more ATP than the wildtype M protein. In light of this, it is anticipated that interaction between the AvrM effector protein, and the M protein, would induce nucleotide exchange thereby converting the protein from the off, ADP bound, state to the on, ATP bound, state. Experiments presented here that aimed to demonstrate nucleotide exchange *in vitro* were, however, inconclusive. The role of ATP hydrolysis, if any, in the activation or resetting process for M is yet to be determined, as is the oligomerisation state of the activated M protein.

The model of M activation is similar to that generalised for the activation of R proteins in the most recent interpretations (Lukasik and Takken, 2009, Takken and Tameling, 2009) although interaction between domains and subdomains (intramolecular) in the M model are not described. Much of the intramolecular interaction work published to date originates from studies of CC-NB-ARC-LRR proteins, in particularly Rx (Bendahmane et al., 2002, Moffett et al., 2002, Rairdan et al., 2008, Rairdan and Moffett, 2006) and at this point there is very little insight provided for the intramolecular interactions involved in TIR-NB-ARC-LRR class. Intramolecular interactions are undoubtedly involved in the regulation and activation of TIR containing proteins, potentially in the same or similar fashion as that described for the CC containing proteins. The involvement of intramolecular interactions during R protein function is discussed in detail in the following review (Lukasik and Takken, 2009).



**Figure 6.1: A working model of M activation**

A) M exists in an autoinhibited conformation with ADP bound. ADP binding is dependent on a functional P-loop. B) Perception of the AvrM effector protein by the M protein occurs inside the flax cell and is facilitated by a direct interaction between the two proteins. C) Recognition of AvrM activates the M protein and it is anticipated that this stimulates exchange of ADP to ATP, however, no experimental data is available to support this. D) The activated M protein is likely to be ATP bound as demonstrated by the analysis of the putative autoactive mutations M<sup>D555V</sup> and M<sup>S492R</sup>. Confirmation of their autoactive phenotypes *in planta* is required before this can be concluded. ATP binding is also dependent on a functional P-loop. Further work is required before the predicted roles of oligomerisation (E) and hydrolysis (F) can be confirmed, however, the data obtained in this study does not warrant their exclusion from the model.

NB: The protein domain labels on M are abbreviated as described in text. Grey text colouring signifies further work required to confirm the role predicted in the model. Domain interactions are likely to occur in both the autoinhibited and active protein states, as demonstrated by studies of Rx (reviewed by (Lukasik and Takken, 2009)), however, as domain interaction work has been predominantly confined to studies of CC-NB-ARC-LRR proteins such interactions are not included in the model.

### **6.3 A Model Stimulates Experimental Questions for Future Research**

The data generated from our biochemical studies of recombinant M protein has provided a clue into the potential activation of the M protein at the molecular level. We wish now to explore further, utilising this system, the molecular function of an R protein. We are currently planning numerous experiments, the details of which are discussed below.

#### **6.3.1 Determining the *in planta* phenotypes**

As highlighted throughout this thesis, the most essential information we now require to support our model of the autoinhibition and activation of M, is the *in planta* phenotypes of the putative gain- and loss-of-function mutations. At the time that this thesis was submitted the *in planta* work was not finalised, however, the testing of the various mutations *in planta* was imminent. The outcome of these experiments will either directly support, or contest, the model of M activation that has been presented (Figure 6.1).

#### **6.3.2 Is ATP hydrolysis and/or oligomerisation involved in M protein function?**

Unfortunately, from this study we have been unable to confirm or disprove the involvement of ATP hydrolysis and protein oligomerisation in the activation process of M. M has all the hallmark characteristics believed to be required of a protein capable of hydrolysing ATP, and the literature of other R proteins is indicative of this function (Tameling et al., 2002, Ueda et al., 2006). The studies of tomato R proteins I-2 and Mi-1 demonstrated that while ATP hydrolysis was a functional property of their NB-ARC domains, hydrolysis itself was not a likely requirement for resistance (Tameling et al., 2006). This prompted the suggestion that ATP hydrolysis may be involved in resetting the R protein ready for refiring of the activation step. In studies of the STAND protein, MalT, such a role of hydrolysis has been demonstrated. Hydrolysis of ATP by MalT while demonstrated was not critical for its activation as a transcription factor but was required for regulatory control of transcription activity. Indeed a recent article that analysed literature pertaining to biochemical and structural studies of STAND proteins supports the recycling/reset function that is provided by ATP hydrolysis (Danot et al., 2009). It is, however, necessary to highlight when discussing ATP hydrolysis by proteins within the STAND classification, that not all STAND proteins studied have ATPase activity. For example, CED-4 has not been demonstrated to have ATPase activity and is activated when the negative regulation of CED-9 is overcome through the interaction of an activating protein called EGL1 (Yan et al., 2005). When addressing the role of hydrolysis in R protein activation, therefore, it appears that we should also focus on determining if there is a need for R protein recycling during activation.

Currently, the biggest challenge faced in determining if the M protein can hydrolyse ATP is to see first, if it can bind ATP, and if it does then, to remove the source of contaminating ATPase activity in the protein preparations. To achieve this we are currently conducting binding experiments with purified M protein and  $P^{32}$  labelled ATP. If binding of ATP can be demonstrated, then we will incorporate a further purification step that involves the use of the anti-M antibody to immunoprecipitate recombinant M protein. During the immunoprecipitation experiments presented in Chapter 5, the use of the anti-M antibody to purify M greatly enhanced protein purity. It is hoped that such a targeted purification step will remove the contaminant ATPase from R protein preparations, and the subsequent background activity that has plagued our experiments to date. Interestingly, evidence has emerged from the study of STAND proteins that suggests slow turnover rates of ATP are analogous with the ATPase activities of STAND proteins when compared to other ATPase proteins (Danot et al., 2009). This finding highlights the need for the removal of any background activity to observe R protein-related ATP hydrolysis, as the assays may require long incubation times. Also, any future ATPase study will need to include the M $\Delta$ LRR protein, as it resembles most closely the LRR truncated forms of I-2 and Mi-1 for which published data exists (Tameling et al., 2002, Tameling et al., 2006). To achieve many of these outcomes, however, alternative purification schemes will need to be developed.

Like ATPase activity, oligomerisation is another molecular function that appears to be linked with the activation of STAND proteins. Biochemical studies have demonstrated its involvement in the activation of NALP1, Apaf-1, MalT and CED-4 (Faustin et al., 2007, Kim et al., 2005, Marquet and Richet, 2007, Yan et al., 2005). It also appears to be a feature of the activation of the tobacco N protein, that controls resistance to TMV (Mestre and Baulcombe, 2006). Analysis of the GF profiles of M<sup>D555V</sup> presented in this thesis, and in work that has continued in the laboratory, has given some indication that this protein may form some higher order structure, most likely a dimer. As oligomerisation is likely to require ATP binding, we are currently investigating methods to trap the M<sup>D555V</sup> protein in an ATP bound state. We are also investigating whether a TIR truncated M<sup>D555V</sup> protein would have any tendency to oligomerise. This may provide clues into the role of the TIR domain in oligomerisation, considering that for tobacco N, the TIR domain was proposed to facilitate oligomer formation (Mestre and Baulcombe, 2006).

In summary, further work is required to resolve the current questions regarding ATP hydrolysis and oligomerisation during M activation. The answers to these questions will add significantly to the model presented in Figure 6.1.

### 6.3.3 Interaction and recognition

The recognition of a pathogenic effector protein by an R protein is arguably one of the most important mechanisms to decipher in plant's disease resistance, as it is fundamental to a plants ability to detect a pathogen. Data presented in this thesis support the conclusion obtained from Y2H analysis that AvrM and M interact directly. We believe that the Co-IP system used here, and the inclusion of other protein-protein interaction methods such as biacore, which can be used to determine the kinetics of an interaction, will lend further support for direct interaction between flax rust effectors and their corresponding flax R proteins. There are several questions that we are currently addressing using the co-immunoprecipitation technique: What are the molecular requirements for M/AvrM interaction? What areas/domains of the M protein are required for AvrM interaction, with a focus towards defining residues involved in specificity? How does the M protein convert from an autoinhibited to an activated state, as a consequence of the interaction with AvrM?

A) *What are the molecular requirements for M/AvrM interaction?* It was demonstrated here that a mutation in the invariant P-loop lysine was shown to prevent M/AvrM interaction, which had also previously been shown in Y2H interactions studies (P. Dodds, personal communication). This same mutation also prevents ADP binding in the M protein and ATP binding in the M<sup>D555V</sup> protein. It is therefore highly likely that the binding of a nucleotide is required for the M protein to take on a conformation that is conducive for AvrM recognition and interaction. Our initial interpretation of this result suggested that ADP-bound M creates the conformation necessary for interaction. However, concurrent studies within the laboratory have demonstrated that mutations within the conserved motifs of the NB-ARC domain that are shown to prevent interaction in Y2H experiments (P. Dodds, personal communication) are still capable of associating with ADP (P. Sornaraj, personal communication). Whilst these experiments require further replication, they are somewhat surprising and suggest potentially that the ability of the M protein to bind ATP, rather than ADP maybe be a requirement for the interaction with AvrM. In the interaction assays presented here, ATP and magnesium were included in the interaction buffers. Future work will utilise the Co-IP approach to define if an M·ADP association is required for interaction, and if the addition of nucleotide and/or the type of nucleotide affects the protein-protein interaction.

B) *What areas/domains of the M protein are required for AvrM interaction, with a focus towards defining residues involved in specificity?* A series of domain swap experiments involving alleles of the flax L locus, have supplied a host of data suggesting that the LRR region of the flax R proteins has the primary role in specifying the interaction with rust effector molecules (reviewed by (Ellis et al., 2007a)) (see 1.7.3). This was further extended by modelling of the LRR domain of

L5 to enable docking of the known structure of the effector protein, AvrL567A. This identified key residues that could be predicted to form the surface of the interacting face of both the LRR and the effector proteins (Wang et al., 2007). The importance of these residues in the L5/AvrL567A interaction was then shown by Y2H analysis (Wang et al 2007). We intend to use the *in vitro* protein-protein interaction assay developed here, to test the findings of the Y2H experiments of L5 and AvrL567A.

Interestingly, from the Y2H data reported by Dodds *et al*, (2006), there is a strong correlation between the strength of the interaction observed in Y2H experiments, and the strength of the *in planta* HR (Dodds et al., 2006). This is most likely a result of the interaction dynamics between the flax rust effector and flax R protein. It is postulated by Wang et al., (2007), that either a strong continuous interaction helps to maintain the protein in an activated state, or alternatively, short repeated interactions facilitate the recycling/amplification of the resistance response. We would hypothesize from the data obtained in this study, that the interaction between effector and R protein is unlikely to involve a strong continuous association, at least for M and AvrM, as we were unable to resolve a complex over GF. Nevertheless, further work is required to support or refute the theory of Wang et al., (2007).

Future investigation of the M/AvrM interaction will focus on the binding/disassociation kinetics of the interaction. Activation of the M protein may not be governed by interaction *per se*, but rather by the kinetics of the interaction. This would be governed by the critical residues in both M and AvrM that are involved in the interaction event. If the strength of the interaction were to control M protein activation, we could potentially determine if minimal binding kinetics were required for activation to occur. Understanding the interaction at such a level would have major consequences for the potential engineering of rust resistance genes, which is a potential application of this work.

*C) How does the M protein convert from an autoinhibited state to an activated state?*

The model I presented above (Figure 6.1) regarding M proteins activation predicts that ADP to ATP exchange is associated with the transition from an autoinhibited to an activated state. This prediction has, however, not been demonstrated experimentally for any R protein, and whilst attempts were made in this study, we were unable to demonstrate any nucleotide-based effects associated with M and AvrM interaction. Determining the outcome of the interaction between M and AvrM on the, identity of the nucleotide in the NB pocket, and the conformation of the M protein, will be one of the primary objectives of our research in the coming few years. Currently there are limited data from the study of other R proteins to assist us in our experimental design, however, we are guided by the investigation of other STAND proteins. Of particular interest is the work generated from studies of Apaf-1. This protein oligomerises into the caspase recruitment

and activator, apoptosome, which is stimulated by cytochrome c in the presence of dATP/ATP. In the absence of the nucleotide, however, Apaf-1 forms an inactive irreversible aggregate that does not stimulate caspase activity (Kim et al., 2005). More recently, it was demonstrated that the ability of Apaf-1 to form functional oligomers was affected by a number of other proteins. These proteins, which included the chaperone protein Hsp70, functioned together to accelerate nucleotide exchange which is a requirement of Apaf-1 activation. In doing so, these proteins help to prevent the formation of an inactive Apaf-1/cytochrome c aggregate (Kim et al., 2008). This later study therefore demonstrates that other proteins assist during the apoptosome formation in a mammalian cell. It is predicted that the proteins play a role by either, directly improving nucleotide exchange, or alternatively maintaining the stability of Apaf-1 during the nucleotide exchange event, which enables the protein reorientation that is necessary for Apaf-1 to assume an active conformational state. It is of interest that chaperone and co-chaperone proteins have been implicated in interacting with R proteins and maintaining an R protein in an autoinhibited conformation (reviewed by (Lukasik and Takken, 2009, Shen and Schulze-Lefert, 2007)).

For investigators of R protein molecular function these findings stimulates a very important question; are other plant proteins involved/required for converting an R protein from an inactive to an active conformation? We are currently defining experiments that may assist in answering this question.

#### **6.3.4 Further characterisation of the NB-ARC domain, including other R proteins**

While an array of gain- and loss-of-function mutations are localised to motifs within the NB-ARC domain, the consequences that these mutations have on molecular function remains predominantly uncharacterised. Work within the laboratory is currently focusing on utilising the established protein production and assay methods to further characterise a number of potentially critical residues in R protein function. The residues that we are currently focusing on are within the kinase 2 (D364E and D364A), GLPL (G449E) and RNBS-B (R392A) motifs. We have incorporated these mutations into M and the putative autoactive protein M<sup>D555V</sup> to investigate their effects on nucleotide association (P. Somaraj, personal communication).

The current mutations we have targeted in M are based on interesting results gained from functional and biochemical studies of R proteins and STAND proteins. Within the kinase 2 motif, mutations of the second aspartate have been shown to cause autoactivity in three R proteins, RPS5, I-2 and Mi-1 (Ade et al., 2007, Tameling et al., 2006, van Ooijen et al., 2008a), with biochemical data linking this particular mutation to an ATP bound state (Tameling et al., 2006). A mutation of the glycine in the well conserved GLPL motif of the flax R protein P2, abolished specific resistance to flax rust, however, the effect that this mutation has on the identity

and binding efficiency of the nucleotide within the NB pocket remains to be characterised (Dodds et al., 2001b). The highly conserved arginine in the RNBS-B domain is predicted to be a critical residue in ATP binding, as the equivalent residue in CED-4 is involved in an interaction with the  $\gamma$ -phosphate of the bound ATP (Yan et al., 2005). This arginine has been classified as the sensor-I in the NB-ARC domain, and the mutation of this arginine in I-2 is suggested to inhibit protein function (van Ooijen et al., 2008b). It is anticipated that this type of mutational approach will help to further define the functional roles of these motifs within the NB-ARC domain, which will, in turn, strengthen our understanding of R protein function.

In order to generalise the information gained for M into a model of R protein function, repetition of the *in vitro* work described here is most likely required in other R proteins. To this end, we are currently working towards repeating the mutational *in vitro* work in the closely related flax R protein, L6. Broadening this work to include non-TIR proteins is also of particular interest, as a large amount of biochemical information is already available for a number of these proteins, in particular Rx, I-2 and Mi-1. With this in mind, expression trials of the CC-NB-ARC-LRR *A. thaliana* R protein RPS2 in *P. pastoris* have been performed within the laboratory. Unfortunately to date, the production of full-length and numerous truncated variants of RPS2 have proved unsuccessful (E. deCourcy-Ireland, personal communication). We currently have no indication as to the reasons behind the failure of these proteins to express in the *P. pastoris* system, however, work is continuing. It is hoped that in the future we, or others, will be able to take the expression and purification procedure optimised for the production of flax R proteins and use it to produce other NB-ARC-LRR R proteins. This would ideally help broaden the biochemical study of R proteins, which will undoubtedly resolve many of the questions regarding R protein molecular function.

### 6.3.5 Intramolecular interactions and protein structure

While the structure of the NB-ARC domain of Apaf-1 has provided important insights into the function of this region in an R protein, a clearer picture would obviously be obtained from the structure of an NB-ARC-LRR R protein. Our ultimate aim, when we began expression and purification studies, was to solve the three dimensional structure of a flax R protein, and this continues to be a research objective within the collaborative group. A structure of the full-length M and/or L6 protein in both their autoinhibited and activated states would provide the maximum insight into the function of these proteins. This information would identify the important intramolecular interactions that exist between domains and subdomains during regulation and activation. In the flax-flax rust system it would also enable docking studies of the direct interaction between effectors and R proteins to be obtained. To enter crystallisation trials

confidently, we need to refine our expression and purification system to improve issues such as protein stability, protein concentration and protein purity. We are currently working on numerous methods to improve these issues and Appendix 7 highlights one of these. Using an increased selection approach during transformation, it is possible to get multiple insertion of the expression cassette into the *P. pastoris* genome. This can, in turn, increase the levels of recombinant protein without increasing the scale of expression. This was trialled with the cassette driving the expression of the M protein and at this stage, while it is not clear if multiple insertion events were obtained as a Southern blot has not been used to confirm this, I do see a clear increase in protein expression (Appendix 7). While a three dimensional structure of a flax R protein is our ultimate goal, a number of other methods are available to investigate structure, and the induction of structural changes. The use of circular dichroism and light scattering techniques may prove useful in further understanding the structural changes that are associated with autoinhibited and activated M protein. An important experiment would involve comparing the profiles of the M and M<sup>D555V</sup> proteins.

### 6.3.6 Intermolecular interactions and signalling

A combination of genetic and biochemical approaches have been used to identify a number of plant proteins that interact with various R proteins (reviewed by (Lukasik and Takken, 2009)). Currently, these proteins can be divided into those that act in, pathogen perception (guardee or decoys), stabilisation duties (chaperon/co-chaperon) or downstream signalling. Of those so far identified, only a few have been identified that are anticipated to be involved in downstream signalling.

Isolation of downstream partners in the resistance response is not a trivial exercise for a number of reasons. Utilising a genetic approach to identify downstream proteins may limit the net within which these genes can be captured, as only proteins encoded by genes that don't lead to lethal phenotypes when over-expressed or knocked-out can be identified. Also, whilst Co-IP approaches have been successful in isolating numerous interacting proteins, including chaperones/co-chaperones and signalling components, the use of this methodology to isolate downstream signalling molecules may be technically challenging. To elaborate, presumably an activated R protein interacts with at least some different proteins compared to when it exists in its autoinhibited form. To capture an interacting protein that only interacts when the protein is activated would be extremely difficult if it oscillates rapidly between inactive and active states. In the future it may be possible to use recombinant R proteins as a bait to isolate interacting protein, potentially solving some of these issues. Indeed the use of recombinant protein as a bait to isolate interacting partners is not a novel technique and has been used in the past to identify

interacting partner proteins in plants (Ok et al., 2005). The experimental design to perform such an experiment would involve the use of both the M and M<sup>D555V</sup> purified recombinant protein to pulldown interacting partners from a soluble extract of flax proteins. If these recombinant proteins resemble the autoinhibited state and the activated state of M, then we would predict a different profile of interacting proteins would be obtained. Mass spectral identification of these proteins would require Y2H and genetic analysis to verify such results. Also, the experiment itself has many design hurdles; however, it is an example of what could potentially be achieved using the recombinant R protein expression system outlined here.

Identifying a downstream signalling partner would potentially have great advantages for the future biochemical studies of NB-ARC-LRR R proteins. A major advantage in studying proteins such as Apaf-1 and the NLR protein, NALP1, is that the respective caspase-9 and caspase-1 proteins, that these proteins activate, have activities that are measurable (Bao et al., 2007, Faustin et al., 2007). Put simply, the level of caspase activation can be correlated to the activity of the activating protein, thus providing a simple and specific assay. This makes it possible to define external factors that inhibit or promote function or identify mutations that affect function. Obviously, in the case of R proteins this would only be possible if the downstream protein had a measurable or definable activity *in vitro*. Ultimately we can envisage the ability to reconstitute the resistance complex “resistasome” *in vitro*.

The signalling pathways utilised by NB-ARC-LRR R proteins appear to be dependent on the structural class from which the R protein originates. The non-TIR class, require the NDR1 signalling component, while the TIR class require the signalling components EDS1/PAD4 (Aarts et al., 1998). These results, initially supported the early suggestion that the N-terminal domain of R proteins was responsible for signalling; however, more recent data implicated the NB-ARC domain as the functional signalling unit of R proteins (Lukasik and Takken, 2009). Direct evidence for this originated from studies in Rx, where the NB subdomain was sufficient to initiate downstream signalling (Rairdan et al., 2008). This report also identified that nucleotide binding was not required for this signalling event. Within the NB region exists the RNBS-A motif which is one of two motifs known to differ in consensus sequence between non-TIR and TIR proteins in the NB-ARC domain (Table 1.2/Figure 1.3)). Given the observed differences in the signalling pathways between TIR and non-TIR proteins (Aarts et al., 1998) it has been argued that the RNBS-A motif may interact with downstream signalling components (Lukasik and Takken, 2009). Mutations that affect function have been localised to the RNBS-A motif. Using our *in vitro* system we could test if mutations in this motif that affect function have any functional consequence on nucleotide binding. Presumably, if they are purely involved in signalling then such a mutation would have no nucleotide binding consequence. It may therefore be possible to develop a

system that uncouples mutations that affect recognition/interaction and/or downstream signalling from those that directly impact on nucleotide binding. One could also explore the use of the Y2H technique as a screening tool using M and M<sup>D555V</sup> as baits.

#### **6.4 Conclusion**

The NB-ARC-LRR R proteins are critical components of plant defence to pathogen attack and are required for proper function of the plant immune system. Before genetic engineering, or other application sciences, can successfully exploit R proteins for the improvement of crop plants, we require precise knowledge of their function at the molecular level. Contributing to a system that enables the investigation of the molecular function of flax R proteins has been the overall goal of this thesis. The data outlined here demonstrate that functional flax R protein can be expressed and purified in a heterologous protein expression system. It is hoped that this work will inspire and guide biochemical studies to further elucidate the mechanisms that R proteins use to protect the plant against pathogen attack.

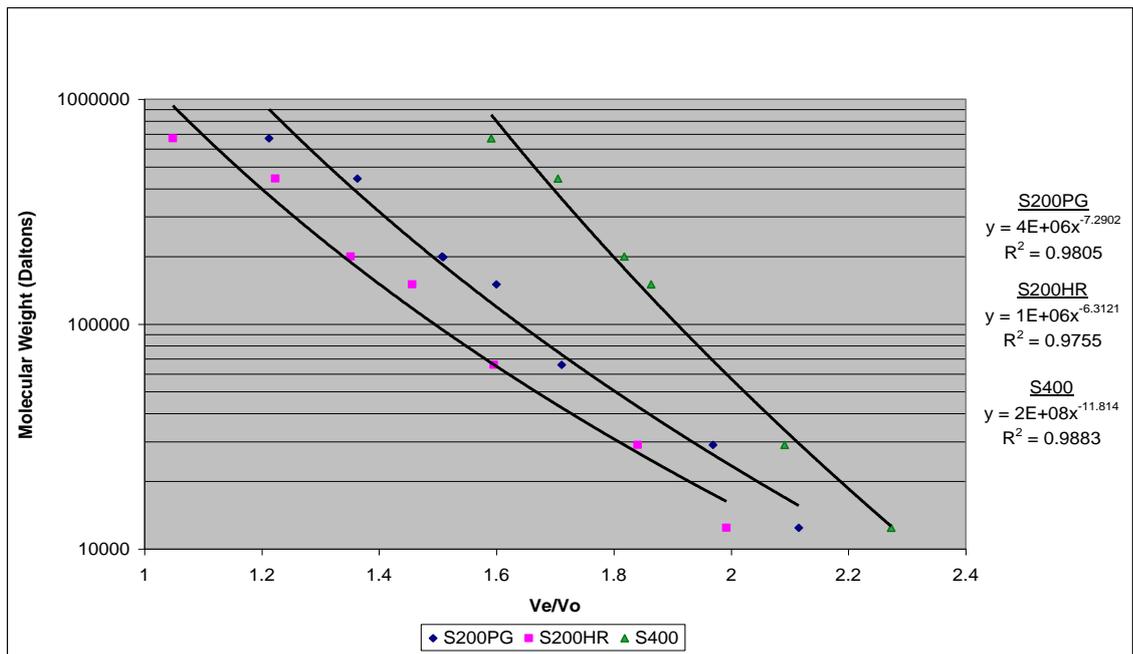


# Appendix

## Appendix 1: Sequencing Oligonucleotides

| Oligonucleotide name                | Sequence (5'-3')          |
|-------------------------------------|---------------------------|
| <b>M sequencing primers</b>         |                           |
| MForward1                           | CAGACTGGACCTTATCGAAAGGC   |
| MForward2                           | CGATGAGCGAACAACATTTCGC    |
| MForward3                           | ATGCTATATGCTTGGGAGTCGG    |
| MForward4                           | GGACTCATAGAGCTTCGTCTCG    |
| MForward5                           | ATGCGAACTCCACGACCAAAC     |
| MForward6                           | GATCTGGATGTGATTGGATCCC    |
| MReverse1                           | GATTGCTCTGAGGAGGTTGACC    |
| MReverse2                           | CCTTCTGCTCTTGCATTGCTC     |
| MReverse3                           | CCGATGAAGAAGCAGGCTATATC   |
| MReverse4                           | AAGAAAGTCTTTGGGAGACCCC    |
| MReverse5                           | TCTCTATTCTGGAAGCCACG      |
| MReverse6                           | CGTGCAACCCTCTAAAATCAACC   |
| <b>AvrM/avrM sequencing primers</b> |                           |
| AvrMForward1                        | CGAAGAGGTCAAAGATGGTGTA    |
| AvrMReverse1                        | TACAAATTCAACTCTGGATCGTCTG |
| <b>Primers for pPICz</b>            |                           |
| 3'AOX                               | GCAAATGGCATTCTGACATCC     |
| 5'AOX                               | GACTGGTTCCAATTGACAAGC     |
| <b>Primers for pET</b>              |                           |
| T7 promoter primer                  | TAATACGACTCACTATAGG       |

## Appendix 2: Gel Filtration Calibration Curves



| <b>S200PG</b>         |           |                     |             |             | Vt=120ml |
|-----------------------|-----------|---------------------|-------------|-------------|----------|
| Protein Standards     | size (Da) | elution volume (ml) | log(mw)     | Ve/Vo       |          |
| Dextran blue          | 4000000   | 45                  | 6.602059991 | 1           |          |
| Thyroglobulin         | 669000    | 54.55               | 5.825426118 | 1.212222222 |          |
| Apoferretin           | 443000    | 61.34               | 5.646403726 | 1.363111111 |          |
| B-Amylase             | 200000    | 67.832              | 5.301029996 | 1.507377778 |          |
| Alcohol dehydrogenase | 150000    | 72                  | 5.176091259 | 1.6         |          |
| Albumin               | 66000     | 77                  | 4.819543936 | 1.711111111 |          |
| Carbonic anhydrase    | 29000     | 88.6                | 4.462397998 | 1.968888889 |          |
| Cytochrome C          | 12400     | 95.186              | 4.093421685 | 2.115244444 |          |

| <b>S200HR</b>         |           |                     |             |             | Vt=24ml |
|-----------------------|-----------|---------------------|-------------|-------------|---------|
| Protein Standards     | size (Da) | elution volume (ml) | log(mw)     | Ve/Vo       |         |
| Dextran blue          | 4000000   | 9.1                 | 6.602059991 | 1           |         |
| Thyroglobulin         | 669000    | 9.54                | 5.825426118 | 1.048351648 |         |
| Apoferretin           | 443000    | 11.13               | 5.646403726 | 1.223076923 |         |
| B-Amylase             | 200000    | 12.3                | 5.301029996 | 1.351648352 |         |
| Alcohol dehydrogenase | 150000    | 13.25               | 5.176091259 | 1.456043956 |         |
| Albumin               | 66000     | 14.52               | 4.819543936 | 1.595604396 |         |
| Carbonic anhydrase    | 29000     | 16.75               | 4.462397998 | 1.840659341 |         |
| Cytochrome C          | 12400     | 18.13               | 4.093421685 | 1.992307692 |         |

| <b>S400</b>           |           |                     |             |             | Vt=120ml |
|-----------------------|-----------|---------------------|-------------|-------------|----------|
| Protein Standards     | size (Da) | elution volume (ml) | log(mw)     | Ve/Vo       |          |
| Dextran blue          | 4000000   | 44                  | 6.602059991 | 1           |          |
| Thyroglobulin         | 669000    | 70                  | 5.825426118 | 1.590909091 |          |
| Apoferretin           | 443000    | 75                  | 5.646403726 | 1.704545455 |          |
| B-Amylase             | 200000    | 80                  | 5.301029996 | 1.818181818 |          |
| Alcohol dehydrogenase | 150000    | 82                  | 5.176091259 | 1.863636364 |          |
| Carbonic anhydrase    | 29000     | 92                  | 4.462397998 | 2.090909091 |          |
| Cytochrome C          | 12400     | 100                 | 4.093421685 | 2.272727273 |          |

NB: Ve is elution volume; Vo is void volume; Vt is total column volume.

### **Appendix 3: Overview of the Truncation/Mutations of M Expressed in *P. pastoris***

MSYL RDVATAVALLLDNLCC**G**RPNLNNDNEDIQQT DSTSPVDPSSSSQSM DSTSVD AISDSTNPSASFPSVEYDVFLSFRGP  
DTRYQITDILYRFLCRSKIHTFKDDDELHKGEEIKVNLRAIDQSKIYVPIISRGYADSKWCLMELAKIVRHQKLDTRQIIPIFYMVDPKD  
VRHQTGPYRKAFQKHSTRYDEMTIRSWKNALNEVGANKGWHVKNND**E**Q**G**AIAD~~EV~~SANI**W**SHISKENFILETDELVGIDDHVEVIL  
EMLSLSKSVTMVGLY**G**MG**G**IG**K**TTAKAVYNKISSHFDRCCFVDNVRAMQE QKDGFILQKLVSEILRMDSVGFTNDSSGGRKMI  
KERVSKSKILVLDVDEKFKFEDILGCPKDFDSGTRFIITSRNQNVLSRLNENQCKLYEVGSMSEQHSLELFSKHAFKKNTPPSDY  
ETLANDIVSTTGGLPLTLKVTGSFLFRQEIGVWEDTLEQLRKTLDLDEVYDRLKI**S**YDALKAEAKEIFLDIACFFIGRNKEMPYYMWS  
ECKFYPKSNIIFLIQRCEMIQVGGDGVLEM**H**DQLRDMGREIVRREDVQRPWKRSRIWSREEGIDLLLNNKGSSQVKAISIPNNMLYA  
WESGVKYEFKSECFNLSELRLFFVGSTLLTGDFNNLLPNLKWLDLPRYAHGLYDPPVTNFTMKKLVILVSTNSKTEWSHMIKM  
APRLKVVRLYSDYGVSRQLSFCWRFPKSI EVLMSGIEIKEVDIGELKNLKTLDLTSCRIQKISGGTFGMLKGLIELRLDSIKCTNLR  
EVDVIGLQSSSLKVLKTEGAQEVQFEFPLAKELSTSSRIPNLSQLLDLEVLKVYGCNDGFDIPPAKSTEDEGSVVWKASKLKSL  
KLYRTRININVDASSGGRYLLPSSLTSL EIWCKEPTWLPGIENLENLTSLVDDVDIFQTLGGDL DGLQGLRSLETLTITEVNGLT  
RIKGLMDLLCSSTCKLEKLEIKACHDLTEILPCELHDQTVVVPSEKLTIRDCPRLEVGP MIRSLPKFPMLKKLDLAVANITKEEDL  
DVIGSLQELVDLRIELDDTSSGIERIASLSK LKLTTRVKVPSLREIEELAALKSLQRLILEGCTSLERLRLKLEKLEKPEDIGGCPDLTE  
LVQTVVWCPSELVELTIRDCPRLEVGP MIRSLPKFPMLKKLDLAVANIEEDLDVIGSLEELVILSLKLD DTTSSSIERISFLSKLQKLF  
LRVKVSSLREIEGLAELKSLQLLFLKGCTSLERLWPDEQQLDNNKSMRIDIRGCKSLSDVHLSALKSTLPPNVKIRWPDEKYK

#### **Key**

**G**: Represents the first amino acid of M after the 9x histidine tag that enabled near full-length expression of M.

**CGRP**NLNNDNEDIQ: The M epitope for the anti-M antibody.

**Q**: The first amino acid of the M $\Delta$ TIR after the 9x histidine tag (L6 $\Delta$ TIR was truncated at the same residue in its coding sequence).

**W**: A later alignment demonstrates that this residue is most likely the last residue that makes up the TIR domain.

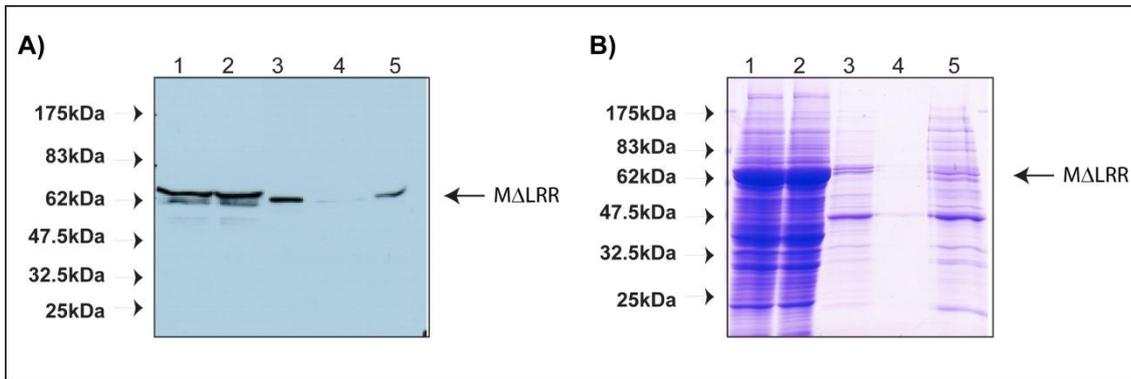
**GMGGIGKTT**: The P-loop motif with the invariant lysine highlighted. This residue was mutated to a leucine in this study.

**SYD**: The VIII motif with the highly conserved serine highlighted. This residue was mutated to an arginine in this study.

**MHD**: The MHD motif with the highly conserved aspartate. This residue was mutated to a valine in this study.

The C-terminal bold section represents the LRR domain. This section was removed to generate the M $\Delta$ LRR construct.

## Appendix 4: Purification of M $\Delta$ LRR



Using the same purification strategy that proved successful for the purification of M and M $\Delta$ TIR, low yields and low purity was achieved for M $\Delta$ LRR, preventing its inclusion in biochemical analysis: A) Western blot analysis using an anti-M antibody showing the presence and absence of M $\Delta$ LRR protein during purification. Lane 1, crude lysate; lane 2, NiA flowthrough; lane 3, NiA 55mM imidazole wash (10c/v); lane 4, NiA elution; lane 5, concentrated elution. B) Coomassie stained gels of protein separated by SDS-PAGE with the same loading arrangement as western blot analysis, arrow indicates M $\Delta$ LRR protein.

## Appendix 5: Supplementary Information for ATP and ADP Quantification Experiments

### Nucleotide extraction methods

| Protein |                    | Nucleotide concentration (M)                  | Fold difference |
|---------|--------------------|---|-----------------|
| M       | Acid Precipitation | $2.3 \times 10^{-8} \pm 0.055 \times 10^{-8}$ | ~5.5            |
|         | Boiling            | $12.6 \times 10^{-8} \pm 0.35 \times 10^{-8}$ |                 |

The nucleotide concentrations were the average of three separate extraction procedures

### ATP/ADP quantification data for M proteins

| Protein                  | ATP Occupancy      | ADP Occupancy      |
|--------------------------|--------------------|--------------------|
| M                        | $2.08 \pm 0.08\%$  | $34.89 \pm 2.95\%$ |
| M <sup>K286L</sup>       | $0.54 \pm 0.12\%$  | $2.70 \pm 1.30\%$  |
| M <sup>D555V</sup>       | $17.96 \pm 4.15\%$ | $5.40 \pm 2.03\%$  |
| M <sup>K286L+D555V</sup> | $0.76 \pm 0.34\%$  | $7.31 \pm 3.07\%$  |
| M <sup>S492R</sup>       | $8.49 \pm 1.57\%$  | $4.69 \pm 2.50\%$  |
| M $\Delta$ TIR           | $0.7 \pm 0.3\%$    | $40.82 \pm 6.81\%$ |

Percentage ATP and ADP occupancies were calculated by comparing nucleotide concentration (M) to protein concentration (M) (as determined by the Sypro Ruby staining technique). The percentage ATP and ADP occupancy values represent the mean, which was calculated from 4 independent purifications originating from 2 independently grown cell culture expressions. Error bars represent standard error.

## Appendix 6: AvrM/avrM/avrMC\* Alignment

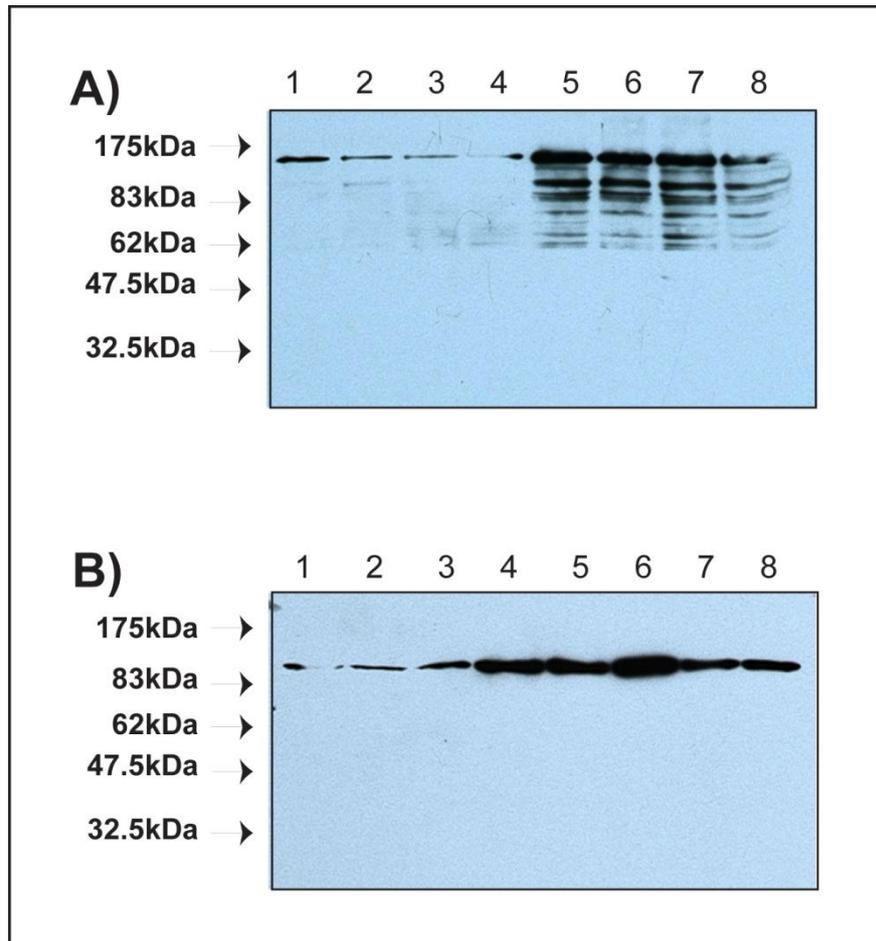
|        |   |  |   |     |
|--------|---|--|---|-----|
| AvrM   | : | MGSSHHHHHSSGLVPRGSHMLEDPNSIKMHPMNSAKLAEVVKDGLNAAHGDSLNNLGTVPDVPHQIPNDKSGTPAIEDPKDMKGFNKALKSTPESEKL   | : | 100 |
| avrM   | : | .....  | : | 47  |
| avrMC* | : | .....  | : | 47  |
| AvrM   | : | GTSSVEGIPQPEFDRGFLRPFGAKMKFLKPDQVQKLSTDDLITYMAEKDKNVRDLAIKLRDAKQDSTKNGTPEIKQTYDKAYEKTAAAEEKLVSEESLTR | : | 200 |
| avrM   | : | .....K.....  | : | 138 |
| avrMC* | : | .....K.....  | : | 138 |
| AvrM   | : | DALLKLTTEEQYVEKAALFDKDVYRNKLRQTYEKLRSSETDVLVREVARIFIAREGEPALTAKIERLALTLENNADTRSKPIDYLAIAADFLKNQANLHA | : | 300 |
| avrM   | : | ...E.....Q.....S.....T.....N.....  | : | 237 |
| avrMC* | : | ...E.....Q.....S.....T.....N.....  | : | 237 |
| AvrM   | : | DDPELNLYKAEIKAREIEANRAMKEALKGADKLFKRNKILKSPDM-----   | : | 345 |
| avrM   | : | .....T.....K.....E.....E.....RYKSAGFQAFIDKMMMAULSKIMTTRSUYIKSLAKP                                    | : | 316 |
| avrMC* | : | .....T.....K.....E.....E.....  | : | 282 |

Multiple-sequence alignment of AvrM, avrM and the generated avrMC\* proteins.

### Key

- 6x histidine tag
- Thrombin cleavage site
- Start of coding sequence

## Appendix 7: Improved Expression



**Increased selection during transformation increased the expression of the recombinant M and M $\Delta$ TIR proteins:** Cleared lysates from the independent test expression of M (A) and M $\Delta$ TIR (B) clones were subjected to western analysis with an anti-M and an anti-His antibody, respectively. 30  $\mu$ g of protein was loaded into each lane and separated by SDS-PAGE before western blot analysis. Lane's 1-4 independent clones screened at 1x zeocin (100 $\mu$ g/ml) selection and lanes 5-8 independent clones screened at 6x zeocin (600 $\mu$ g/ml) selection.



# Bibliography

- Aarts, N., Metz, M., Holub, E., Staskawicz, B. J., Daniels, M. J. & Parker, J. E. (1998) Different requirements for *EDS1* and *NDR1* by disease resistance genes define at least two *R* gene-mediated signalling pathways in Arabidopsis. *Proc. Natl. Acad. Sci. USA*, **95**, 10306-10311.
- Ade, J., DeYoung, B. J., Golstein, C. & Innes, R. (2007) Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease. *Proc. Natl. Acad. Sci. USA*, **104**, 2531-2536.
- Albrecht, M. & Takken, F. L. W. (2006) Update on the domain architectures of NLRs and R proteins. *Biochemical and Biophysical Research Communications*, **339**, 459-462.
- Altenbach, D. & Robatzek, S. (2007) Pattern recognition receptors: From the cell surface to intracellular dynamics. *Molecular Plant-Microbe Interactions*, **20**, 1031-1039.
- Anderson, P. A., Lawrence, G. J., Morrish, B. C., Ayliffe, M. A., Finnegan, E. J. & Ellis, J. G. (1997) Inactivation of the flax rust resistance gene M associated with loss of a repeated unit within the leucine-rich repeat coding region. *Plant Cell*, **9**, 641-651.
- Asai, T., Tena, G., Plotnikova, J., Willmann, M. R., Chiu, W. L., Gomez-Gomez, L., Boller, T., Ausubel, F. M. & Sheen, J. (2002) MAP kinase signalling cascade in Arabidopsis innate immunity. *Nature*, **415**, 977-983.
- Axtell, M. J., McNellis, T. W., Mudgett, M. B., Hsu, C. S. & Staskawicz, B. J. (2001) Mutational analysis of the Arabidopsis RPS2 disease resistance gene and the corresponding *Pseudomonas syringae* avrRpt2 avirulence gene. *Molecular Plant-Microbe Interactions*, **14**, 181-188.
- 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.
- Bailey, S. (1994) The CCP4 Suite - Programs for protein crystallography. *Acta Crystallographica Section D-Biological Crystallography*, **50**, 760-763.
- Bao, Q., Lu, W. Y., Rabinowitz, J. D. & Shi, Y. G. (2007) Calcium blocks formation of apoptosome by preventing nucleotide exchange in Apaf-1. *Mol. Cell*, **25**, 181-192.
- Bao, Q., Riedl, S. J. & Shi, Y. G. (2005) Structure of Apaf-1 in the auto-inhibited form - A critical role for ADP. *Cell Cycle*, **4**, 1001-1003.
- Bao, Q. & Shi, Y. (2007) Apoptosome: a platform for the activation of initiator caspases. *Cell Death and Differentiation*, **14**, 56-65.
- Bella, J., Hindle, K., McEwan, P. & Lovell, S. (2008) The leucine-rich repeat structure. *Cellular and Molecular Life Sciences*, **65**, 2307-2333.
- 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 Journal*, **32**, 195-204.
- Bent, A. F. & Mackey, D. (2007) Elicitors, effectors, and R genes: The new paradigm and a lifetime supply of questions. *Annual Review of Phytopathology*, **45**, 399-436.
- Bieri, S., Mauch, S., Shen, Q. H., Peart, J. & Devoto, A. (2004) RAR1 positively controls steady state levels of barley MLA resistance proteins and enables sufficient MLA6 accumulation for effective resistance. *Plant Cell*, **16**, 3480.
- Birch, P. R. J., Rehmany, A. P., Pritchard, L., Kamoun, S. & Beynon, J. L. (2006) Trafficking arms: oomycete effectors enter host plant cells. *Trends in Microbiology*, **14**, 8-11.
- Block, A., Li, G. Y., Fu, Z. Q. & Alfano, J. R. (2008) Phytopathogen type III effector weaponry and their plant targets. *Current Opinion in Plant Biology*, **11**, 396-403.
- Blum, H., Beier, H. & Gross, H. J. (1987) Silver staining of proteins in polyacrylamide gels. *Electrophoresis*, **8**, 93-99.

- Boyes, D. C., Nam, J. & Dangl, J. L. (1998) The *Arabidopsis thaliana* RPM1 disease resistance gene product is a peripheral plasma membrane protein that is degraded coincident with the hypersensitive response. *Proc. Natl. Acad. Sci. USA*, **95**, 849.
- Burch-Smith, T. M., Schiff, M., Caplan, J. L., Tsao, J., Czymmek, K. & Dinesh-Kumar, S. (2007) A Novel Role for the TIR Domain in Association with Pathogen-Derived Elicitors. *PLOS Biology*, **5**, 501-514.
- Büttner, D. & Bonas, U. (2002) Getting across—bacterial type III effector proteins on their way to the plant cell. *Embo J.*, **21**, 5313-5322.
- Caplan, J. L., Mamillapalli, P., Burch-Smith, T. M., Czymmek, K. & Dinesh-Kumar, S. P. (2008) Chloroplastic protein NRIP1 mediates innate immune receptor recognition of a viral effector. *Cell*, **132**, 449-462.
- 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.
- Chinchilla, D., Bauer, Z., Regenass, M., Boller, T. & Felix, G. (2006) The *Arabidopsis* receptor kinase FLS2 binds flg22 and determines the specificity of flagellin perception. *Plant Cell*, **18**, 465-476.
- Chinchilla, D., Zipfel, C., Robatzek, S., Kemmerling, B., Nurnberger, T., Jones, J. D. G., Felix, G. & Boller, T. (2007) A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. *Nature*, **448**, 497-U12.
- 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.
- Dangl, J. L. & Jones, J. D. G. (2001) Plant pathogens and integrated defence responses to infection. *Nature*, **411**, 826-833.
- Danot, O., Marquenot, E., Vidal-Ingigliardi, D. & Richet, E. (2009) Wheel of life, wheel of death: A mechanistic insight into signaling by STAND proteins. *Structure*, **17**, 172-182.
- de la Fuente van Bentem, S., Vossen, J. H., de Vries, K. J., van Wees, S., Tameling, W. I. L., Dekker, H. L., 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 Journal*, **43**, 284-298.
- Deslandes, L., Olivier, J., Peeters, N., Feng, D. X., Khounlotham, M., Boucher, C., Somssich, I., Genin, S. p. & Marco, Y. (2003) Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. *Proc. Natl. Acad. Sci. USA*, **100**, 8024-8029.
- Dinesh-Kumar, S. P., Tham, W. H. & 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. & Thrall, P. (2009) Recognition events and host-pathogen co-evolution in gene-for-gene resistance to flax rust. *Funct. Plant Biol.*, **36**, 395-408.
- 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 recognized inside plant cells. *Plant Cell*, **16**, 755-768.
- Dodds, P. N., Lawrence, G. J., Catanzariti, A. M., Teh, T., Wang, C. I. A., Ayliffe, M. A., Kobe, B. & Ellis, J. G. (2006) Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proc. Natl. Acad. Sci. USA*, **103**, 8888-8893.
- Dodds, P. N., Lawrence, G. J. & Ellis, J. G. (2001a) Contrasting modes of evolution acting on the complex N locus for rust resistance in flax. *Plant Journal*, **27**, 439.
- Dodds, P. N., Lawrence, G. J., Pryor, T. & Ellis, J. (2001b) Six amino acid changes confined to the leucine-rich repeat beta-strand/beta-turn motif determine the difference between the P and P2 rust resistance specificities in flax. *Plant Cell*, **13**, 163-178.

- Ellis, J., Dodds, P. & Pryor, T. (2000) Structure, function and evolution of plant disease resistance genes. *Current Opinion in Plant Biology*, **3**, 278-284.
- Ellis, J. G., Dodds, P. N. & Lawrence, G. J. (2007a) Flax rust resistance gene specificity is based on direct resistance-avirulence protein interactions. *Annual Review of Phytopathology*, **45**, 289-306.
- Ellis, J. G., Lawrence, G. J. & Dodds, P. N. (2007b) Further analysis of gene-for-gene disease resistance specificity in flax. *Molecular Plant Pathology*, **8**, 103-109.
- 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.
- Faustin, B., Lartigue, L., Bruey, J. M., Luciano, F., Sergienko, E., Bailly-Maitre, B., Volkmann, N., Hanein, D., Rouiller, I. & Reed, J. C. (2007) Reconstituted NALP1 inflammasome reveals two-step mechanism of caspase-1 activation. *Mol. Cell*, **25**, 713-724.
- Fliegmann, J., Mithofer, A., Wanner, G. & Ebel, J. (2004) An ancient enzyme domain hidden in the putative beta-glucan elicitor receptor of soybean may play an active part in the perception of pathogen-associated molecular patterns during broad host resistance. *J. Biol. Chem.*, **279**, 1132-1140.
- Flor, H. H. (1956) The complementary genic systems in flax and flax rust. *Adv. Gen.*, **8**, 2954.
- Flor, H. H. (1971) Current status of gene-for-gene concept. *Annual Review of Phytopathology*, **9**, 275-296.
- Frost, D., Way, H., Howles, P., Luck, J., Manners, J., Hardham, A., Finnegan, J. & Ellis, J. (2004) Tobacco transgenic for the flax rust resistance gene L expresses allele-specific activation of defense responses. *Molecular Plant-Microbe Interactions*, **17**, 224-232.
- Gabriels, S. H. E. J., Vossen, J. H., Ekengren, S. K., van Ooijen, G., Abd-El-Halim, 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 Journal*, **50**, 14-28.
- Gaffney, T., Friedrich, L., Vernooij, B., Negrotto, D., Nye, G., Uknes, S., Ward, E., Kessmann, H. & Ryals, J. (1993) Requirement of salicylic-acid for the induction of systemic acquired-resistance. *Science*, **261**, 754-756.
- Galan, J. E. & Collmer, A. (1999) Type III secretion machines: Bacterial devices for protein delivery into host cells. *Science*, **284**, 1322-1328.
- Gomez-Gomez, L., Bauer, Z. & Boller, T. (2001) Both the extracellular leucine-rich repeat domain and the kinase activity of FLS2 are required for flagellin binding and signaling in Arabidopsis. *Plant Cell*, **13**, 1155-1163.
- Grant, M. R., Godiard, L., Straube, E., Ashfield, T., Lewald, J., Sattler, A., Innes, R. W. & Dangl, J. L. (1995) Structure of the Arabidopsis RPM1 gene enabling dual-specificity disease resistance. *Science*, **269**, 843-846.
- Greenberg, J. T. (1997) Programmed cell death in plant-pathogen associations. *Annual Review of Phytopathology*, **48**, 525-545.
- Hahn, M. & Mendgen, K. (2001) Signal and nutrient exchange at biotrophic plant-fungus interfaces. *Current Opinion in Plant Biology*, **4**, 322-327.
- Hammond-Kosack, K. E. & Jones, J. D. G. (1997) Plant disease resistance genes. *Annual Review of Plant Physiology and Plant Molecular Biology*, **48**, 575-607.
- Hammond-Kosack, K. E. & Parker, J. E. (2003) Deciphering plant-pathogen communication: fresh perspectives for molecular resistance breeding. *Current Opinion in Plant Biology*, **14**, 177-193.
- He, P., Shan, L., Lin, N. C., Martin, G. B., Kemmerling, B., Nurnberger, T. & Sheen, J. (2006) Specific bacterial suppressors of MAMP signaling upstream of MAPKKK in Arabidopsis innate immunity. *Cell*, **125**, 563-575.
- Heath, M. (2000) Hypersensitive response-related cell death. *Plant Molecular Biology*, **44**, 321-334.

- Heath, M. C. (1997) Signalling between pathogenic rust fungi and resistant or susceptible host plants. *Ann Bot*, **80**, 713-720.
- Hogenhout, S. A., van der Hoorn, R. A. L., Terauchi, R. & Kamoun, S. (2009) Emerging concepts in effector biology of plant-associated organisms. *Molecular Plant-Microbe Interactions*, **22**, 115-122.
- Holt, B. F., Hubert, D. A. & Dangl, J. L. (2003) Resistance gene signaling in plants -- complex similarities to animal innate immunity. *Current Opinion in Immunology*, **15**, 20-25.
- Houterman, P. M., Cornelissen, B. J. C. & Rep, M. (2008) Suppression of plant resistance gene-based immunity by a fungal effector. *PLoS Pathog.*, **4**, 6.
- Howles, P., Lawrence, G., Finnegan, J., McFadden, H., Ayliffe, M., Dodds, P. & Ellis, J. (2005) Autoactive alleles of the flax L6 rust resistance gene induce non-race-specific rust resistance associated with the hypersensitive response. *Molecular Plant-Microbe Interactions*, **18**, 570-582.
- Hwang, C.-F., Bhakta, A. V., Truesdell, G. M., Pudlo, W. M. & Williamson, V. M. (2000) Evidence for a role of the N terminus and leucine-rich repeat region of the Mi gene product in regulation of localized cell death. *Plant Cell*, **12**, 1319-1330.
- Hwang, C. F. & Williamson, V. M. (2003) Leucine-rich repeat-mediated intramolecular interactions in nematode recognition and cell death signaling by the tomato resistance protein Mi. *Plant Journal*, **34**, 585-593.
- 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 Journal*, **19**, 4004-4014.
- Jones, J. D. G. & Dangl, J. L. (2006) The plant immune system. *Nature*, **444**, 323-329.
- Kaku, H., Nishizawa, Y., Ishii-Minami, N., Akimoto-Tomiya, C., Dohmae, N., Takio, K., Minami, E. & Shibuya, N. (2006) Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 11086-11091.
- Kamoun, S. (2006) A catalogue of the effector secretome of plant pathogenic oomycetes. *Annu. Rev. Phytopathol.*, **44**, 41-60.
- Kay, S., Hahn, S., Marois, E., Hause, G. & Bonas, U. (2007) A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science*, **318**, 648-651.
- Kim, H. E., Du, F. H., Fang, M. & Wang, X. D. (2005) Formation of apoptosome is initiated by cytochrome c-induced dATP hydrolysis and subsequent nucleotide exchange on Apaf-1. *Proc. Natl. Acad. Sci. U. S. A.*, **102**, 17545-17550.
- Kim, H. E., Jiang, X. J., Du, F. H. & Wang, X. D. (2008) PHAPI, CAS, and Hsp70 promote apoptosome formation by preventing Apaf-1 aggregation and enhancing nucleotide exchange on Apaf-1. *Mol. Cell*, **30**, 239-247.
- Kobe, B. & Deisenhofer, J. (1993) Crystal-structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. *Nature*, **366**, 751-756.
- Kobe, B. & Kajava, A. V. (2001) The leucine-rich repeat as a protein recognition motif. *Current Opinion in Structural Biology*, **11**, 725-732.
- Koonin, E. V. & Aravind, L. (2000) The NACHT family - a new group of predicted NTPases implicated in apoptosis and MHC transcription activation. *Trends in Biochemical Sciences*, **25**, 223-224.
- Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.
- Lawrence, G. J., Anderson, P. A., Dodds, P. N. & Ellis, J. G. (2009) Relationships between rust resistance genes at the M locus in flax. *Molecular Plant Pathology*, **in press**.
- Lawrence, G. J., Dodds, P. N. & Ellis, J. G. (2007) Rust of flax and linseed caused by *Melampsora lini*. *Molecular Plant Pathology*, **8**, 349-364.
- Lawrence, G. J., Finnegan, E. J., Ayliffe, M. A. & Ellis, J. G. (1995) The L6 gene for flax rust resistance is related to the Arabidopsis bacterial resistance gene RPS2 and the tobacco viral resistance gene N. *Plant Cell*, **7**, 1195.

- 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. *Journal of Molecular Biology*, **343**, 1-28.
- Liu, J. & Coaker, G. (2008) Nuclear trafficking during plant innate immunity. *Molecular Plant*, **1**, 411-422.
- Luck, J. E., Lawrence, G. J., Dodds, P. N., Shepherd, K. W. & Ellis, J. G. (2000) Regions outside of the leucine-rich repeats of flax rust resistance proteins play a role in specificity determination. *Plant Cell*, **12**, 1367-1377.
- Lukasik, E. & Takken, F. L. W. (2009) STANDING strong, resistance proteins instigators of plant defence. *Current Opinion in Plant Biology*, **12**, 427-436.
- Mackey, D., Holt, B. F., 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.
- Marquenet, E. & Richet, E. (2007) How integration of positive and negative regulatory signals by a STAND signaling protein depends on ATP hydrolysis. *Mol. Cell*, **28**, 187-199.
- Martin, G. B., Bogdanove, A. J. & Sessa, G. (2003) Understanding the functions of plant disease resistance proteins. *Annual Review of Plant Biology*, **54**, 23-61.
- Martin, G. B., Brommonschenkel, S. H., Chunwongse, J., Frary, A., Ganai, M. W., Spivey, R., Wu, T. Y., Earle, E. D. & Tanksley, S. D. (1993) Map-based cloning of a protein-kinase gene conferring disease resistance in tomato. *Science*, **262**, 1432-1436.
- Martin, M. U. & Wesche, H. (2002) Summary and comparison of the signaling mechanisms of the Toll/interleukin-1 receptor family. *Biochimica et Biophysica Acta Molecular Cell Research*, **1592**, 265-280.
- Meng, F. G., Hong, Y. K., He, H. W., Lyubarev, A. E., Kurganov, B. I., Yan, Y. B. & Zhou, H. M. (2004) Osmophobic effect of glycerol on irreversible thermal denaturation of rabbit creatine kinase. *Biophysical Journal*, **87**, 2247-2254.
- 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 protein family within the nucleotide binding superfamily. *Plant Journal*, **20**, 317-332.
- Meyers, B. C., Kozik, A., Griego, A., Kuang, H. & Michelmore, R. W. (2003) Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. *Plant Cell*, **15**, 809.
- Mindrinis, M., Katagiri, F., Yu, G. L. & Ausubel, F. M. (1994) The *A-thaliana* disease resistance gene RPS2 encodes a protein containing a nucleotide-binding site and leucine-rich repeats. *Cell*, **78**, 1089-1099.
- Miya, A., Albert, P., Shinya, T., Desaki, Y., Ichimura, K., Shirasu, K., Narusaka, Y., Kawakami, N., Kaku, H. & Shibuya, N. (2007) CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.*, **104**, 19613-19618.
- Moffett, P., Farnham, G., Peart, J. & Baulcombe, D. C. (2002) Interaction between domains of a plant NBS-LRR protein in disease resistance-related cell death. *Embo Journal*, **21**, 4511-4519.
- Monosi, B., Wisser, R., Pennill, L. & Hulbert, S. (2004) Full-genome analysis of resistance gene homologues in rice. *Theoretical and Applied Genetics*, **109**, 1434-1447.
- Morgan, W. & Kamoun, S. (2007) RXLR effectors of plant pathogenic oomycetes. *Curr. Opin. Microbiol.*, **10**, 332-338.
- Mucyn, T. S., Clemente, A., Andriotis, V. M. E., Balmuth, A. L., Oldroyd, G. E. D., Staskawicz, B. J. & Rathjen, J. P. (2006) The tomato NBARC-LRR protein Prf interacts with Pto kinase *in vivo* to regulate specific plant immunity. *Plant Cell*, **18**, 2792-2806.

- Nicholas, K.B., Nicholas H.B. Jr., and Deerfield, D.W. II. (1997) GeneDoc: Analysis and visualization of genetic variation, *EMBLNEWS* 4:14.
- Nurnberger, T., Brunner, F., Kemmerling, B. & Piater, L. (2004) Innate immunity in plants and animals: striking similarities and obvious differences. *Immunological Reviews*, **198**, 249-266.
- Ogura, T., Whiteheart, S. W. & Wilkinson, A. J. (2004) Conserved arginine residues implicated in ATP hydrolysis, nucleotide-sensing, and inter-subunit interactions in AAA and AAA+ ATPases. *Journal of Structural Biology*, **146**, 106-112.
- Ok, S. H., Jeong, H. J., Bae, J. M., Shin, J. S., Luan, S. & Kim, K. N. (2005) Novel CIPK1-associated proteins in Arabidopsis contain an evolutionarily conserved C-terminal region that mediates nuclear localization. *Plant Physiology*, **139**, 138-150.
- Padmanabhan, M., Cournoyer, P. & Dinesh-Kumar, S. P. (2009) The leucine-rich repeat domain in plant innate immunity: a wealth of possibilities. *Cell Microbiol.*, **11**, 191-198.
- Pan, Q., Wendel, J. & Fluhr, R. (2000) Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes *Nature Immunology*, **7**, 1243-1249.
- Panstruga, R. & Dodds, P. N. (2009) Terrific protein traffic: The mystery of effector protein delivery by filamentous plant pathogens. *Science*, **324**, 748-750.
- Peart, J. R., Mestre, P., Lu, R., Malcuit, I. & Baulcombe, D. C. (2005) NRG1, a CC-NB-LRR protein, together with N, a TIR-NB-LRR protein, mediates resistance against tobacco mosaic virus. *Current Biology*, **15**, 968-973.
- Rairdan, G. & 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 signaling. *Plant Cell*, **20**, 739-751.
- 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.
- Rathjen, J. P., Chang, J. H., Staskawicz, B. J. & Michelmore, R. W. (1999) Constitutively active Pto induces a Prf-dependent hypersensitive response in the absence of avrPto. *Embo Journal*, **18**, 3232-3240.
- 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.
- Rivas, S. & Thomas, C. M. (2005) Molecular interactions between tomato and the leaf mold pathogen *Cladosporium fulvum*. *Annu. Rev. Phytopathol.*, **43**, 395-436.
- Ron, M. & Avni, A. (2004) The receptor for the fungal elicitor ethylene-inducing Xylanase is a member of a resistance-like gene family in tomato. *Plant Cell*, **16**, 1604-1615.
- Ryals, J. A., Neuenschwander, U. H., Willits, M. G., Molina, A., Steiner, H. Y. & Hunt, M. D. (1996) Systemic acquired resistance. *Plant Cell*, **8**, 1809-1819.
- Salmeron, J. M., Oldroyd, G. E. D., Rommens, C. M. T., Scofield, S. R., Kim, H. S., Lavelle, D. T., Dahlbeck, D. & Staskawicz, B. J. (1996) Tomato Prf is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the Pto kinase gene cluster. *Cell*, **86**, 123-133.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A laboratory manual*, New York, Cold Spring Harbour Laboratory Press.
- Saraste, M., Sibbald, P. R. & Wittinghofer, A. (1990) The P-loop - a common motif in ATP-binding and GTP-binding proteins. *Trends in Biochemical Sciences*, **15**, 430-434.
- Schmidt, S. (2002) Analysis of the M flax rust resistance gene, its transcripts and protein products. *Biological Sciences*. Adelaide, Flinders University.
- Schmidt, S., Lombardi, M., Gardiner, D. M., Ayliffe, M. & Anderson, P. A. (2007a) The M flax rust resistance pre-mRNA is alternatively spliced and contains a complex upstream untranslated region. *Theoretical and Applied Genetics*, **115**, 373-382.

- Schmidt, S. A., Williams, S. J., Wang, C. I. A., Sornaraj, P., James, B., Kobe, B., Dodds, P. N., Ellis, J. G. & Anderson, P. A. (2007b) Purification of the M flax-rust resistance protein expressed in *Pichia pastoris*. *Plant Journal*, **50**, 1107-1117.
- Schwessinger, B. & Zipfel, C. (2008) News from the frontline: recent insights into PAMP-triggered immunity in plants. *Current Opinion in Plant Biology*, **11**, 389-395.
- Scofield, S. R., Tobias, C. M., Rathjen, J. P., Chang, J. H., Lavelle, D. T., Michelmore, R. W. & Staskawicz, B. J. (1996) Molecular basis of gene-for-gene specificity in bacterial speck disease of tomato. *Science*, **274**, 2063-2065.
- Shen, Q. H., Saijo, Y., Mauch, S., Biskup, C., Bieri, S., Keller, B., Seki, H., Ulker, B., Somssich, I. E. & Schulze-lefert, P. (2007) Nuclear activity of MLA immune receptors links isolate-specific and basal disease-resistance responses. *Science*, **315**, 1098-1103.
- Shen, Q. H. & Schulze-Lefert, P. (2007) Rumble in the nuclear jungle: Compartmentalization, trafficking, and nuclear action of plant immune receptors. *Embo Journal*, **26**, 4293-4301.
- Shirano, Y., Kachroo, P., Shah, J. & Klessig, D. F. (2002) A gain-of-function mutation in an Arabidopsis Toll Interleukin1 receptor-nucleotide binding site-leucine-rich repeat type R gene triggers defense responses and results in enhanced disease resistance. *Plant Cell*, **14**, 3149-3162.
- Sohn, J. r., Voegelé, R. T., Mendgen, K. & Hahn, M. (2000) High level activation of vitamin B1 biosynthesis genes in haustoria of the rust fungus *Uromyces fabae*. *Molecular Plant-Microbe Interactions*, **13**, 629-636.
- 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.
- Street, T. O., Bolen, D. W. & Rose, G. D. (2006) A molecular mechanism for osmolyte-induced protein stability. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 13997-14002.
- Swiderski, M. R., Birker, D. & Jones, J. D. G. (2009) The TIR domain of TIR-NB-LRR resistance proteins is a signaling domain involved in cell death induction. *Molecular Plant-Microbe Interactions*, **22**, 157-165.
- Takeda, K. & Akira, S. (2005) Toll-like receptors in innate immunity. *International Immunology*, **17**, 1-14.
- Takken, F. L. W., Albrecht, M. & Tameling, W. I. L. (2006) Resistance proteins: Molecular switches of plant defense. *Current Opinion in Plant Biology*, **9**.
- Takken, F. L. W. & Tameling, W. I. L. (2009) To nibble at plant resistance proteins. *Science*, **324**, 744-746.
- Tameling, W. I. L. & Baulcombe, D. C. (2007) Physical association of the NB-LRR resistance protein Rx with a ran GTPase-activating protein is required for extreme resistance to Potato virus X. *Plant Cell*, **19**, 1682-1694.
- Tameling, W. I. L., Elzinga, S. D. J., Darmin, P. S., Vossen, J. H., Takken, F. L. 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**, 2929-2939.
- 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 Physiology*, **140**, 1233-1245.
- Tang, X., Frederick, R. D., Zhou, J., Halterman, D. A., Jia, Y. & Martin, G. B. (1996) Initiation of plant disease resistance by physical interaction of AvrPto and Pto kinase. *Science*, **274**, 2060-2063.
- Tao, Y., Xie, Z. Y., Chen, W. Q., Glazebrook, J., Chang, H. S., Han, B., Zhu, T., Zou, G. Z. & Katagiri, F. (2003) Quantitative nature of Arabidopsis responses during compatible and incompatible interactions with the bacterial pathogen *Pseudomonas syringae*. *Plant Cell*, **15**, 317-330.
- 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.

- Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994) Clustal-W- Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673-4680.
- 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.
- Towbin, H., Staehelin, T. & Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. U. S. A.*, **76**, 4350-4354.
- Traut, T. W. (1994) The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide binding sites. *European Journal of Biochemistry*, **222**, 9-19.
- Tuskan, G. A., DiFazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R. R., Bhalerao, R. P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G.-L., Cooper, D., Coutinho, P. M., Couturier, J., Covert, S., Cronk, Q., Cunningham, R., Davis, J., Degroove, S., Dejardin, A., dePamphilis, C., Detter, J., Dirks, B., Dubchak, I., Duplessis, S., Ehling, J., Ellis, B., Gendler, K., Goodstein, D., Gribskov, M., Grimwood, J., Groover, A., Gunter, L., Hamberger, B., Heinze, B., Helariutta, Y., Henrissat, B., Holligan, D., Holt, R., Huang, W., Islam-Faridi, N., Jones, S., Jones-Rhoades, M., Jorgensen, R., Joshi, C., Kangasjarvi, J., Karlsson, J., Kelleher, C., Kirkpatrick, R., Kirst, M., Kohler, A., Kalluri, U., Larimer, F., Leebens-Mack, J., Leple, J.-C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D. R., Nelson, C., Nieminen, K., Nilsson, O., Pereda, V., Peter, G., Philippe, R., Pilate, G., Poliakov, A., Razumovskaya, J., Richardson, P., Rinaldi, C., Ritland, K., Rouze, P., Ryaboy, D., Schmutz, J., Schrader, J., Segerman, B., Shin, H., Siddiqui, A., Sterky, F., Terry, A., Tsai, C.-J., Uberbacher, E., Unneberg, P., Vahala, J., Wall, K., Wessler, S., Yang, G., Yin, T., Douglas, C., Marra, M., Sandberg, G., van de Peer, Y. & Rokhsar, D. (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*, **313**, 1596-1604.
- Tyler, B. M., Tripathy, S., Zhang, X., Dehal, P., Jiang, R. H. Y., Aerts, A., Arredondo, F. D., Baxter, L., Bensasson, D., Beynon, J. L., Chapman, J., Damasceno, C. M. B., Dorrance, A. E., Dou, D., Dickerman, A. W., Dubchak, I. L., Gabelotto, M., Gijzen, M., Gordon, S. G., Govers, F., Grunwald, N. J., Huang, W., Ivors, K. L., Jones, R. W., Kamoun, S., Krampis, K., Lamour, K. H., Lee, M.-K., McDonald, W. H., Medina, M., Meijer, H. J. G., Nordberg, E. K., Maclean, D. J., Ospina-Giraldo, M. D., Morris, P. F., Phuntumart, V., Putnam, N. H., Rash, S., Rose, J. K. C., Sakihama, Y., Salamov, A. A., Savidor, A., Scheuring, C. F., Smith, B. M., Sobral, B. W. S., Terry, A., Torto-Alalibo, T. A., Win, J., Xu, Z., Zhang, H., Grigoriev, I. V., Rokhsar, D. S. & Boore, J. L. (2006) Phytophthora genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science*, **313**, 1261-1266.
- Ueda, H., Yamaguchi, Y. & Sano, H. (2006) Direct interaction between the tobacco mosaic virus helicase domain and the ATP-bound resistance protein, N factor during the hypersensitive response in tobacco plants. *Plant Molecular Biology*, **61**, 31-45.
- van der Biezen, E. A. & Jones, J. D. G. (1998a) The NB-ARC domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. *Current Biology*, **8**, R226-R228.
- van der Biezen, E. A. & Jones, J. D. G. (1998b) Plant disease-resistance proteins and the gene-for-gene concept. *Trends in Biochemical Sciences*, **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., Albrecht, M., Cornelissen, B. J. C. & Takken, F. L. W. (2008a) Transcomplementation, but not Physical Association of the CC-NB-ARC and LRR Domains of Tomato R Protein Mi-1.2 is Altered by Mutations in the ARC2 Subdomain. *Mol Plant*, **1**, 401-410.
- van Ooijen, G., Mayr, G., Kasiem, M. M. A., Albrecht, M., Cornelissen, B. J. C. & Takken, F. L. W. (2008b) Structure-function analysis of the NB-ARC domain of plant disease resistance proteins. *Journal of Experimental Botany*, **59**, 1383-1397.
- Verberne, M. C., Hoekstra, J., Bol, J. F. & Linthorst, H. J. M. (2003) Signaling of systemic acquired resistance in tobacco depends on ethylene perception. *Plant Journal*, **35**, 27-32.
- Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982) Distantly related sequences in the  $\alpha$ - and  $\beta$ -subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *Embo Journal*, **1**, 945-51.
- Wang, C. I. A., Guncar, G., Forwood, J. K., Teh, T., Catanzariti, A. M., Lawrence, G. J., Loughlin, F. E., Mackay, J. P., Schirra, H. J., Anderson, P. A., Ellis, J. G., Dodds, P. N. & Kobe, B. (2007) Crystal structures of flax rust avirulence proteins AvrL567-A and -D reveal details of the structural basis for flax disease resistance specificity. *Plant Cell*, **19**, 2898-2912.
- Weaver, L. M., Swiderski, M. R., Li, Y. & 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 Journal*, **47**, 829-840.
- 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.
- Wirthmueller, L., Zhang, Y., Jones, J. D. G. & Parker, J. E. (2007) Nuclear accumulation of the Arabidopsis immune receptor RPS4 is necessary for triggering EDS1-dependent defense. *Current Biology*, **17**, 2023-2029.
- Xu, Y., Tao, X., Shen, B., Horng, T., Medzhitov, R., Manley, J. L. & Tong, L. (2000) Structural basis for signal transduction by the Toll/interleukin-1 receptor domains. *Nature*, **408**, 111-115.
- Yan, N., Chai, J., E, S., Gu, L., Liu, Q., He, J., Wu, J., 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.
- Zhang, Y., Dorey, S., Swiderski, M. & Jones, J. D. G. (2004) Expression of RPS4 in tobacco induces an AvrRps4-independent HR that requires EDS1, SGT1 and HSP90. *Plant Journal*, **40**, 213-224.
- Zipfel, C. (2008) Pattern-recognition receptors in plant innate immunity. *Current Opinion in Immunology*, **20**, 10-16.
- Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J. D. G., Boller, T. & Felix, G. (2006) Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts Agrobacterium-mediated transformation. *Cell*, **125**, 749-760.
- Zou, H., Li, Y. C., Liu, H. S. & Wang, X. D. (1999) An APAF-1 cytochrome c multimeric complex is a functional apoptosome that activates procaspase-9. *J. Biol. Chem.*, **274**, 11549-11556.