The structure and function of RPW8.1 and RPW8.2, powdery mildew disease resistance proteins from *Arabidopsis thaliana* (L.) Heynh.

Daniel Andrew William Jaggard BA(Hons), MA Cantab

School of Biological Sciences
University of East Anglia
Norwich

A thesis submitted to the University of East Anglia for the degree of Doctor of Philosophy

June 2002

© This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that no quotation from the thesis, nor any information derived therefrom, may be published without the author’s prior, written consent.
Dedicated to my parents, Joan and Keith Jaggard.

In memory of
Joan Hartley (1917 – 2000) and Carmen Olive Jaggard (1922 – 2001)
Acknowledgements

I am indebted to several people for their help and support, both in the laboratory and whilst writing this thesis. My parents have helped and advised me and, without their support I might not have been able to submit this thesis. Dr Shunyuan Xiao (who cloned *RPW8* and manages several other *RPW8* research projects) was always pleased to discuss theories and, share his knowledge of *RPW8* and his extensive technical expertise. Mr Andrew Bottley grew the *A. thaliana* plants which were used to extract DNA for the Southern blots. Dr Brent Emerson did McDonald-Kreitman tests and phylogenic analyses on *RPW8.1* and *RPW8.2* alleles. Mr David Alden, the BIO greenhouse technician took exceptional care of a considerable number of *N. benthamiana* plants, necessary for RPW8 transient expression tests. Dr Rosamund Hembry and Dr Christian Roghi provided essential technical advice on epifluorescence microscopy, used to detect the autofluorescence associated with the *RPW8*-induced hypersensitive response in *N. benthamiana*. The virus-induced gene silencing tests, for essential signal transduction components of the *RPW8*-induced hypersensitive response in *N. benthamiana* were conducted with Dr Jack Peart, in Professor David Baulcombe’s laboratory, at the John Innes Centre. Dr Ian Moore (at the Department of Plant Sciences, Oxford University) did the laser scanning confocal microscopy to subcellularly localise GFP fusions with RPW8.1 and RPW8.2. I would also like to thank Dr David Hanke and Dr Ken Korth for their support and guidance. Finally, I am grateful to Dr Mark Coleman and Professor John Turner (my Supervisor and Adviser, respectively) for giving me the opportunity to work on such an interesting project.
Preface

This thesis records the strategies I have employed to investigate RPW8.1 and RPW8.2 protein structures and function.

Research into \textit{RPW8} proteins poses four questions.

What are the structures of RPW8.1 and RPW8.2 proteins?

How does the structure of RPW8.1 and RPW8.2 proteins relate to their function?

What are the functional requirements of RPW8.1 and RPW8.2?

Where are RPW8.1 and RPW8.2 proteins located within the cell?

In order to answer these questions I:

— expressed RPW8.1 and RPW8.2 in bacterial and yeast systems, in an attempt to obtain sufficient protein for X-Ray crystallography and the co-immunoprecipitation of interacting proteins

— sequenced \textit{RPW8.1} and \textit{RPW8.2} alleles from ecotypes of \textit{A. thaliana} to find nucleotide and amino acid polymorphisms, for insights into structural conservation in RPW8.1 and RPW8.2 predicted polypeptides that may imply functional constraints on the sequence

— transiently expressed combinations of RPW8.1 and RPW8.2 in tobacco to assess their potentials to generate hypersensitive response cell death, search for signal transduction components and find their subcellular localisations
Abstract

*RPW8* is a broad-spectrum powdery mildew disease resistance gene locus in *Arabidopsis thaliana* (Xiao *et al.*, 2001). The *RPW8* locus contains two genes, *RPW8.1* and *RPW8.2*, encoding two unique predicted polypeptides, RPW8.1 and RPW8.2, predicted to have coiled-coils and trans-membrane helices (Xiao *et al.*, 2001). However, the trans-membrane helix predictions were probably incorrect and a secretory signal peptide was predicted in RPW8.2. Analysis of the predicted polypeptides of *RPW8.1* and *RPW8.2* alleles from various resistant and susceptible *Arabidopsis thaliana* ecotypes indicated that amino acid polymorphisms T64S and D116G in RPW8.2 and S40I in RPW8.1 were responsible for susceptibility. Interestingly, T64 was predicted to be a yin-yang regulation site. Transient expression of the *RPW8* locus in *Nicotiana benthamiana* generated an *RPW8*-induced, avirulence-independent hypersensitive response. This was *Sgt1*-dependent, did not require protein kinase C or cyclic nucleotide-dependent protein kinases and was taxonomically restricted to *Nicotiana* sp.. Transient expression tests in *Nicotiana benthamiana* showed that RPW8.1 and RPW8.2 were both necessary, but probably not sufficient for the *RPW8*-induced, avirulence-independent hypersensitive response; suggesting a third component, Factor X, was also necessary. Factor X could be encoded by *RPW8fx*, a putative open reading frame at the *RPW8* locus. RPW8.1 was localised in the cytoplasm and nucleoplasm and RPW8.2 was probably localised in the apoplast; the expected site of elicitor recognition of biotrophic fungal pathogens. Distinct RPW8.1 and RPW8.2 subcellular localisations suggests they perform different functions, which could explain why both were necessary for the *RPW8*-induced, avirulence-independent hypersensitive response in *Nicotiana benthamiana*. 
I: Cloned plant disease resistance genes: structure, function and evolution

I.1: Plant disease resistance

I.1.1: The hypersensitive response and autofluorescence

I.1.2: The gene-for-gene hypothesis

I.2: The structures of cloned resistance gene predicted polypeptides

I.2.1: secretory signal peptides and signal anchors

I.2.2: Trans-membrane helices

I.2.3: Coiled-coils

I.2.4: Leucine zippers

I.2.5: Toll/interleukin-1 receptor domains

I.2.6: Nucleotide binding sites

I.2.7: Nucleotide binding domain shared by Apaf-1, certain R gene products and CED-4 domains

I.2.8: Leucine rich repeat domains

I.2.9: WRKY domains

I.2.10: Protein kinase domains

I.2.11: Glycosylation

I.2.12: Myristoylation

I.3: Resistance protein function

I.3.1: Elicitor recognition

I.3.1.1: The site of elicitor recognition

I.3.1.2: Resistance protein subcellular localisation

I.3.1.3: Direct and indirect elicitor recognition

I.3.1.4: Elicitor recognition specificity

I.3.2: Plant disease resistance protein signal transduction

I.3.2.1: Restricted taxonomic functionality

I.3.2.2: Signal transduction pathway components

I.3.2.2.1: Signal transduction pathways initiated by
I.3.2.2.2: Mitogen-activated protein kinase and calcium-dependent protein kinase signal transduction pathway components ........................................... 39
I.3.2.2.3: EDS1 and NDR1 signal transduction pathways ........................................... 43
I.3.2.2.4: Unidentified signal transduction pathways... 48
I.3.2.2.5: RAR1 & SGT1 signal transduction pathway 48
I.3.2.2.6: Resistance protein mediated synthesis of the resistance protein and the signal transduction pathway components ........................................... 51

I.4: Cloned resistance gene evolution ........................................... 52
I.4.1: Duplication and resistance gene clusters ................................. 52
I.4.2: Rearrangement and exchange within resistance gene clusters... 52
I.4.3: Transposon mutations and their secondary evolutionary effects 56
I.4.4: Insertion and deletion ........................................... 57
I.4.5: Positive, diversifying, adaptive selection of resistance gene sequences ........................................... 58
I.4.6: Resistance gene recognition specificity evolution: the “boom and bust” cycle ........................................... 60
I.4.7: Resistance gene relationships ........................................... 63

I.5: Summary ........................................... 65

II: Powdery mildew disease resistance genes in Arabidopsis thaliana (L.) Heynh ........................................... 66
II.1: Powdery mildews ........................................... 66
II.2: Arabidopsis as a model plant for studying plant pathogen interactions... 68
II.3: Resistance to powdery mildews in Arabidopsis thaliana .................. 69
II.4: RPW8 ........................................... 73
II.4.1: Map-based cloning of RPW8 ........................................... 73
II.4.2: RPW8-mediated broad spectrum disease resistance ........... 75
III: Structural predictions of RPW8.1 and RPW8.2 predicted polypeptides

III.1: Coiled-coil predictions

III.2: Transmembrane domain predictions

III.3: Predicted subcellular targeting of the RPW8 predicted polypeptides

III.4: Predicted post-translational modification of the RPW8 predicted polypeptides

III.5: RPW8 predicted polypeptide tertiary structure predictions

III.6: Conclusion

IV: Expression of RPW8.1 and RPW8.2 proteins in *Escherichia coli*

IV.1: Introduction

IV.2: Methods

IV.2.1: RPW8.1 expression test

IV.2.2: The solubility of expressed RPW8.1

IV.2.3: The effects of temperature and inducer concentration on the solubility of RPW8.1

IV.2.4: RPW8.2 expression and solubility test

IV.3: Results

IV.3.1: RPW8.1 expression test

IV.3.2: The solubility of expressed RPW8.1

IV.3.3: The effects of temperature and inducer concentration on the solubility of RPW8.1

IV.3.4: RPW8.2 expression and solubility test

IV.4: Discussion

V: Expression of RPW8.1 and RPW8.2 proteins in *Schizosaccharomyces pombe*

V.1: Introduction

V.2: Methods

V.3: Results
VI: Polymorphism of *RPW8.1* and *RPW8.2* allele sequences in ecotypes of *Arabidopsis thaliana*

VI.1: Introduction

VI.2: Methods

VI.2.1: Sequencing genomic *RPW8* alleles

VI.2.2: Southern blot

VI.3: Results & Discussion

VI.3.1: *RPW8* allele sequences

VI.3.1.1: Insertions and deletions

VI.3.1.2: Single nucleotide polymorphisms

VI.3.1.3: Duplication and deletion of *RPW8.1* and *RPW8.2* alleles

VI.3.1.4: *RPW8* homologues

VI.3.2: RPW8 predicted polypeptides

VI.3.3: Conclusion

VII: Transient expression of the *RPW8* locus generated the hypersensitive response in *Nicotiana benthamiana*

VII.1: Introduction

VII.2: Methods

VII.2.1: Transient expression of the *RPW8* locus in *Nicotiana benthamiana*

VII.2.2: Histochemical detection of transient expression in *Nicotiana benthamiana* with β-glucuronidase

VII.2.3: Epifluorescence microscopical detection of autofluorescence

VII.2.4: Transient expression of the *RPW8* locus in various species

VII.3: Results & Discussion

VII.3.1: RPW8.1 and RPW8.2 are necessary for the *RPW8*-induced,
avirulence-independent hypersensitive response in *Nicotiana benthamiana* ..........................................................................................................................174

VII.3.2: RPW8.1 and RPW8.2 are probably not sufficient for the *RPW8*-induced, avirulence-independent hypersensitive response in *Nicotiana benthamiana* ..........................................................................................................................177

VII.3.3: The restricted taxonomic functionality of *RPW8*........................182

VII.3.4: Conclusion ..............................................................................................183

VIII: Complete suppression of the *RPW8*-induced hypersensitive response by virus-induced gene silencing of *Sgt1* ..........................................................................................................................184

VIII.1: Introduction ..............................................................................................184

VIII.2: Methods .....................................................................................................186

VIII.3: Results ........................................................................................................187

VIII.4: Discussion ..................................................................................................188

IX: Expression and subcellular localisation of RPW8.1 and RPW8.2 proteins in *Nicotiana benthamiana* ..........................................................................................................................190

IX.1: Introduction ..................................................................................................190

IX.2: Methods ........................................................................................................194

IX.2.1: Cloning GFPs and RPW8 cDNAs into binary vector T-DNAs ..........................................................................................................................194

IX.2.2: Transient expression of RPW8.1::EGFP and RPW8.2::ECFP and detection of EGFP and ECFP fluorescence in *Nicotiana benthamiana* ..........................................................................................................................211

IX.3: Results & Discussion ....................................................................................213

IX.3.1: Subcellular localisation of RPW8.1::EGFP ............................................213

IX.3.1.1: RPW8.1::EGFP was cytoplasmic and nucleoplasmic .........................213

IX.3.1.2: Bright-spots ............................................................................................215

IX.3.2: Subcellular localisation of RPW8.2::ECFP ............................................220

IX.3.2.1: RPW8.2::ECFP was probably extracellular ........................................220
IX.3.2.2: Conclusively proving whether RPW8.2 is extracellular………….221
IX.3.3: Validating RPW8.1 and RPW8.2 subcellular localisations............222

X: General Discussion............................................................................223
X.1: The functions of RPW8 proteins.................................................223
X.2: The structures of RPW8 proteins................................................224

References.............................................................................................226
Plant disease resistance is the ability to prevent a pathogen from invading, colonising and reproducing in a host plant. Plant disease (susceptibility) is the exception, not the rule, despite the vast range of potential pathogens (Dangl & Holub, 1997; Hammond-Kosack & Jones, 1996; Jackson & Taylor, 1996; Keen, 1999; Lamb et al., 1989; Staskawicz et al., 1995; Staskawicz, 2001; Vanderplank, 1984). Thus, plants are resistant to the majority of potential pathogens, although resistance (incompatibility) to a particular pathogen can be due to “non-host” or “host-specific” defences.

Non-host resistance refers to interactions where the potential pathogen species is unable to cause disease on a certain species of host plant (Heath, 2000; Lamb et al., 1989), because the pathogen lacks basic compatibility. Basic compatibility is the ability of a pathogen to successfully invade, colonise and reproduce on certain host plants by eliminating, overcoming, avoiding or escaping from static host defences using methods conferred by pathogenicity determinants (de Wit, 1992; Gabriel & Rolfe, 1990; Keen, 1982; Keen, 1992; Lamb et al., 1989; Staskawicz, 2001). Static host defences (such as preformed and constitutive anti-microbial compounds, the cell wall or the cuticle) are effective chemical and physical barriers against most pathogens (Bell, 1981; Dangl & Holub, 1997; Keen, 1992; Keen, 1999; Knogge, 1996; Newton & Andrivon, 1995; Osbourn, 1996; Walton, 1994). Non-host pathogens have gained basic compatibility by acquiring pathogenicity determinants (compatibility factors), e.g. genes for a cutinase (Dickman et al., 1989) or coronatine synthesis (Mittal & Davis, 1995).
Host-specific resistance occurs where a particular species of pathogen is capable of causing disease on a certain host species but interactions between specific races of the pathogen and specific cultivars of the host are incompatible. Host-specific resistance is inducible and composed of active defence responses, superimposed upon basic compatibility in a gene-for-gene manner (Bonas & Van den Ackerveken, 1999; de Wit, 1992; Gabriel & Rolfe, 1990; Keen, 1982; Keen, 1990; Staskawicz et al., 1995).

Resistance (R) genes in the plant and avirulence (Avr) genes in the pathogen determine the gene-for-gene specificity of host-specific resistance.

I.1.1: The hypersensitive response and autofluorescence

Resistance is often accompanied by small necrotic lesions at the site of pathogen attack. Stakman (1915) first observed the association of rapid host cell death in the vicinity of *Puccinia graminis* invasions and used the term “hypersensitive” to define this extreme incompatible interaction.

The hypersensitive response (HR) is a rapid and localised programmed cell death at the site of pathogen attack, usually associated with non-host and host-specific active defences (Atkinson, 1993; Dangl et al., 1996; Dixon et al., 1994; Gilchrist, 1998; Goodman & Novacky, 1994; Greenberg, 1996; Heath, 1998; Heath, 1999; Heath, 2000; Hutcheson, 1998; Jones & Dangl, 1996; Keen, 1992; Keen, 1999; Lam et al., 1999; Morel & Dangl, 1997; Pennell & Lam, 1997; Pontier et al., 1998; Richberg et al., 1998; Scheel, 1998; Somssich & Hahlbrock, 1998). Coincident with the HR in responding cells are multiple and complex biochemical changes: Ca$^{2+}$ influx, H$^+$/K$^+$ exchange, generation of reactive oxygen species (ROS, *e.g.* H$_2$O$_2$ and O$_2$•$^-$), protein phosphorylation, salicylic acid accumulation, lignin and callose deposition, gene expression (including pathogenesis-related genes) and phytoalexin biosynthesis (Atkinson, 1993; Dangl, 1995; Dangl et al., 1996; Dangl & Holub, 1997; de Wit, 1992; Dixon & Lamb, 1990; Dixon et al., 1994; Gilchrist, 1998; Greenberg, 1996; Heath, 1998; Heath, 2000;
Jones & Dangl, 1996; Keen, 1992; Keen, 1999; Lam et al., 1999; Lamb et al., 1989; Morel & Dangl, 1997; Pennell & Lam, 1997; Pontier et al., 1998; Richael & Gilchrist, 1999; Richberg et al., 1998; Scheel, 1998; Somssich & Hahlbrock, 1998; Yang et al., 1997). Dead HR cells collapse and, if in sufficiently large groups, subsequently desiccate and melanise to form necrotic lesions (Bell, 1981; Heath, 2000; Goodman & Novacky, 1994). Whether the HR acts as a direct defence mechanism, limiting pathogen growth, is debatable. However, it is responsible for local and systemic acquired resistance (SAR) by triggering systemic gene induction (Dangl et al., 1996; Dangl & Holub, 1997; Graham & Graham, 1999; Heath, 2000; Hutcheson, 1998; Morel & Dangl, 1997; Pontier et al., 1998; Richberg et al., 1998; Scheel, 1998; Staskawicz et al., 1995).

Resistance and the HR are separable. Rx-mediated resistance to potato virus X does not involve the HR (Bendahmane et al., 1999). Resistance and the extent of the HR (from absent to systemic cell death) to CaMV in Nicotiana sp. segregated independently (Cole et al., 2001). The HR is absent in dnd1 (defence, no death 1) A. thaliana plants, which retain (albeit less effective) resistance to Pseudomonas syringae; they even have enhanced resistance to other pathogens, due to increased salicylic acid accumulation (Yu et al., 1998; Yu et al., 2000). Pc82-mediated HR, but not resistance, is suppressed by the Rds locus in Avena sp. challenged with Puccinia coronata f. sp. avenae (Yu et al., 2001). The Rih (resistance-independent HR) locus in Avena plants causes HR in resistant (Pc82 present) and susceptible (Pc82 absent) interactions (Yu et al., 2001).

Cell death is not always associated with resistance. Cell death can occur in compatible interactions too, but as the result of disease (Heath, 2000; Goodman & Novacky, 1994; Lam et al., 1999). Cell death induced by the HR differs from disease induced cell death in speed and scope; the HR is extremely fast (identifiable microscopically within 24 hours) and limited to just a few selected host cells (Goodman & Novacky, 1994).

Phytoalexins are a chemically diverse group of phenolic secondary metabolites whose de novo biosynthesis is induced by pathogen attack (Bell, 1981; Darvill & Albersheim, 1984; Dixon et al., 1983; Dixon, 2001; Ebel, 1986; Hammerschmidt, 1999; Kuc & Rush, 1985; Nicholson & Hammerschmidt, 1992; Smith, 1996; VanEtten et al., 1994). Paxton (1981) defined phytoalexins as, “low molecular weight antimicrobial compounds that are both synthesised by and accumulated in
plants after exposure to microorganisms”. High concentration phytoalexin accumulation (Nicholson & Hammerschmidt, 1992) occurs late in the HR just before necrosis (Bell, 1981). Hence, the accumulation of phytoalexins to high concentrations is a natural indicator of the HR later stages.

An intense UV autofluorescence is indicative of HR cell death (Bent, 1996; Dempsey et al., 1997; Érsek et al., 1981; Grundler et al., 1997; Heath, 1999; Heath, 2000; Hinrichs-Berger et al., 1999; Holliday et al., 1981; Jones, 1994; Koga et al., 1980; Mayama & Shishiyama, 1978; Nicholson & Hammerschmidt, 1992) and has been used to specifically identify HR cell death in plant-pathogen interactions (Belefant-Miller et al., 1994; Koga, 1994; Koga et al., 1988; Somssich & Hahlbrock, 1998; Stadnik & Buchenauer, 1999; Tang et al., 1996; Yu et al., 1998; Yu et al., 2001; Zeyen et al., 1995). Copeman (1969) spectrophotometrically identified scopoletin (a phytoalexin) as the autofluorescent material accumulating in the xylem, xylem parenchyma and endodermis of Nicotiana tabacum infected with Pseudomonas solanacerum. Therefore an intense UV autofluorescence of accumulated scopoletin in Nicotiana cells is a natural and specific indicator of HR cell death.

I.1.2: The gene-for-gene hypothesis

The gene-for-gene hypothesis was proposed by Flor (1942) to explain the “interaction between factors for pathogenicity [dominant Avr genes] in the pathogen and those for resistance [dominant R genes] in the host” in host-specific interactions. The gene-for-gene hypothesis states that an incompatible interaction will occur between the host and the pathogen if the pathogen is carrying an Avr gene and the host contains the corresponding R gene (Table I.1). Absence of either the host R gene or the corresponding pathogen Avr gene permits a compatible interaction.

Table I.1: A quadratic of the potential host and pathogen genotype combinations for an R gene and its corresponding Avr gene, indicating the gene-for-gene specificity that determines host-specific resistance.

<table>
<thead>
<tr>
<th>Pathogen genotype</th>
<th>Host genotype</th>
<th>Avr</th>
<th>Incompatible</th>
<th>Compatible</th>
<th>Compatible</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td></td>
<td></td>
<td>Compatible</td>
<td>Compatible</td>
</tr>
<tr>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The elicitor-receptor model (Keen, 1982; Keen, 1990) is the most plausible molecular representation of the gene-for-gene hypothesis. It is an extension of the protein-for-protein hypothesis (Vanderplank, 1984), which assumes that R and Avr, gene-for-gene specificity is determined by the interaction of their
protein products (directly or indirectly). In the elicitor-receptor model (Figure I.1) Avr gene products or metabolites resulting from their activity (elicitors) are recognised by R gene products or associated proteins (receptors).

All genetic factors responsible for resistance in a gene-for-gene manner have been reported as R genes. However, the resistance mediated by some of these genetic factors (e.g. the HM1, T-urf13 and Vb genes) does not involve active defence responses. Also the resistance mediated by genes that do not involve active defence responses has a genetic interaction (Table I.2) the opposite of R-Avr gene interactions where active defence responses are induced (Briggs & Johal, 1994; Hammond-Kosack & Jones, 1997).

Table I.2: A quadratic of the potential host and pathogen genotype combinations for an-R gene that does not indue active defence responses and its corresponding pathogenicity determinant (Path D) in the pathogen. Note that the loss of the R gene (r) results in resistance to a virulent (Path D<sup>+</sup> pathogen).

<table>
<thead>
<tr>
<th>Pathogen genotype</th>
<th>Path D+</th>
<th>path D-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host genotype</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>Compatible</td>
<td>Incompatible</td>
</tr>
<tr>
<td>r</td>
<td>Incompatible</td>
<td>Incompatible</td>
</tr>
</tbody>
</table>

Cloned R genes that do not indue active defence responses are detoxifying enzymes or toxin target proteins. HM1 codes for HC toxin reductase and confers resistance in maize to Cochliobolus carbonum, a pathogen which requires the HC toxin for infection (Johal & Briggs, 1992). T-urf13 encodes URF13, a protein responsible for sensitivity to Bipolaris maydis T-toxin; mutation of URF13 dicyclohexylcarbodiimide binding sites eliminates T-toxin sensitivity (Braun et al., 1989). The Vb locus controls sensitivity to victorin, a toxin produced by Cochliobolus victoriae (Wolpert et al., 1995), and probably encodes the 100 kDa victorin target protein (Wolpert & Macko, 1989). Therefore, R genes that do not indue active defence responses function by preventing pathogens from using pathogenicity determinants to gain basic compatibility, whereas R genes initiate active defence responses against pathogens already possessing basic compatibility. R genes that do not indue active defence responses are not considered further in this review.
I.2 : The structures of cloned resistance gene predicted polypeptides

Forty R genes have been cloned (Table I.3a & b) and, except Cre3 and P2, have been shown to confer resistance. No R protein structures have been determined, but the predicted polypeptides of cloned R genes are known. Surprisingly, considering the cloned R genes confer resistance to a wide array of pathogens (aphids, bacteria, fungi, nematodes and viruses) there are often shared structural motifs between R gene predicted polypeptides (Figure I.4a & b). These structural motifs are secretory signal peptide (SP), signal anchor (SA), trans-membrane (TM), coiled-coil (CC), leucine zipper (LZ), Toll/interleukin-1 receptor (TIR), nucleotide binding site (NBS), nucleotide binding domain shared by Apaf-1, certain R gene products and CED-4 (NB-ARC), leucine-rich repeats (LRR), WRKY and protein kinase (PK) domains.

Analysis of deletion or termination codon mutants of several R genes has shown that there are no large dispensable regions of R gene predicted polypeptides (Axtell et al., 2001; Dinesh-Kumar et al., 2000; Dodds et al., 2001; Grant et al., 1995; McDowell et al., 1998; Mindrinos et al., 1994; Tao et al., 2000; Thomas et al., 2000; Van der Hoorn et al., 2001; Wulf et al., 2001). However, L6, MLA6, N, RPS4 and Rpl-D genes are alternatively spliced, potentially producing different predicted polypeptides that may be functional, although the effect of alternative splicing on resistance has only been studied for the N and L6 genes. Susceptible flax plants transformed with full length L6 cDNA – intronless and thus incapable of alternative splicing – gained complete resistance to rust fungus carrying A-L6 (Ayliffe et al., 1999). The N gene produces two alternatively spliced N_s (short) and N_l (long) transcripts (Dinesh-Kumar & Baker, 1999). The N_s transcript is predicted to encode the full-length protein whereas N_l is predicted to encode a truncated protein lacking the C terminus, including 13 of the 14 LRRs. The transcript ratio between N_s and N_l varies (without the total transcript abundancy changing) during TMV infection, with the N_l transcript only present at greater abundancy than the N_s between four and eight hours post TMV inoculation (Dinesh-Kumar & Baker, 1999). Presumably the ratio of N_s/N_l abundancy is directly proportional to the ratio of protein concentrations. Surprisingly, susceptible tobacco plants transformed with cDNA of the N_s transcript, the cDNA of the N_l transcript, or both cDNAs were all only partially resistant to TMV (Dinesh-Kumar & Baker, 1999). Complete resistance
was achieved with cDNA of the Ns transcript modified to include the alternatively spliced third intron sequence and 1.4 kb of the 3’ UTR – necessary for and regulating transcript alternative splicing (Dinesh-Kumar & Baker, 1999). Therefore the N gene requires both protein products at different relative abundancies during infection to combat viral attack by TMV, whereas a full length L6 transcript is sufficient for complete resistance to rust fungus. Hence, all of the R predicted polypeptides are required for function, although other R predicted polypeptides could be functional.

The structural motifs predicted in R gene predicted polypeptides (Figure I.4a & b), here and in the literature, must not be thought of as rigidly fixed. Some structural predictions were made years ago and with prediction algorithms which may now be obsolete. For example Ellis & Jones (1998) comment that the prediction by Cai et al. (1997) of a SP and a TM helix in Hs1pro-1 are too short and contain charged, polar residues. Indeed TMHMM2.0 and SignalP2.0 algorithms did not find TM helix or SP motifs, respectively, in Hs1pro-1. Cai et al. (1997) also predicted a LRR motif in Hs1pro-1 even though it is not sufficiently homologous to LRR proteins in the databases. Therefore, I have re-assessed the likelihood of all R gene predicted polypeptides containing the structural motifs claimed by authors in the literature with the consensi for SP, SA, TM, CC, LZ, TIR, NBS, NB-ARC, LRR, WRKY and PK motifs.

1.2.1: Secretory signal peptides and signal anchors

Secretory signal peptides (SPs) and signal anchors (SAs) both target eukaryotic proteins to the endoplasmic reticulum (ER), but differ in their primary structure. A SP is a 15-30 residue sequence at the N terminus that targets proteins to the ER and subsequently the secretory pathway (Gierasch, 1989; von Heijne, 1985; von Heijne, 1990). Once the protein is translocated to the ER the SP is cleaved off and the mature protein is released into the lumen of the ER. A SA also targets proteins to the ER but the N terminus is not cleaved, thereby anchoring the protein to the membrane (a type II membrane protein) with the C terminus in the ER lumen (von Heijne, 1990).
I.2.2: Trans-membrane helices

Transmembrane (TM) helices are hydrophobic, α-helices of approximately 20 residues in length – the number necessary to span the approximately 30Å width of the lipid bilayer membrane (Engelman et al., 1986).

I.2.3: Coiled-coils

Coiled-coils (CCs) are bundles of two to five solvent-exposed α-helices, with a left-handed superhelical twist, that wind around each other in parallel forming a superhelix – literally a coiled-coil. They were proposed by Crick (1952) and Pauling et al. (1951) in α-keratin. Since then CCs have been discovered in fibrous proteins such as keratin, myosin and fibrinogen; more recently they have been found in globular proteins like the bZIP transcription factors – for a recent structural review of CCs see Lupas (1996). A superhelical twist means that CCs have 3.5 residues per turn and therefore heptad periodicity of residue physico-chemical properties at equivalent positions in the helix. Of the seven positions in the heptad repeat, labelled a-g, positions a and d are typically hydrophobic and form the “core” of the helix-helix interface. Interactions at the helix-helix interface have a distinctive side-chain packaging geometry, known as “knobs-into-holes” (Crick, 1953), illustrated in Figure I.5. Positions b, c, e, f and g of the CC are typically occupied by hydrophilic residues, forming the solvent-exposed “coat”. Substituted residues were found in the RPM1 predicted polypeptide of non-functional sequences (Tornero et al., 2002).

I.2.4: Leucine zippers

A leucine zipper (LZ) motif is a parallel CC of approximately 30 amino acids in each α-helix with leucine residues in a heptad repeat (Alber, 1992; Ellenberger et al., 1992; O’Shea et al., 1989; O’Shea et al., 1991). Homo or heterodimers of LZs are formed (Stryer, 1995) by stable association of the helices within a CC – the LZ of GCN4 homodimerises between 10µM-2mM (O’Shea et al., 1989). LZs are found in transcription factors where there is a DNA binding, basic region of approximately 30 residues N terminal to the LZ motif (Alber, 1992). For example, in GCN4 the basic region forms an α-
helix when in contact with the DNA major groove (Ellenberger et al., 1992). Basic regions in LZ transcription factors have the 16 amino acids consensus sequence (K/R)RxR-NxxAXx(R/K)xRx(R/K)xRx(R/K)xRx(R/K) where invariant residues are underlined and the leucine is the first in the heptad repeat. Basic regions were not found N terminal to the LZs in HRT, R1, RPP8, RPS2 and RPS5 predicted polypeptides, implying that these predicted LZs do not bind DNA.

1.2.5: Toll/interleukin-1 receptor domains

The TIR domain corresponds to the cytoplasmic regions of interleukin-1 receptors (IL-1RI, Sims et al., 1989) and the Drosophila receptor Toll (Hashimoto et al., 1988), which are 26% identical and 43% similar (Gay & Keith, 1991).

The TIR domain is necessary for Toll and IL-1RI signal transduction; with functional sites conserved between the predicted polypeptides. Amino acid regions 508-521 (Croster et al., 1995) and, 515-529 and 364-474 (Kuno et al., 1993) of the IL-1RI cytoplasmic region – residues 335-552 – were necessary for function. Also substitution by Heguy et al. (1992) of Arg431, Phe513, Trp514, Lys515, Tyr517 and Arg518 IL-1RI residues abolished function; substitution of Pro521 decreased the functional ability. The Toll cytoplasmic region mutants F859I, H893Y and V911M were non functional (Schneider et al., 1991). Substitution of non-conserved residues in IL-1RI by Heguy et al. (1992) did not affect function. Therefore residues common to the cytoplasmic regions of both Toll and IL-1RI are necessary for function and amino acid substitutions in the N predicted polypeptide, corresponding to the H893Y and V911M substitutions, abolished N function (Dinesh-Kumar et al., 2000).

Xu et al. (2000) have determined the structure of the cytoplasmic (TIR) domain from Toll-like receptors 1 and 2 (TLR1 and TLR2). The structures of the TIR domains from TLR1 and TLR2 have a central five stranded parallel β-sheet surrounded by five α-helices (Figure I.6). The BB-loop, which protrudes from the TIR domain and is solvent-exposed, is conserved between the structures of TLR1 and TLR2. The P681H amino acid substitution in the BB-loop of TLR2 did not disrupt its TIR domain structure but did abolish the in vitro binding of TLR2 and MyD88 – a signal transduction pathway component (Figure I.7). Hence, the BB-loop of the TIR domain is probably a contact surface for TIR
domain containing R proteins to interact with homologues of MyD88.

*Toll* and IL-1RI signal transduction cascades are composed of common and homologous downstream effectors (Figure I.7), implying an evolutionary conserved signalling pathway (Belvin & Anderson, 1996; Daun & Fenton, 2000; Hoffman et al., 1999; Means et al., 2000). Hence, TIR domains in R gene predicted polypeptides might transduce the elicitor signal in a plant pathway homologous to the *Toll* and IL-1RI pathways.

1.2.6: Nucleotide binding sites

Direct transfer of the -P from adenosine triphosphate (ATP) or guanosine triphosphate (GTP) to an acceptor requires a nucleotide binding site (NBS) in the acceptor. A NBS is usually composed of three motifs: P-loop, kinase-2 and kinase-3a (Traut, 1994). The ATP and GTP purine base and pentose or phosphates are bound by kinase-3a and P-loop motifs respectively. The kinase-2 motif contains a conserved aspartate (D), necessary to coordinate the divalent metal ion for phosphotransfer. Kinase-2 and kinase-3a motifs are usually found 50-70 and 120-145 amino acids, respectively, C terminally to the P-loop (Traut, 1994). The consensus sequences for the P-loop (Saraste et al., 1990) kinase-2 and kinase-3a motifs (Traut, 1994) are in Table I.4. G4 and G5 motifs are characteristic features of GTPases (Bourne, 1991). R gene predicted polypeptides are more likely to bind ATP than GTP at the NBS. R predicted polypeptides do not have G4 (Meyers et al., 1999) or G5 motifs. Substitution or deletion of amino acids in the P-loop, kinase-2 or kinase-3a motifs in RPS2 or N abolished function (Dinesh-Kumar et al., 2000; Tao et al., 2000), although substitution in the P-loop of RPS2 did not prevent complex formation with AvrRpt2 or AvrB *in planta* (Leister & Katagiri, 2000).

Table I.4: Consensus sequences, in single letter amino acid code, of the P-loop, kinase-2 and kinase-3a NBS motifs (x = any amino acid, h = hydrophobic amino acids such as V, I or L, invariant residues are in bold).

<table>
<thead>
<tr>
<th>NBS Motif</th>
<th>Consensus sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-loop</td>
<td>GxxxxGK(S/T)</td>
</tr>
<tr>
<td>Kinase-2</td>
<td>hhhhD</td>
</tr>
<tr>
<td>Kinase-3a</td>
<td>(E/A/G/P/F)(T/S/G/F)xxx(Y/R)</td>
</tr>
<tr>
<td>G4</td>
<td>NKxD</td>
</tr>
<tr>
<td>G5</td>
<td>HhE(A/C/S/T)SA(K/L)</td>
</tr>
</tbody>
</table>
The NBS core structure was predicted by Milner-White et al. (1991) to be four parallel β-sheet strands sandwiched between four α-helices. Alternatively, Rigden et al. (2000) modelled the NBS domain of R gene predicted polypeptides and concluded that their structure was homologous to the receiver domains of proteins in the histidine-aspartate phosphotransfer pathway, which they claim correlates better with the results of mutational analysis. However, RPS2 binds ATP and GTP (A. Bent, pers. comm.) meaning that R proteins do contain NBSs and are unlikely to be involved in histidine-aspartate phosphotransfer.

I.2.7: Nucleotide binding domain shared by Apaf-1, certain R gene products and CED-4 domains

A region (92-412 amino acids) of Apaf-1 (Zou et al., 1997) was noticed by van der Biezen & Jones (1998) to be 50% similar to CED-4 (Yuan & Horvitz, 1992) and also L6, N, Prf, RPM1, RPP5 and RPS2 predicted polypeptides. This region of similarity was called the nucleotide binding domain shared by Apaf-1, certain R gene products and the CED-4 (NB-ARC) domain by van der Biezen & Jones (1998) included the NBS (kinase-1, kinase-2 and kniase-3a) and five other structural motifs – the hydrophobic domain (HD) and motifs 2-5. Motifs 3 and 4 are weakly conserved and might not play any direct functional role. The HD and motifs 2 and 5 are highly conserved and have the following consensus sequences CxGLPLhL, SYDxL and MHDLh. Amino acid substitutions in the NB-ARC region, although not in NB-ARC motifs, of P2 (Dodds et al., 2001), RPM1 (Grant et al., 1995; Tornerro et al., 2002), RPP8 (McDowell et al., 1998), RPS2 (Axtell et al., 2001; Tao et al., 2000) and RPS4 (Gassman et al., 1999) predicted polypeptides, or deletion of part of the NB-ARC regions in the N predicted polypeptide (Dinesh-Kumar et al., 2000) all abolished function. An amino acid substitution in the HD domain of P2 abolished function (Dodds et al., 2001). Fold-recognition models of NB-ARC domains by Jaroszewski et al. (2000) found α/β structure around the P-loop and then two α-helical regions to the C terminal side of the P-loop (Figure I.8).

Apaf-1 and CED-4 share structural and functional similarities. They both activate apoptotic caspases (cysteine proteases) in mammals and C. elegans respectively (Chinnaiyan et al., 1997a; Zou et al., 1997). Caspase activation by Apaf-1 and CED-4 is inhibited by Bel-2 (Zou et al., 1997) and CED-9 (James et al., 1997), respectively. Bel-2 and CED-9 are structurally similar and Bel-XL (a Bel-2) compliments CED-9 inhibition of CED-4 (Chinnaiyan et al., 1997b). ATP binding at the P-loop and
hydrolysis is required for caspase activation by Apaf-1 and CED-4 (Chaudhary et al., 1998; Chinnaiyan et al., 1997a; James et al., 1997b; Zou et al., 1999). ATP and oligomerisation of Apaf-1 and CED-4 are necessary for caspase activation (Yang et al., 1998; Zou et al., 1999). Amino acids 171-435 in CED-4 (the NB-ARC domain) are required for oligomerisation, although ATP binding isn’t (Yang et al., 1998). Apaf-1 forms self associating oligomers through the NB-ARC region (Srinivasula et al., 1998). CED-9 binds the P-loop of CED-4 (Chaudhary et al., 1998) and competes with CED-4 for CED-4 binding during oligomerisation (Yang et al., 1998). Therefore, Apaf-1 and CED-4 recruit caspases, by oligomerising and then activates them via phosphorylation.

Caspase activation by Apaf-1 and CED-4 functional homologues occurs during active defence responses in plants. Transcription of the potato cysteine protease encoding gene (CYP) was induced during resistance to Phytophthora infestans (Avrova et al., 1999). Cysteine protease activity was found in tissue developing the HR (del Pozo & Lam, 1998; D’Silva et al., 1998; Solomon et al., 1999) and cysteine protease inhibitors prevented the HR in N. tabacum (del Pozo & Lam, 1998) and soybean (Solomon et al., 1999). Over-expression of either Bcl-X_L or CED-9 in N. tabacum suppressed the N-mediated, TMV coat protein-dependent HR and N-mediated resistance to TMV (Mitsuhara et al., 1999). Therefore, NB-ARC domain containing R protein recruited and activated caspases were necessary for the HR and resistance in N. tabacum and soybean.

1.2.8: Leucine-rich repeat domains

Leucine-rich repeats (LRRs) are versatile binding domains involved in protein-protein interactions (Buchanan & Gay, 1996; Kobe & Deisenhofer, 1994, 1995; Kobe & Kajava, 2001). LRR proteins operate in functionally diverse roles, but approximately half are connected to signal transduction (Kobe & Diesenhofer, 1994). Also, LRR proteins are known to bind proteinaceous and non-protein ligands, like lipopolysaccharides and DNA (Kobe & Deisenhofer, 1995).

A LRR domain is composed of tandem repeats of the leucine-rich consensus sequence, LxxLxLxx(N/C)xL, which corresponds to the β-sheets in each repeat unit (Kajava, 1998; Kobe & Deisenhofer, 1995) and between the consensus sequence the variable, interstrand residues forming an α-helical or extended region (Kobe & Kajava, 2001). LRRs are classified by length and consensus sequence into
subfamilies, which imply distinct evolutionary origins (Kajava, 1998); they are now divided into seven subfamilies (Kobe & Kajava, 2001). Plant LRRs belong to cysteine-containing and plant-specific subfamilies with consensus sequences \((C/L)xxLxxLxLxxCxxITDxxoxL(A/G)xx\) and \(LxxLxxLxLxxNxL(I/S)GxIPxxLx\), respectively. Cysteine-containing LRRs (present in RPS2 and L6 predicted polypeptides) are 25-27 residues long and are intracellular (Kajava, 1998). Plant-specific LRRs (present in the predicted polypeptide of Cf-2, Cf-4, Cf-5 and Cf-9) are 23-25 residues long and are extracellular (Kajava, 1998). Table I.5 lists all the LRR containing R gene predicted polypeptides and if possible their subfamily classification. Substitution of amino acids in the LRR consensus sequence or interstrand regions were found in non-functional M (Anderson et al., 1997), N (Dinesh-Kumar et al., 2000), P2 (Dodds et al., 2001), RPM1 (Grant et al., 1995; Tornerro et al., 2002), RPP8 (McDowell et al., 1998); RPS2 (Axtell et al., 2001; Bent et al., 1994; Mindrinos et al., 1994), RPS4 (Gassman et al., 1999) and RPS5 (Warren et al., 1998) predicted polypeptides. Deletion of one or two LRRs from any region in the LRR domain of N (Dinesh-Kumar et al., 2000) or RPS2 (Tao et al., 2000) predicted polypeptides abolished function. Also, an in frame duplication of 90 residues in the LRR domain of RPP5, which extended the predicted LRR domain by four repeats, retained function (Parker et al., 1997).

The structures of 10 LRR proteins have been determined (reviewed by Kobe & Kajava, 2001). The first LRR structure determined was porcine ribonuclease inhibitor (RI), a protein with 15 LRRs of alternating 28 or 29 amino acid repeats (Kobe & Diesenhofer, 1993). RI is non-globular and horseshoe (solenoid) shaped, with a curved solvent-exposed parallel \(\beta\)-sheet lining the (inner) concave face and \(\alpha\)-helices flanking the convex face (Figure I.9). Each repeat corresponds to \(\beta\)-\(\alpha\) structural segments with a \(\beta\)-strand and an \(\alpha\)-helix parallel to each other. Repeats align adjacent forming the curved solenoid with \(\beta\)-strands on the concave face and \(\alpha\)-helices on the convex face. The solvent-exposed \(\beta\)-sheet provides a large surface area \((2500\text{Å}^2)\) for strong ligand binding interactions (Kobe & Deisenhofer, 1995). Structures of RI and U2A’ LRR motifs bound to their ligands confirm that binding takes place at the solvent-exposed, concave face (Kobe & Kajava, 2001; Figure I.9). The structure of a cysteine-containing LRR protein, Skp2, has been determined but no examples of plant-specific structures are known (Kobe & Kajava, 2001). Despite variation in the lengths of interstrand regions between the repeat consensus sequences and loop-out regions, the structures of the LRR domains
which have been determined are very similar (Kobe & Kajava, 2001; Figure I.10). They all have a
curved, solenoid shape, with a parallel β-sheet on the concave face and α-helical or extended elements
on the convex face.

The LRR domain is a good candidate for a receptor domain (Jones & Jones, 1997). LRR motifs are
highly variable; there is variation in the length of each repeat, number of repeats and the type of repeat.
LRR domains are also the most plastic of all the structural features in R predicted polypeptides.

1.2.9: WRKY domains

WRKY proteins belong to a transcription factor superfamily that, although they are structurally
divergent, contain highly conserved, 60 residue, DNA-binding domains (Eulgem et al., 2000). WRKY
domains have a WRKYGQK motif at the N terminal end followed by a zinc-finger-like motif; many
WRKY proteins also have CC and LZ motifs. Zinc fingers are DNA binding motifs with variable
structures, but they always involve a β-hairpin and α-helix folded around (and coordinating) Zn\(^{2+}\) with
four, either cysteine or histidine, residues (Mackay & Crossley, 1998).

WRKY proteins specifically bind W-box, TTTGAC(C/T) DNA sequences. Mutation of the W-box
sequence abolished specific binding by WRKY1 and WRKY3 (Eulgem et al., 1999). The WRKY
domain of WRKY1 was sufficient for W-box specific binding (Eulgem et al., 1999). The WRKY
motifs of four NtWRKY proteins specifically bound the W-box, which were disrupted by substitutions
in the WRKYGQK and zinc-finger-like motifs and required Zn\(^{2+}\) (Maeo et al., 2001).

The WRKY domain in the RRS1-R predicted polypeptide (Deslandes et al., 2002) is identical to the
WRKY domain of AtWRKY52 (Eulgem et al., 2000), a group III WRKY protein. Group III WRKY
proteins have only one WRKY domain and the group III zinc-finger-like consensus motif
Cx\(_7\)Cx\(_{23}\)HxC.

WRKY proteins are involved in active defence responses (Eulgem et al., 2000). Chen & Chen (2000),
Dellagi et al. (2000), Robatzek & Somssich (2001), Rushton et al. (1996) and Yang et al. (1999) report
that accumulation of WRKY gene mRNA was induced during active defence responses in *A. thaliana*, tobacco, potato and parsley. The protein products of these WRKY genes have NLSs (Rushton et al., 1996), were localised to the nucleus (Robatzek & Somssich, 2001; Rushton et al., 1996) and specifically bound the W-box, including the W-box in the promoters of PR genes (Chen & Chen, 2000; Euglem et al., 1999; Rushton et al., 1996; Yang et al., 1999). *NPR1* mRNA abundance was increased in *A. thaliana* plants over-expressing *AtWRKY6* (Robatzek & Somssich, 2002). Hence, WRKY proteins are probably involved in regulating *NPR1* transcription by binding the W-box in the *NPR1* promoter (Yu et al., 2001b). TDBA12 (a WRKY protein) specific binding affinity for the PR-3 promoter was increased by phosphorylation (Yang et al., 1999). Therefore, WRKY proteins, once activated by phosphorylation, induce transcription of *NPR1* and PR genes by binding the W-box in their promoters.

### I.2.10: Protein kinase domains

Protein kinases (PKs) were defined by Hunter (1991) as,

> “enzymes that transfer a phosphate group from a phosphate donor onto an acceptor amino acid in a substrate protein. Generally the ~P of ATP or another nucleotide triphosphate is the donor, but individual enzymes may have other ~P donors.”

There are two broad classes of PKs, tyrosine and serine-threonine PKs, divided depending on the substrate amino acids that the PKs phosphorylate (Hanks et al., 1988; Hanks & Quinn, 1991). The 250-300 amino acid catalytic domains of all PKs are highly conserved (Hanks et al., 1988; Hanks & Quinn, 1991), forming a bilobal structure with the active site located in the cleft between the two lobes (Bossmeyer, 1995; The protein kinase resource, 1997; Figure I.11). Triphosphates are bound in the active site to the P-loop (Hanks et al., 1988), with the ~P oriented outwards and the substrate bound to the mouth of the cleft (Bossmeyer, 1995; Goldsmith & Cobb, 1994). Phosphorylation at the lip of the cleft can activate many PKs (Goldsmith & Cobb, 1994).

### I.2.11: Glycosylation

N-glycosylation is the post-translational addition of glucose to asparagine (N) at sites with the consensus sequence Nx(S/T) in secreted proteins, where x is unlikely to be proline or glutamate
N-glycosylation sites were predicted in Cf-2 (31), Cf-5 (25), Cf-9 (22), Cre3 (6), N (8), Pi-ta (4), R1 (9) RPM1 (3), RPS2 (6) and Xa-1 (22) predicted polypeptides. Cf-9 was proven to be heavily glycosylated when treatment with glucosidase caused a 55kDa change in protein mass (Piedras et al., 2000). N-glycosylation of the rat luteinizing hormone receptor is essential for ligand binding (Zhang et al., 1995). Thus, N-glycosylation of R proteins may effect the binding of Avr protein elicitors.

I.2.12: Myristoylation

Myristoylation is the co-translational, covalent attachment of myristic acid by an amide bond to the N terminus of a protein (Boutin, 1997; McIlhinney, 1990; Towler et al., 1988). Myristoylation occurs at the invariant glycine residue at position two in the protein (Boutin, 1997; Resh, 1999; Towler et al., 1988) and according to Resh (1999) has the consensus sequence MGxxx(S/T). Proteins that are myristoylated are associated with membranes and require myristoylation for function (Boutin, 1997; McIlhinney, 1990; Resh, 1999). Myristoylation sites were predicted in I2 (2) and Pto (1) predicted polypeptides (Figure I.4a & b). Yet myristoylation of Pto is not functionally necessary for Pto-mediated resistance (Loh et al., 1998).

I.3: Resistance protein function

R proteins must perform two functions: a) specifically recognise pathogens carrying the Avr gene and b) transduce the recognition signal to initiate active defence responses (Feys & Parker, 2000; Hammond-Kosack & Jones, 1997; Jones, 1996; Jones & Jones, 1997; Martin, 1999; Pryor & Ellis, 1993). Pathogen recognition and signal transduction, although independent functions of R proteins, are part of a continuous chain of biochemical and molecular changes leading to active defence responses. The specificity of gene-for-gene interactions and the uniformity of active defence responses in host-specific resistance, implies that R proteins are located at the site of elicitor recognition and initiate signal transduction pathways and active defence responses common to all host-specific resistances.
I.3.1: Elicitor recognition

R genes encode sentinel proteins (Fluhr, 2001). In order to resist pathogen attack plants must have a surveillance system to recognise when the plant is under attack. This surveillance system must be constitutive and able to initiate defence responses in all plant cells. Cf-9 (Jones et al., 1994), HRT (Cooley et al., 2000), Hsp pro-1 (Cai et al., 1997), Mi (Milligan et al., 1998), Pi-ta (Bryan et al., 2000), Pto (Martín et al., 1993), Rp1 (Collins et al., 1999), RPP5 (Parker et al., 1997), RPW8 (Xiao et al., 2001) and Sw-5 (Brommonschenkel et al., 2000) mRNAs were detected, albeit in low abundance, in uninfected plant cells. Also, single-cell expression assays have demonstrated that Mla1 (Zhou et al., 2001) and Mla6 (Halterman et al., 2001) were functional in individual cells. Therefore R proteins are constitutively present (probably at low concentrations) and cell-autonomous, implying that R proteins recognise Avr protein elicitors and initiate active defence responses in all plant cells.

I.3.1.1: The site of elicitor recognition

The site of elicitor recognition depends on the type of pathogen. Viruses replicate in the host cytoplasm (Agrios, 1988) and hence the site of elicitor recognition is expected to be within the host plant cell. Bacteria reside in the intercellular spaces of host plants, against the plant cell walls (Alfano & Collmer, 1996) and hence the site of elicitor recognition was expected to be outside the host cell. However, bacterial Avr proteins depend on the Hrp (Hypersensitive response and pathogenicity) type III protein secretion pathway, which is necessary to elicit the HR and thought to be responsible for the transfer of Avr proteins into the host plant cytoplasm (Alfano & Collmer, 1996; Bonas & Van den Ackerveken, 1997). Hence, the site of elicitor recognition for bacterial pathogens is expected to be within the host plant cell. Nematodes pierce and rupture host cell walls with their stylets, make a small hole in the plasma membrane and, through the stylet orifice, insert feeding tubes into the host’s cytoplasm, where oesophageal gland secretions are injected (Williamson & Hussey, 1996). Aphids insert their stylets into host plant cells (Tjallingii et al., 1993). Hence, the site of elicitor recognition for nematodes and aphids is expected to be the host plant cytoplasm. Biotrophic fungi grow in the intercellular spaces of leaves and produce haustoria within the plant cell wall, but they do not rupture the plasma membrane (Hudson, 1992); hence the site of elicitor recognition is expected to be outside the host cell.
Bacterial, viral and hemibiotrophic fungal elicitors are recognised within the plant cell. Transient expression of *AvrBs3* in pepper leaves elicits a *Bs3*-mediated HR, but only when the functional nuclear localisation signals (NLSs) are intact (Van den Ackerveken *et al*., 1996). Casper-Lindley *et al.* (2002) demonstrated that the targeting peptides of AvrBs2 fused to a reporter protein were translocated to the host plant cytoplasm by *Xanthomonas campestris* pv. *vesicatora*. Pepper plants containing *Bs2* were inoculated with *X. campestris* pv. *vesicatora* carrying *avrBs2* fused to the reporter protein sequence generated an *avrBs2*-dependent, *Bs2*-mediated HR (Casper-Lindley *et al*., 2002). PVX carrying *avrPto* generated *Pto*-mediated HRs in tomato (Tobias *et al*., 1999). The TMV helicase elicits an *N*-mediated HR when expressed in *N. tabacum* (Abbink *et al*., 1998) even with ATPase activity abolished (Erickson *et al*., 1999). Expression of *AvrB* or *avrRpt2* in *A. thaliana* leaf cells elicits an *RPM1*-mediated or *RPS2*-mediated HR, respectively (Gopalan *et al*., 1996; Leister *et al*., 1996). *AVR-Pita*176 – mature protein with N terminal SP and pro-protein sequences removed – transiently expressed in rice leaves generated a *Pi-ta*-mediated HR; whereas vacuum infiltration of *AVR-Pita*176 into rice leaves did not generate a *Pi-ta*-mediated HR (Jia *et al*., 2000). Therefore Avr proteins from bacterial, viral and hemibiotrophic fungal pathogens elicit active defence responses (and thus are recognised) within the plant cell.

Bacterial elicitors were myristoylated and palmitoylated. Induced expression of *avrPto*, mutated in its myristoylation site, failed to elicit an *avrPto*-dependent, *Pto*-mediated HR in tomato (Shan *et al*., 2000). Induced expression of either *avrRpm1* or *avrB*, mutated in the myristoylation and palmitoylation sites did not elicit an *avrRpm1* or *avrB*-dependent, respectively, *RPM1*-mediated HR in *A. thaliana* (Nimchuk *et al*., 2000). Functional myristoylation and palmitoylation motifs were also required for AvrPto, AvrRpm1 and AvrB plasma membrane localisation (Nimchuk *et al*., 2000; Shan *et al*., 2000). Mutation of the myristoylation site of AvrPphB abolished membrane fraction localisation (Nimchuk *et al*., 2000) and elicitation of *RPS5*-mediated resistance (Warren *et al*., 1998). Therefore Avr proteins from *P. syringae* are myristoylated and palmitoylated for targeting to the plasma membrane where they are recognised and elicit active defence responses at the cytoplasmic face of the plasma membrane.

The elicitors of biotrophic fungi are recognised in the apoplast. Transient expression of *Avr9* in *N.*
*N. tabacum* elicited a *Cf*-9-mediated HR (Kamoun *et al*., 1999), implying CF-9 recognises AVR9 in the cytoplasm. However, infiltration of the AVR9 peptide into leaves of tomato transformed with *Cf*-9 elicited a *Cf*-9-mediated HR (Hammond-Kosack *et al*., 1998; Hennin *et al*., 2001), indicating that CF-9 recognises AVR9 in the apoplast. These results appear contradictory but the predicted polypeptide of *Avr9* was 63 amino acids long with a predicted 23 amino acid SP at the N terminus (Van den Ackervecken *et al*., 1992). Van den Ackervecken *et al*., (1993) demonstrated that a 34 amino acid AVR9 was secreted by *C. fulvum*, which was subsequently cleaved to the 28 amino acid, AVR9 peptide elicitor by plant factors. Hence, expression of *Avr9 in planta* produced the 63 amino acid AVR9 peptide, which was secreted to the apoplast where the 40 amino acid AVR9 peptide was cleaved by plant factors to produce the AVR9 peptide elicitor. The predicted polypeptide of *Avr4* also has a predicted SP at its N terminus (Joosten *et al*., 1994). Therefore the AVR9 peptide and probably the AVR4 peptide elicits active defence responses in the apoplast. Transient expression of CF-9 minus the E,F and G domains in *N. tabacum*, transformed with *Avr9*, did not generate an *Avr9*-dependent, *Cf*-9 mediated HR (Thomas *et al*., 2000), demonstrating that the E,F and G domains of CF-9 are required and indicating that the site of elicitor recognition for the AVR9 peptide is the apoplastic face of the plasma membrane.

I.3.1.2: Resistance protein subcellular localisation

The subcellular localisation of R proteins can be predicted from the R gene predicted polypeptide structures. All R gene predicted polypeptides, except those with SPs and TM helices, are expected to be cytoplasmic (Figure I.4a & b). The R gene predicted polypeptides with SPs and TM helices (CF-2, CF-4, CF-5, CF-9 and Xa21) are expected to span the plasma membrane (I.4a).

The expected subcellular localisations for CF-9 and RPM1 predicted polypeptides is consistent with the known subcellular localisations of CF-9 and RPM1. C terminally, c-myc tagged RPM1 was functional and enriched in plasma membrane fractions by associating peripherally with the plasma membrane in *A. thaliana* (Boyes *et al*., 1998). Functional, N and C terminally, c-myc tagged CF-9 were found in plasma membrane fractions of tobacco plants (Piedras *et al*., 2000). Immunogold labelling demonstrated that the N terminally, c-myc tagged CF-9 was present on the surface of tobacco protoplasts (Piedras *et al*., 2000). Benghezal *et al*., (2000) found that GFP fused to the region of CF-9
containing E, F and G (TM and cytosolic) domains and N terminally, triple hemagglutinin tagged Cf-9 was localised to the ER in tobacco cells and retained in the ER by the putative ER retrieval dilysine motif, KKRY at the Cf-9 C terminus (Jones et al., 1994). However, Benghezal et al. (2000) did not test if these Cf-9 fusion proteins were functional. Van der Hoorn et al. (2001) mutated the KKRY motif in Cf-9 to AARY and found that it did not abolish the Cf-9-mediated HR after infiltration with the AVR9 peptide. A SP was predicted in the Cf-9 predicted polypeptide (Jones et al., 1994) so Cf-9 would be expected to be targeted to the ER (I.2.1.1); c-myc tagged RPM1 was detected in the ER (Boyes et al., 1998). Also, over-expression of the Cf-9 fusion proteins by Benghezal et al. (2000), using the CaMV 35S promoter, could have resulted in fusion protein aggregation in the ER. Therefore the plasma membrane is the subcellular localisation of functional Cf-9 and RPM1; RPM1 is a peripheral plasma membrane protein and Cf-9 is an integral plasma membrane protein, protruding from the apoplastic face.

The sites of elicitor recognition coincide with R protein or R gene predicted polypeptide subcellular locations. Bacteria and viruses are recognised within the plant cell, i.e. the cytoplasm, which is where their corresponding R gene predicted polypeptides are expected to be located and where RPM1 is located. Biotrophic fungi are recognised in the apoplast, which is where their corresponding R gene predicted polypeptides are expected to be located and where Cf-9 is located. Aphids and nematodes pierce the plasma membrane and come into contact with the cytoplasm, which is where their corresponding R gene predicted polypeptides are expected to be located. Hemibiotrophic fungi are extracellular and although they are not thought to breach the plasma membrane they do release toxins and enzymes (Dickinson & Lucas, 1982). These toxins and enzymes could come into contact with the cytoplasm, the expected location of their corresponding R gene predicted polypeptides. Hence, R protein subcellular localisation is dependent on the modus operandi of the pathogen.

I.3.1.3: Direct and indirect elicitor recognition

The elicitor-receptor model (Figure I.1) depicts R proteins acting as receptors for their corresponding Avr protein elicitors at the site of elicitor recognition. Hence, R proteins are expected to interact with their corresponding Avr proteins and elicitor recognition is direct, as proposed by Parker & Coleman (1997). If R and Avr proteins do not interact then elicitor recognition must be indirect.
Direct elicitor recognition has only been demonstrated for Pto and Pi-ta. Pto mutants that were unable to interact with AvrPto in yeast were also unable to confer avrPto-dependent, Pto-mediated resistance in tomato (Scofield et al., 1996; Tang et al., 1996). Substitution T38A in Pto abolished the interaction between Pto and AvrPto in yeast and the avrPto-dependent, Pto-mediated HR in N. tabacum (Sessa et al., 2000). The LRR domain of Pi-ta interacts with AVR-Pita$_{176}$ in vitro and in yeast (Jia et al., 2000). Substitution of A918S in the Pi-ta LRR domain disrupted the interaction with AVR-Pita$_{176}$ in yeast and also abolished avrPita-dependent, Pi-ta-mediated resistance in rice (Jia et al., 2000). Therefore interactions between AVR-Pita and Pi-ta (through the LRR domain) and AvrPto and Pto were necessary for Pi-ta and Pto, respectively, to initiate active defence responses.

Several R proteins did not interact with their corresponding Avr proteins. HRT did not interact with the TCV capsid protein in yeast (Ren et al., 2000). RPM1 did not interact with either AvrRpm1 or AvrB in yeast (Boyes et al., 1998). RPS5 did not interact with AvrPphB in yeast (Swiderski & Innes, 2001). Cf-9 and AVR9 did not interact in yeast, nor was Cf-9 specific binding of the AVR9 peptide observed in tobacco cell membrane preparations (Luderer et al., 2001). Therefore elicitor recognition of the TCV capsid protein by HRT, AVR9 by Cf-9, AvrRpm1 or AvrB by RPM1 and AvrPphB by RPS5 is indirect.

Indirect elicitor recognition implies that other proteins, besides R and Avr proteins, are involved. AvrRpt2 and a 75 kDa protein from A. thaliana cell extracts co-immunoprecipitated with RPS2 (Leister & Katagiri, 2000); the 75 kDa protein was also co-immunoprecipitated by the LZ and NBS region of RPS2, prevented by deletions of the LZ region (Tao et al., 2000). This implies that the 75 kDa protein is the AvrRpt2 elicitor receptor component of the RPS2 recognition complex. RIN4 (RPM1-interacting protein 4) interacted independently with RPM1 (though the CC motif) and AvrB, but not RPS2 or RPP5, in yeast (Mackey et al., 2001). RPM1, AvrRpm1 or AvrB were co-immunoprecipitated independently from A. thaliana cell extracts by RIN4, which was membrane localised (Mackey et al., 2001). Hence, RIN4 interacted with AvrRpm1, AvrB and specifically with RPM1, indicating RIN4 is the RPM1 specific receptor for the AvrRpm1 and AvrB elicitors. The TCV capsid protein elicits HRT-mediated resistance to TCV and interacts with TIP (TCV-interacting protein), a NAC protein and
probably a transcription factor, in yeast (Ren et al., 2000). Mutations in the TCV capsid protein that disrupt the interaction with TIP in yeast also permitted TCV to evade inducing HRT-mediated HR and resistance (Ren et al., 2000). Therefore, TIP is the TCV capsid protein elicitor receptor component of the HRT indirect recognition complex, although TIP’s connection to HRT is unknown. Hence, RIN4, TIP and the 75 kDa protein are probably specific Avr protein elicitor receptor components of RPM1, HRT and RPS2 indirect recognition complexes, respectively, which interact with the R protein LZs.

Cf-4 (Rivas et al., 2002b) and Cf-9 (Rivas et al., 2002) constitutively form protein complexes (without disulphide bridges or dimerisation) of approximately 400 and 420 kDa, respectively, in tobacco microsomal membranes with unknown glycoproteins. The mass of these complexes extracted from N. benthamiana leaves did not change detectably after infiltration of the AVR4 peptide and AVR9 peptide elicitors, implying that the AVR4 and AVR9 peptide elicitors do not bind Cf-4 and Cf-9, membrane localised complexes, respectively. However, the AVR4 and AVR9 peptides may only be bound transiently or weakly to Cf-4 and Cf-9, respectively.

I.3.1.4: Elicitor recognition specificity

The recognition of Avr proteins by their corresponding R proteins is specific for the R and Avr proteins involved in a gene-for-gene manner. Mutation (of R genes) or domain swapping – exchanges of sequence between related, i.e. paralogous, R genes to form fused R genes – techniques have been used to find the Avr protein elicitor recognition specificity determinants within the R proteins.

Fen is an autophosphorylatable serine-threonine PK (Loh & Martin, 1995), with 80% identity and 87% similarity to Pto, which was responsible for fenthion sensitivity, but was not able to confer resistance to P. syringae pv. tomato carrying avrPto (Martin et al., 1994). Domain swaps between Pto and Fen by Scofield et al. (1996) and Tang et al. (1996) identified residues 188-217 and 202-209, respectively, in Pto as necessary for a) interaction between Pto and AvrPto in yeast and b) an avrPto-dependent, Pto-mediated HR in N. tabacum. There are only four differences between residues 202-209 of Pto and Fen (Martin et al., 1994). Frederick et al. (1998) substituted Pto residues for Fen residues for each of the four different residues and discovered that two substitutions, T204N and L205I, effected the interaction between Pto and AvrPto in yeast. L205I attenuated the interaction of Pto and AvrPto in yeast, but
T204N completely abolished interaction in yeast and an *avrPto*-dependent, *Pto*-mediated HR in *N. tabacum* (Frederick *et al*., 1998). Substitution of N204T, but not I205L, in Fen permitted Fen to generate an *avrPto*-dependent HR in *N. tabacum* (Frederick *et al*., 1998). Therefore Thr204 is essential for *Pto* recognition specificity.

The predicted polypeptide of Cf-4 differs with respect to Cf-9 by 78 amino acid polymorphisms (AAPs), a 46 residue deletion of two complete LRRs (11th and 12th) and a 10 residue deletion at the N terminus (Thomas *et al*., 1997). Domain swapping and amino acid substitution between Cf-4 and Cf-9 abolished their ability to initiate *Avr4* and *Avr9*-dependent active defence responses, *e.g.* insertion in Cf-4 or deletion in Cf-9 of the 11th and 12th LRRs produces non-functional R proteins (Van der Hoorn *et al*., 2001b; Wulf *et al*., 2001). Van der Hoorn *et al*., (2001b) discovered that by deletion of the 11th and 12th LRRs of Cf-9 and substituting Y389W, E411G and L457F amino acids in Cf-9 that they produced an R protein able to generate an *Avr4*-dependent HR in *N. benthamiana*. Therefore, the absence of the 11th and 12th LRRs and presence of amino acids W389, G411 and F457 in the LRR domain were essential for Cf-4 specificity and function. Tomato plants transformed with the 34 N terminal LRR encoding region of Cf-2 fused to the C terminal encoding region of Cf-9 (including the three Cf-9 C terminal LRRs) conferred Cf-2 specific resistance (Krüger *et al*., 2002). Hence, the LRR domains of Cf-2, Cf-4 and Cf-9 were sufficient to determine specificity.

Ellis *et al*., (1999) and Luck *et al*., (2000) investigated the specificity of *L*-mediated resistance with domain swaps between paralogues at the *L* locus. Flax plants transformed with fusions of the TIR and NBS encoding region from either *L6* or *L10* to the LRR encoding region of *L2* conferred *L2* specific resistance (Ellis *et al*., 1999). This indicated that the LRR domain of L2 was sufficient to determine L2 specificity. However, *L6* and *L7* confer different resistance specificities but their predicted polypeptides only differ in the TIR region (Luck *et al*., 2000). Indeed, flax plants transformed with the TIR encoding region of either *L2* or *LH* fused to the NBS and LRR encoding region of *L6* conferred *L7* specific resistance, as did plants transformed with the TIR and NBS encoding region of *L2* fused to the LRR encoding region of *L6* (Luck *et al*., 2000). This indicated that S47P, S83F and L99R amino acids in the TIR domain of L7 were determinants of L7 specific resistance. Novel resistance specificities
were conferred by the TIR and NBS encoding regions of either L2 or L9 fused to the LRR encoding region of L10 (Ellis et al., 1999; Luck et al., 2000), rather than L10 specific resistance. The TIR encoding region of L9 fused to the NBS and LRR encoding region of L10 also conferred the novel resistance specificity, indicating that two amino acids, E209K and I214A, in the TIR domain of L10 were determinants of L10 specific resistance (Luck et al., 2000). Hence, the TIR and LRR domains of L contain determinants of L specific resistance.

Elicitor recognition can be determined by TIR and LRR domains and, for direct recognition of the Avr protein then requires interaction with the Avr protein. Also, the specificity of elicitor recognition can be determined by a single amino acid. Importantly, R proteins determine the specificity of elicitor recognition, even if it is indirect.

I.3.2: Plant disease resistance protein signal transduction

I.3.2.1: Restricted taxonomic functionality

R genes are functional in heterologous species. Infection with P. syringae pv. tabaci carrying avrPto induces Pto-mediated HR and resistance in N. benthamiana (Rommens et al., 1995) and N. tabacum (Thilmony et al., 1995). TMV or ToMV infection of tomato, transformed with N, induces N-mediated HR and resistance (Whitham et al., 1996). Transient expression of Cf-2 generated an Avr2-dependent, Cf-2-mediated HR in tobacco (Dixon et al., 2000). Co-expression of Cf-4 and Avr4 generated an Avr4-dependent, Cf-4-mediated HR in N. benthamiana, N. glutinosa, N. rustica, N. tabacum, Petunia hybridia and lettuce (Thomas et al., 2000; Van der Hoorn et al., 2000). Co-expression of Cf-9 and AVR9 generated an Avr9-dependent, Cf-9-mediated HR in N. benthamiana, N. glutinosa, N. rustica, N. tabacum, Petunia hybridia, potato and Brassica napus (Hammond-Kosack et al., 1998; Hennin et al., 2001; Kamoun et al., 1999; Thomas et al., 2000; Van der Hoorn et al., 2000). Infiltration of the AVR9 peptide into leaves of N. tabacum, potato, tomato, oilseed rape or Brassica napus transformed with Cf-9 elicited a Cf-9-mediated HR (Hammond-Kosack et al., 1998; Hennin et al., 2001; Hennin et al., 2001b). Mla6 conferred resistance to Blumeria graminis f. sp. hordei heterologously in wheat – a member of the Triticeae (Halterman et al., 2001). Co-expression of Bs2 and avrBs2 in pepper, tomato and N. benthamiana – members of the Solanaceae – generated an avrBs2-dependent, Bs2-mediated HR
Transient expression of Rx generated a PVX coat protein-dependent, Rx-mediated HR in *Nicotiana* sp. (Bendahmane et al., 1999). *N. tabacum* transformed with Rx2 is resistant to PVX (Bendahmane et al., 2000). Hence, R genes are functional in heterologous species, which implies that the signal transduction pathways are conserved between the native and heterologous species.

Heterologous R gene function is taxonomically restricted. Bs2 was not able to generate an *avrBs2*-dependent HR in *A. thaliana*, turnip, cucumber or broccoli – non-solanaceous species (Tai et al., 1999). *Cf-4* and *Cf-9* were not able to generate an *Avr4*-dependent or an *Avr9*-dependent HR, respectively, in *N. clevelandii*, *A. thaliana*, radish, lupin, pea or flax (Van der Hoorn et al., 2000). Thus, solanaceous R genes *Bs2*, *Cf-4* and *Cf-9* are functionally restricted to the *Solanaceae* and conversely *RPS2*, from *A. thaliana*, was not functional in tomato (Tai et al., 1999). There are two exceptions to the solanaceous taxonomic restriction. Firstly, *Cf-4* generated an *Avr4*-dependent HR in lettuce – a member of the *Compositae* (Van der Hoorn et al., 2000). Secondly, *Cf-4* and *Cf-9* were not able to generate an *Avr4*-dependent or *Avr9*-dependent HR, respectively, in *N. clevelandii*, a member of the *Solanaceae* (Van der Hoorn et al., 2000). *N. clevelandii*, unlike *N. benthamiana* or *N. tabacum*, was resistant to *P. syringae pv. tabaci* carrying *avrPto* even though *N. clevelandii* did not have a *Pto* homologue (Rommens et al., 1995). *N. clevelandii*, despite being a member of the *Solanaceae* does not seem to possess the same R gene and signal transduction pathway components as other solanaceous species, which may explain why *Cf-4* and *Cf-9* were not functional in *N. clevelandii*. Surprisingly, *Cf-9* was not able to generate an *Avr9*-dependent HR in lettuce (Van der Hoorn et al., 2000) despite the similarities between *Cf-4* and *Cf-9* heterologous function. Perhaps the homology between *Cf-4* signal transduction pathway components in tomato and lettuce is greater than *Cf-9*.

Heterologous R gene function is correlated with the presence of homologous sequences. *Bs2* was functional in heterologous species that had *Bs2* hybridising sequence in the genomic DNA (Tai et al., 1999). *Pto* has hybridising sequence in *N. benthamiana*, *N. tabacum*, potato, *A. thaliana*, bean, soyabean, pea, rice, maize, barley, wheat and sugarcane (Martin et al., 1993; Rommens et al., 1995). *PtI1* hybridising sequences were also found in tobacco, potato, *A. thaliana*, bean, *Petunia*, pepper and sugarcane (Zhou et al., 1995). *MlaI* – which is a homologue of *Mla6* – has a wheat orthologue, *TaMla* (Zhou et al., 2001). Hence, heterologous R gene function is restricted taxonomically to species with hybridising sequences of R genes and their signal transduction pathway components.
I.3.2.2: Signal transduction pathway components

R genes initiate defence responses, including the HR, via signal transduction pathways. Components of R gene signal transduction pathways are described here to illustrate their connection to the biochemical changes associated with the HR and resistance (I.1.1).

I.3.2.2.1: Signal transduction pathways initiated by resistance protein kinases

He et al. (2000) demonstrated that the PK domain of Xa21 could initiate the HR. Rice cells were transformed with the region of Xa21 encoding the PK domain fused to the region of BRI1 (a brassinosteroid receptor) encoding the extracellular LRR domain and TM helix. ROS accumulation, defence gene induction and the HR specifically occurred in rice cells transformed with the BRI1-Xa21 fusion protein and treated with brassinosteroid. Therefore the PK domain of Xa21 is sufficient to initiate elicitor-dependent active defence responses; the result of this experiment also implies that the LRR domain of Xa21 is the elicitor receptor domain.

Pto is a serine-threonine PK that autophosphorylated \textit{in vitro} (Loh & Martin, 1995) at eight serine and threonine residues (Sessa et al., 2000) – Ser\textsuperscript{17}, Thr\textsuperscript{38}, Thr\textsuperscript{133}, Thr\textsuperscript{190}, Thr\textsuperscript{195}, Ser\textsuperscript{198}, Thr\textsuperscript{199} and Thr\textsuperscript{288}. Several signal pathway components interacted with Pto in yeast, including Pti1 (Pto-interacting 1), Pti4, Pti5 and Pti6 (Zhou et al., 1995; Zhou et al., 1997). Sessa et al. (2000) discovered that two amino acid substitutions, T38A and S198A, abolished the \textit{avrPto}-dependent, Pto-mediated HR in \textit{N. tabacum}. Thr\textsuperscript{38}, but not Ser\textsuperscript{198}, was necessary Pto PK activity \textit{in vitro} and for the interaction of Pto with Pti1 and Pti4 in yeast (Sessa et al., 2000). Ser\textsuperscript{198} was necessary for the interaction between Pto and Pti10 in yeast and the S198A substitution impaired the interaction between Pto and Pti3 in yeast. Therefore, Pto phosphorylation is necessary for its PK activity, interaction with Pti1, Pti3, Pti4 and Pti10 and the generation of an \textit{avrPto}-dependent, Pto-mediated HR in \textit{N. tabacum}.

Pti1 is a serine-threonine PK, specifically phosphorylated at Thr\textsuperscript{233} (Sessa et al., 2000b) by either Pto
(Zhou et al., 1995) or the product of the susceptible allele \textit{pto} (Sessa et al., 1998). Pti1 interaction with and phosphorylation by Pto was disrupted by the T233A substitution in Pti1 (Sessa et al., 2000b), although the interaction of PK inactive Pti1 with Pto was not altered (Zhou et al., 1995). Over-expression of \textit{Pti1} in \textit{N. tabacum} accelerated and enhanced the \textit{Pto}-mediated, \textit{avrPto}-dependent HR (Zhou et al., 1995).

\textit{Pti4, Pti5} and \textit{Pti6} encode proteins with ethylene-responsive element binding protein (EREBP) transcription factor characteristics and Pto interacted with EREBP-2 in yeast (Zhou et al., 1997). Pti4 and Pti5 specifically bind GCC-boxes (PR boxes) \textit{in vitro} (Gu et al., 2000) and GCC-box promoters in tobacco (Wu et al., 2000). Transcript of GCC-box genes accumulated during \textit{Pto}-mediated resistance to \textit{P. syringae pv. tomato} carrying \textit{avrPto} (Jia & Martin, 1999). Over-expression of \textit{Pto} or \textit{Pti4} in tomato leaves leads to increased GCC-box, PR gene transcript (Gu et al., 2000; Tang et al., 1999). In \textit{A. thaliana} over-expressed \textit{Pti4} induced at least 28, mostly PR protein encoding (18 had GCC-boxes) genes and caused an ethylene-like phenotype (Wu et al., 2000). Constitutive expression of \textit{Pti5} in tomato leaves did not lead to increased accumulation of PR gene mRNA (Gu et al., 2000).

\textit{Pti4} function is distinct to \textit{Pti5} and \textit{Pti6}. Pto specifically phosphorylates \textit{Pti4} at four threonine residues \textit{in vitro}, but does not phosphorylate \textit{Pti5} (Gu et al., 2000). Wounding, ethylene, SA and \textit{P. syringae pv. tomato} infection induce \textit{Pti4} transcription, whereas \textit{Pti5} and \textit{Pti6} transcription is not induced by wounding, ethylene or SA (Gu et al., 2000; Thara et al., 1999).

Fenthion sensitivity (conferred by the \textit{Pto} paralogue, \textit{Fen}) and \textit{Pto}-mediated resistance both require \textit{Prf} (\textit{Pseudomonas} resistance and fenthion sensitivity), implying that \textit{Prf} is a signal transduction pathway common to \textit{Fen} and \textit{Pto} (Salmeron et al., 1994). \textit{Prf} encodes a protein with an 18 repeat LRR domain, LZ and NBS motifs (Salmeron et al., 1996) that did not interact with Pto in yeast (Bogdanove & Martin, 2000). Over-expression of \textit{Prf} in tomato plants resulted in increased fenthion sensitivity, SA accumulation, constitutive SAR and increased resistance to a broad spectrum of normally virulent pathogens, but \textit{Pto}-mediated resistance was not affected (Oldroyed & Staskawicz, 1998). Hence, \textit{Prf} is not a component of the \textit{Pto} recognition complex and thus must be a component of a signal transduction pathway specific to \textit{Pto} and \textit{Fen}, which initiates SAR, but not \textit{Pto} specific responses.
Five Adi (AvrPto-dependent, Pto-interactor) proteins interacted with Pto and AvrPto in yeast: Adi1 was identified as tomato catalase 1, Adi2 is a homologue of Pti1 that significantly increased binding in the presence of AvrPto, Adi3 is a putative serine-threonine PK, Adi4 was not identified, but had the strongest interaction and Adi5 was identified as Pti2, a proteosome α-subunit (Bogdanove & Martin, 2000). These Adi proteins are probably components of unidentified signal transduction pathways involved in Pto initiated active defence responses.

I.3.2.2.2: Mitogen-activated protein kinase and calcium-dependent protein kinase signal transduction pathway components

Mitogen-activated protein kinases (MAPKs) are components of three tier signal transduction cascades intermediate between receptors and responses that can amplify signals (Davis, 1993; Hirt, 1997; Meskiene & Hirt, 2000) and filter out background noise, but switch within a narrow input range to induce responses (Huang & Ferrell, 1996). Mitogen-activated protein kinase kinase kinases (MAPKKKs) activate or deactivate Mitogen-activated protein kinase kinases (MAPKKs) by phosphorylation at serine and threonine residues in the S/TxxxS/T motif (Meskiene & Hirt, 2000). MAPKKs from one or more alternative signal transduction cascades can be activated or deactivated by MAPKKKs, suggesting MAPKKs are points of convergence between signal transduction cascades (Meskiene & Hirt, 2000). MAPKs are activated by MAPKK by phosphorylation at threonine and tyrosine residues (Stone & Walker, 1995) in the T-Glu-Tyr motif (Hirt, 1997; Meskiene & Hirt, 2000). MAPKs are serine-threonine protein kinases which phosphorylate protein kinases, phosphatases, enzymes and transcription factors in the cytoplasm and the nucleus at Pro-x-S/T-Pro motifs (Davis, 1993; Hirt, 1997; Meskiene & Hirt, 2000).

Romeis et al. (1999) detected Avr9-dependent, Cf-9-mediated activation of two (46 kDa and 48 kDa) protein kinases in tomato and tobacco, which involved Ca$^{2+}$ influx and tyrosine phosphorylation, but not ROS accumulation. Romeis et al. (1999) identified the 46 kDa and 48 kDa protein kinases as wound-induced protein kinase (WIPK) and SA-induced protein kinase (SIPK), respectively. In tobacco SIPK was rapidly and transiently activated by phosphorylation at threonine and tyrosine residues in response to exogenously applied SA (Zhang & Klessig, 1997). In tobacco WIPK was transiently
activated by phosphorylation at serine, threonine and tyrosine residues in response to wounding (Usami et al., 1995). SIPK and WIPK were activated during N-mediated resistance to TMV (Zhang & Klessig, 1998b). However, Zhang & Klessig (1998) found that SIPK, not WIPK, was activated by phosphorylation at tyrosine residues in response to wounding in tobacco leaves; WIPK activity was undetectable in wounded tobacco leaves. Also, WIPK was activated by other signals (Usami et al., 1995). The activation of WIPK was SA-independent (Zhang & Klessig, 1998b). Transcription of the WRK (wound-induced receptor-like protein kinase) gene, which encodes a serine-threonine protein kinase, was SA-independently induced during N-mediated HR in Nicotiana sp. (Ito et al., 2002). Therefore SIPK and WIPK (WRK) activity was induced during Cf-9 and N-mediated defence responses.

Ca2+ influx is associated with the HR (I.1.1). Calcium-dependent protein kinases (CDPKs) are activated by binding Ca^{2+} and serve to transduce signals involving temporal and spatial changes in Ca^{2+} concentration (Harmon et al., 2000). A membrane-bound CDPK (CDPK2) was phosphorylated in Cf-9 transformed tobacco cells treated with the AVR9 peptide (Romeis et al., 2000). Virus-induced gene silencing (VIGS) of NtCDPK2 in N. benthamiana significantly reduced and delayed Avr9-dependent, Cf-9-mediated and Avr4-dependent, Cf-4-mediated HRs (Romeis et al., 2001). CDPK2 phosphorylation was not affected by MAPK or CDPK inhibitors, but CDPK inhibition did prevent ROS accumulation (Romeis et al., 2000). Therefore Cf-4 and Cf-9-mediated phosphorylation of the membrane-bound CDPK2 does not involve another CDPK or MAPKs but was necessary for ROS accumulation and the HR.

Cf-2-mediated resistance requires Rcr3 (required for Cladosporium resistance 3), which is not required for Cf-5 and Cf-9-mediated resistances (Dixon et al., 2000). RCR3 was localised to the apoplast and the mature RCR3 protein (with the SP and pro-domain removed), but not the pre-protein, was shown to be a cysteine protease (Krüger et al., 2002). Therefore RCR3 becomes active once it reaches the apoplast where it probably either cleaves AVR2 to produce the AVR2 peptide elicitor or is involved in Cf-2 specific signal transduction.

Cf gene mediated resistances have common signal transduction components. Cf-2 and Cf-9-mediated
resistance to *C. fulvum* carrying *Avr2* or *Avr9*, respectively, in tomato plants was independent of SA accumulation (Brading *et al*., 2001). ROS accumulation and MAPK activity was detected during an AVR4 peptide elicited HR in Cf-4 transformed *N. benthamiana* (Rivas *et al*., 2002b), as is the case for Cf-9. Hence, Cf-2, Cf-4, Cf-9 and probably Cf-5-mediated resistances, use the same signal transduction pathway components and probably use the same signal transduction pathway.

### 1.3.2.2.3: EDS1 and NDR1 signal transduction pathway components

Two signal transduction pathways were clearly defined. *RPP1*, *RPP2*, *RPP4*, *RPP5*, *RPP10*, *RPP12*, *RPP14*, *RPP21*, *RPS4* and *RPW8* require *EDS1* (enhanced disease susceptibility 1) for resistance (Aarts *et al*., 1998; Calis, 2001; Parker *et al*., 1996; Xiao *et al*., 2001). *NDR1* (non-race-specific disease resistance 1), but not *EDS1*, is required for *RPM1*, *RPS2* and *RPS5*-mediated resistance (Aarts *et al*., 1998; Century *et al*., 1995). VIGS of *NbEds1* in *N. benthamiana*, by Peart *et al*. (2002), demonstrated that EDS1 was required for *N*-mediated, but not *Pto* or *Rx*-mediated resistance. TIR domain encoding R genes are all *EDS1*-dependent (Aarts *et al*., 1998). Hence, *EDS1* or *NDR1* dependence is specified by R protein structure.

*RPP1*, *RPP5*, *RPP21*, and *RPS4*-mediated (but not *RPM1* or *RPS2*-mediated) resistances were completely compromised by *eds1* and partially by *pad4* (phytalexin-deficient 4) or SA deficiency (Feys *et al*., 2001). Feys *et al*. (2001) found that *Eds1* was essential for the HR, whereas *PAD4* or SA deficiency was not. *RPP4* (van der Biezen *et al*., 2002) and *RPW8*-mediated (Calis, 2001; Xiao *et al*., 2001) resistances were also compromised by *pad4* and SA deficiency. Mutations in *PAD4* caused enhanced disease susceptibility to *P. syringae* pv. *maculicola* in *A. thaliana* and a failure to accumulate the phytalexin, camalexin, or induce PR-1 gene transcription (Glazebrook *et al*., 1997; Zhou *et al*., 1998). SA synthesis was reduced and delayed in *pad4* plants (Zhou *et al*., 1998). Exogenously applied SA partially rescued camalexin and PR-1 mRNA deficiency in *pad4* plants (Zhou *et al*., 1998) and PR-I mRNA accumulation in *eds1* plants (Falk *et al*., 1999). The accumulation of PR-1 mRNA and camalexin during resistance to either *Cochliobolus carbonum* or *P. syringae* pv. *maculicola* carrying *avrRpt2* was unaffected by *pad4* (Glazebrook *et al*., 1997; Zhou *et al*., 1998). Therefore EDS1 is
pleiotropic to PAD4, which regulates camalexin synthesis and PR-1 gene expression via SA.

EDS1 and PAD4 are probably lipases that form part of a complex in planta. The EDS1 N terminus is similar to the catalytic site of euakaryotic lipases (Falk et al., 2001), although site-directed-mutagenesis of the catalytic site by Parker et al. (2001) did not compromise resistance. PAD4 is similar to triacyl glycerol lipases and other sterases (Jirage et al., 1999). EDS1 dimerises and interacts with PAD4 in yeast; mutated EDS1 dimerised but PAD4 interaction was abolished (Feys et al., 2001). EDS1 and PAD4 co-immunoprecipitated in infected and uninfected A. thaliana tissue and the protein content was higher in infected tissue (Feys et al., 2001). Therefore an EDS1-PAD4 complex, which probably signals via a lipid compound, is more abundant in infected tissue and is essential for the EDS1 signal transduction pathway.

RPP4 (van der Biezen et al., 2002) and RPP5 (van der Biezen et al., 2000) interact with RSH1, a RelA/SpoT homologue from A. thaliana, through their NB-ARC domains, in yeast. RSH1 complements relA, but not spoT, mutants in E. coli, demonstrating that RSH1 synthesises guanosine tetraphosphate or guanosine pentaphosphate. RSH1 did not interact with RPM1, RPP1, rpp5, RPS4 or N in yeast. Hence, RSH1 is a specific component of the RPP4 and RPP5 recognition complexes, which synthesises the potential signalling compounds, guanosine tetr phosphate or guanosine pentaphosphate.

RPM1, RPS2 and RPS5-mediated resistances were compromised by ndr1, but these (except RPS2 A. thaliana) plants are still able to generate an HR (Century et al., 1995). RRS1-R-mediated resistance was also compromised by ndr1 (Deslandes et al., 2002). P. syringae carrying AVRpl2, but not AVRpl, infected ndr1 plants accumulated a reduced amount of PR-1 mRNA – an indicator of systemic acquired resistance (Shapiro & Zhang, 2001). SA accumulation, in response to ROS, was impaired in ndr1 plants (Shapiro & Zhang, 2001). P. syringae carrying AvrB, but not avrRpt2, infected ndr1 plants accumulated H2O2 (Shapiro & Zhang, 2001). Therefore, in response to ROS, NDR1 – predicted to have two TM helices and to be membrane associated (Century et al., 1997) – positively regulates SA accumulation and SAR; HR generation by ROS is NDR1-independent.

RPM1 and RPS2-mediated resistances compete for signal transduction components in A. thaliana (Ritter & Dangl, 1996). Independently, an avrRpm1-dependent, RPM1-mediated, strong HR develops
after 5 hours and an *avrRpt2*-dependent, *RPS2*-mediated, weak HR develops after 20 hours. In mixed inoculations the weaker and slower *avrRpt2*-dependent HR masks the stronger and faster *avrRpm1*-dependent HR. Transcription of the defence response genes AIG1 and AIG2 (*avrRpt2*-induced gene 1 and 2) or ELI3 were specifically induced by *RPS2* (Reuber & Ausubel, 1996) or *RPM1*-mediated (*Kiedrowski et al., 1992*) resistances to *P. syringae* in *A. thaliana*. Masking of the *avrRpm1*-dependent HR by the *avrRpt2*-dependent HR corresponded to *RPS2* or *RPM1* specific defence response gene transcription – AIG1 transcript accumulated and ELI3 transcript abundance was at background quantities. Increasing the ratio of *avrRpm1:* *avrRpt2* carrying *P. syringae* to 2.5 fold more *avrRpm1* than *avrRpt2* was sufficient to prevent masking of the *avrRpm1*-dependent HR and specific defence response gene transcription – ELI3, and not AIG1, transcription was induced. An *avrRpm1*-dependent HR is still masked by the *avrRpt2*-dependent HR in an *rps2* background and vice versa. Hence the masking was not due to the R proteins. The *avrRpt2*-dependent HR develops faster (at 6 hours rather than 21 hours) in the *RPM1*-null background. Therefore *avrRpt2*-dependent and *avrRpm1*-dependent HRs compete for components of the *NDRI* signal transduction pathway upstream of NDR1 and, possibly in complex with RPS2 and RPM1.

Warren et al. (1999) found that *RPS5*-mediated resistance, but not *RPM1*, *RPS2* and *RPS4*–mediated resistances, were compromised by *pbs1* (*avrPphB* susceptible 1). PBS1 is an autophosphorylatable serine-threonine PK that did not interact with either RPS5 or AvrPphB in yeast (Swiderski & Innes, 2001). Hence, PBS1 is a signal transduction pathway component that is specific for RPS5 but is not a component of the RPS5 recognition complex.

NPR1 activates PR gene transcription in response to SA. *A. thaliana npr1* (non expresser of PR genes 1) plants are susceptible, SA insensitive and unable to accumulated PR gene mRNA, like SA deficient plants (Cao et al., 1994; Delaney et al., 1995; Shah et al., 1997). SA accumulation in npr1 plants was actually greater than wild type plants (Shah et al., 1997), implying disruption of a negative feedback loop. Over-expression of *NPR1* enhanced resistance, which correlated with increased accumulation of *NPR1* mRNA, NPR1 and PR gene mRNA after infection (Cao et al., 1998; Friedrich et al., 2001). SA deficient plants over-expressing *NPR1* are susceptible (Friedrich et al., 2001), indicating *NPR1* over-expression increases the sensitivity to, but not independence from SA. Also, *AtNPR1* functions heterologously and has homologues in tomato, tobacco and rice (Dong et al., 2001). Therefore NPR1
is downstream of SA, activating PR gene transcription.

NPR1 interacts with AHBP-1b, TGA2, TGA2.1, TGA2.2, TGA3, TGA6 and TGA7 basic leucine zipper transcription factors in yeast (Després et al., 2000; Niggeweg et al., 2000; Zhang et al., 1999; Zhou et al., 2000). Interaction of NPR1 deletion constructs and transcription factors in yeast indicated that the interactions rely on the central, ankyrin repeats of NPR1 (Zhang et al., 1999; Zhou et al., 2000). AHBP-1b, TGA2 and TGA3 bind to the SA-responsive, as-1 element in the A. thaliana PR-1 promoter (Zhang et al., 1999; Zhou et al., 2000). TGA2.1 and TGA2.2 bind to the as-1 element in the tobacco PR-1 promoter (Niggeweg et al., 2000). NPR1, but not non-functional NPR1 increased TGA2 binding to the as-1, LS7 (SA-inducible and positive) and LS5 (negative) elements of the PR-1 gene promoter (Després et al., 2000). NIMIN-1, NIMIN-2 and NIMIN-3 interact with NPR1 in yeast (Weigel et al., 2001). NIMIN-2 demotes whereas NIMIN-1 and NIMIN-3 promote, via NPR1, NPR1 interaction with AHBP-1, TGA2 and TGA6 (Weigel et al., 2001). AHBP-1b, TGA2, TGA3, TGA6, NIMIN-1, NIMIN-2 and NIMIN-3 did not interact with non-functional NPR1 in yeast (Després et al., 2000; Weigel et al., 2001; Zhang et al., 1999; Zhou et al., 2000). NIMIN-1 NIMIN-2 and NIMIN-3 mRNAs only accumulate transiently, in A. thaliana, in response to exogenously applied SA and have functional NLSs (Weigel et al., 2001). Exogenously applied SA caused the nuclear localisation of NPR1 fused to GFP, which was also demonstrated to be essential for inducing transcription of PR mRNA (Kinkema et al., 2000). Therefore NPR1 activates PR gene transcription factors in complexes with NIMIN proteins in the nucleus.

1.3.2.2.4: Unidentified signal transduction pathways

There are two unidentified signal transduction pathways. The first pathway mediates RPP7, RPP8 and RPP13 resistance; the second mediates HRT resistance. RPP7, RPP8 and RPP13-mediated resistances were independent of EDS1, NDRI, NPR1 and SA accumulation (Aarts et al., 1998; Bittner-Eddy & Beynon, 2001; Feys et al., 2001; McDowell et al., 2000). PBS2, PAD4 and RPS5 are not required for RPP13-mediated resistance (Bittner-Eddy & Beynon, 2001). RPP7-mediated resistance was also unaffected by mutations in jasmonate or ethylene signalling pathways (McDowell et al., 2000). HRT-mediated resistance is negatively regulated by RRT (regulates resistance to TCV), independent of jasmonate, ethylene and NPR1, but was SA-dependent (Cooley et al., 2000; Kachroo et al., 2000).
I.3.2.2.5: RAR1 and SGT1 signal transduction pathway components

*Rar1* (required for *Mla12* resistance 1) mutants suppressed *Mla12* resistance in barley (Jørgensen, 1996) and abolished ROS accumulation, HR generation and PR-1 mRNA accumulation (Freialdenhoven *et al*., 1994; Hückelhoven *et al*., 2000; Shirasu *et al*., 1999). *RPP5*-mediated resistance was partially compromised by *rar1*, where ROS accumulation and the HR were absent in 95% of interactions (Muskett *et al*., 2002). Many R genes that initiate resistance via EDS1 and NDR1 signal transduction pathway components are *Rar1*-dependent (Table I.6). Therefore RAR1 is a signal transduction component, involved in generating ROS and the HR for *Eds1* and *Ndr1*-dependent R genes; RAR1 is not a component of the unidentified signal transduction pathway that mediates *RPP8* resistance.

*RPP7*-mediated, but not *RPM1*, *RPS4* or *RPS5*-mediated, resistance was compromised and ROS accumulation was abolished by *sgt1* (Tör *et al*., 2002). Several R genes that generated the HR or initiated resistance via EDS1 and NDR1 signal transduction pathway components were *Sgt1*-dependent (Table I.6). Therefore, SGT1 and RAR1 are components of the same signal transduction pathway, but are differentially required for resistance mediated via the EDS1 signal transduction pathway. SGT1 was also required for resistance mediated by R genes that are dependent on MAPK and PK signal transduction components.

Table I.6: The effect of *sgt1* and *rar1*, independently, on the resistance phenotype of various R gene mediated resistances in *A. thaliana* and the ability of various R genes to generate an Avr-dependent HR after VIGS of *Sgt1* in *N. benthamiana*. R and S denote resistant and susceptible phenotypes, respectively. No HR denotes VIGS induced suppression of the HR.

<table>
<thead>
<tr>
<th>R gene</th>
<th>EDS1/NDR1-dependence</th>
<th><em>Sgt1</em></th>
<th><em>Rar1</em></th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>N</em></td>
<td>EDS1</td>
<td>S &amp; No HR</td>
<td>?</td>
<td>Peart <em>et al</em>. (2002b)</td>
</tr>
<tr>
<td>Gene</td>
<td>EDS1</td>
<td>R</td>
<td>S</td>
<td>Source</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>RPP21</td>
<td></td>
<td></td>
<td>S</td>
<td>Austin et al. (2002) &amp; Muskett et al. (2002)</td>
</tr>
<tr>
<td>RPS4</td>
<td>EDS1</td>
<td>R</td>
<td>S</td>
<td>Austin et al. (2002) &amp; Muskett et al. (2002)</td>
</tr>
<tr>
<td>RPM1</td>
<td>NDR1</td>
<td>R</td>
<td>S</td>
<td>Austin et al. (2002) &amp; Muskett et al. (2002)</td>
</tr>
<tr>
<td>RPS2</td>
<td>NDR1</td>
<td>R</td>
<td>S</td>
<td>Austin et al. (2002) &amp; Muskett et al. (2002)</td>
</tr>
<tr>
<td>RPS5</td>
<td>NDR1</td>
<td>?</td>
<td>S</td>
<td>Muskett et al. (2002)</td>
</tr>
<tr>
<td>RPP8</td>
<td>-</td>
<td>R</td>
<td>R</td>
<td>Austin et al. (2002) &amp; Muskett et al. (2002)</td>
</tr>
<tr>
<td>Cf-4</td>
<td>?</td>
<td>No HR</td>
<td>?</td>
<td>Peart et al. (2002b)</td>
</tr>
<tr>
<td>Cf-9</td>
<td>?</td>
<td>No HR</td>
<td>?</td>
<td>Peart et al. (2002b)</td>
</tr>
<tr>
<td>Pto</td>
<td>?</td>
<td>S &amp; No HR</td>
<td>?</td>
<td>Peart et al. (2002b)</td>
</tr>
<tr>
<td>Rx</td>
<td>?</td>
<td>S &amp; No HR</td>
<td>?</td>
<td>Peart et al. (2002b)</td>
</tr>
</tbody>
</table>

*A. thaliana* plants compromised in RPP5-mediated resistance by *sgt1* had a susceptible phenotype in equivalent to resistance compromised by *rar1*, but not *eds1* (Austin et al., 2002). *A. thaliana* plants compromised in RPP5-mediated resistance by *sgt1* and *rar1* had a susceptible phenotype equivalent to *rpp5*, where ROS accumulation was absent and the HR was only induced in 5% of the cells (Austin et al., 2002). VIGS of *Rar1* and *Sgt1* in barley independently impaired *Mla6*-mediated resistance in 10% and 63% of the cells, respectively, whereas VIGS of *Rar1* and *Sgt1* together in barley impaired *Mla6*-mediated resistance in 83% of cells (Azevedo et al., 2002). The effects on resistance of *sgt1* and *rar1* and of *Sgt1* and *Rar1* VIGS indicate that SGT1 and RAR1 act additively.

Barley and *A. thaliana* RAR1 interacted with *AtSGT1*, through the CHORD-II (but not CHORD-I) motif of RAR1, in yeast (Azevedo et al., 2002). Non-functional barley RAR1 also interacted with *AtSGT1* (Azevedo et al., 2002). SGT1 co-immunoprecipitated with RAR1, non-functional RAR1 or SKP1 and CUL1 members of the SCF ubiquitylation ligase complex and RAR1 or non-functional RAR1 from uninfected extracts from barley (Azevedo et al., 2002). Barley RAR1 did not co-immunoprecipitate with SKP1 and CUL1 (Azevedo et al., 2002). An SGT1-RAR1 complex (of either RAR1 or non-functional RAR1) co-immunoprecipitated with CSN4 and CSN5, COP9 signalosome...
components in extracts from barley (Azevedo et al., 2002). Therefore RAR1 and SGT1 interact and can form complexes with members of ubiquitylation (targeted degradation) and signalosome complexes, through SGT1, even with non-functional RAR1.

Ubiquitin-mediated proteolytic degradation is associated with the HR in tobacco. Karrer et al. (1998) screened a cDNA library for genes whose protein products could induce an HR in *N. tabacum* leaves and found sequences encoding ubiquitin-like proteins. Also, over-expression of a ubiquitin variant with a K48R substitution – a specific inhibitor of ubiquitin-mediated proteolytic degradation – in *N. tabacum* transformed with *N* reduced the number of HRs generated by inoculation with TMV (Becker et al., 1993). Therefore ubiquitin-mediated proteolytic degradation, which might be positively regulated by SGT1, is necessary and sufficient to generate the HR in tobacco.

**I.3.2.6 : Resistance protein mediated synthesis of the resistance protein and the signal transduction pathway components**

Signal transduction pathway component and R gene mRNA accumulation increased during defence responses, *e.g.* *EDSI* (Falk et al., 1999), *Pd4* (Jirage et al., 1999) *HsIp^pro-1* (Cai et al., 1997), *Rpw8.1* and *Rpw8.2* (Xiao et al., 2001), *WIPK* (Seo et al., 1995; Zhang & Klessig, 1998), *WRK* (Ito et al., 2002). There is also positive feedback in R gene signal transduction pathways, *e.g.* in the *EDSI*/*Pd4* signal transduction pathway (Falk et al., 1999; Feys et al., 2001). Mackey et al. (2002) found that the accumulation of RPM1 was positively correlated with RIN4 accumulation, implying positive feedback. Hence, part of R protein function is to positively regulate the synthesis of components from the recognition complex and signal transduction pathways, presumably to increase sensitivity to the attacking pathogen and amplify the active defence responses.
I.4: Cloned resistance gene evolution

I.4.1: Duplication and resistance gene clusters

R gene loci are composed of either single copy sequences or gene clusters (multiple tandem arrays of homologous sequences or paralogues. *Pi-ta, RPM1, RPP13, RPS2* and *Xa1* cloned R genes are present as single copies, although most, e.g. *Cf-2, Cf-4, Cf-5, Cf-9, Cre3, Dm3, Gpa2, I2, L, M, Mi, P2, Pib, Rp1, Pto, R1, RPP1, RPP4, RPP5, RPP8, RPS4, Rx, Sw5* and *Xa21*, are in gene clusters.

R gene clusters could have evolved from single copy R genes by gene duplication. The most likely mechanism of gene duplication into tandem arrays (Figure I.12) is sister chromatid mispairing and unequal crossover at meiosis (Pickersky, 1990) caused by repeated elements in the genome (Jeffreys & Harris, 1982).

Evidence of gene duplication exists within the sequences of R gene clusters. The intergenic regions of *Hcr4C-Cf-4* and *Hcr4A-Hcr4B* are nearly identical, strongly suggesting that the former pair are duplicates of the later (Parniske et al., 1997). The 688 bp 3’ sequences after the stop codon of *RPP8-Ler* and its parologue, *RPH8A*, are identical (McDowell et al., 1998). *Cf-2.1* and *Cf-2.2* gene sequences are identical for 787 bp at the 5’ end and are 95.3% identical for 0.4kb after their polyadenylation sites (Dixon et al., 1996). *Sw5-a* and *Sw5-b* promoter and terminator regions are highly homologous (Spassova et al., 2001). *Gpa2* and *Rx* are in the same gene cluster and have conserved intron positions (van der Vossen et al., 2000). Hence, gene duplication, by mispairing and
unequal crossover at meiosis, is the most likely mechanism for the formation of R gene clusters.

I.4.2: Rearrangement and exchange within resistance gene clusters

Duplicated genes within a cluster are susceptible to intergenic and intragenic recombination. Intergenic recombination (Figure I.12) rearranges the genes within the cluster as well as expanding and contracting the number of genes by “replication slippage” (Lovett & Feschenko, 1996), even if the difference between the sequences is up to 20% (Lewins, 1994). However, the rate of intergenic recombination is probably low because R gene orthologues – R gene homologues separated by speciation – are more similar than paralogues (Michelmore & Meyers, 1998). Intragenic recombination exchanges sequences between genes to form novel gene sequences. There is good evidence that both intergenic and intragenic recombination have occurred during R gene evolution.

Comparison of the intergenic regions between and within the Cf-4 and Cf-9 gene clusters strongly suggests rearrangement and intergenic recombination of R gene clusters. Intergenic regions of Cf-4 and Cf-9 gene clusters share homology with more than one other intergenic region in the same cluster and, where homology switches from one intergenic region to another, indicates a recombination break-point (Parniske et al., 1997; Parniske & Jones, 1999).

Intragenic recombination is implied from R gene cluster sequence. Analysis of sequences for the RPP8 and HRT loci from Ler, Col and Di-17 A. thaliana ecotypes strongly suggests RPP8-Col and HRT sequences were produced by separate intragenic recombination events between RPP8-Ler and its parologue, RPH8A (Cooley et al., 2000; McDowell et al., 1998). Alignment of gene sequences from the Xa21 locus shows sudden changes in sequence homology between paralogues around a highly conserved 233 bp region, implying it is an intragenic recombination break-point (Song et al., 1997). Regions of shared sequence homology between different paralogues at the RPP5 (Noël et al., 1999) and L (Ellis et al., 1999) loci, which implies intragenic recombination. Sequence of the P locus (Dodds et al., 2001) and the Hcr9 gene clusters (Parniske et al., 1997) suggests multiple sequence exchange between paralogues. RRS540 was probably formed by intragenic recombination between RRS1-R and RRS1-S (Deslandes et al., 2002). HR4 in the RPW8 locus from the Col-0 ecotype was probably formed
by intragenic recombination, between RPW8.1 and RPW8.2. Sequence similarity indicates 9DC was formed by intragenic recombination between Hcr9-9D and Cf-9 (Van der Hoorn et al., 2001). Also, direct repeats in the sequences of Rx (Bendahmane et al., 1999) and HR4 (Xiao et al., 2002) genes and in the LRR encoding sequence of I2 (Simons et al., 1998), L6 (Ellis et al., 1999), M (Anderson et al., 1997) and RPP5 (Noël et al., 1999) genes were probably formed by intragenic recombination.

The number of LRRs between R genes within a cluster is variable. NLOE encodes four less LRRs than Cf-9 (Parniske & Jones, 1999). Hcr4B and Cf-4 encodes 23 and 25 LRRs, respectively, whereas the other genes in the Cf-4 and Cf-9 clusters encode 27 (Parniske et al., 1997). RPP1-WsA and RPP1-WsB have 10 LRRs whereas RPP1-WsC has nine (Botella et al., 1998). Hcr2-5D encodes two more LRRs than Cf-5 and the number of LRRs encoded by Cf-2 or Cf-5 homologues from other tomato species varies (Dixon et al., 1998). RPP4 encodes four less LRRs than RPP5 (van der Biezen et al., 2002). I2C-2 encodes one more LRR than I2C-1 (Simons et al., 1998). L2 encodes four direct repeats in its LRR encoding region whereas L6 encodes two (Ellis et al., 1999). The M gene mutant, m-X39, has a direct repeat in its LRR domain (Anderson et al., 1997). An RPP5 mutant has an in-frame 270 bp duplication, which extends its LRR domain by four repeats (Parker et al., 1997). The variation in the number of LRRs in the LRR encoding domains between R gene paralogues is probably due to intragenic recombination.

Experimental evidence that intragenic recombination does occur has been obtained for the Rp1 and L gene clusters. Sudupak et al. (1993) and Richter et al. (1995) demonstrated, by monitoring flanking marker exchange, that intragenic recombination occurred at the Rp1 gene cluster and that (4 out of 143) recombination events generated novel resistance specificities. Islam & Shepherd (1991) analysed the resistance specificity of recombinants produced by intragenic recombination between different L genes and found that 11 out of 56 recombinants conferred novel resistance specificities. Therefore intragenic recombination does occur, albeit rarely, and can generate novel resistance specificities.

I.4.3: Transposon mutations and their secondary evolutionary effects

Transposons – mobile elements that can be integrated into and excised from the genome – are found at R gene loci. The Xa21 cluster contains a diverse group of 17 transposon-like elements (Song et al.,
1998), including two inserted into Xa21D and Xa21E (Song et al., 1997). Retroelements – a type of transposon – are present in the RPP5 cluster (Nöel et al., 1999; van der Biezen et al., 2002), the Hcr9 NL cluster (Parniske & Jones, 1999) and the Rp1 cluster (Collins et al., 1999). The Bs2 locus contains two Tam-like transposons and two retroelements (Tai et al., 1999). The RPW8 loci from ecotypes Ms-0 and Col-0 contain retrotransposons and possible footprint sequences (Xiao et al., 2002). The integration of transposons into or near a gene will alter the structure of the gene or, possibly, its regulation (Bennetzen, 2000; Döring & Starlinger, 1986; Li Graur, 1991).

Transposons also have secondary effects. They act as repeated elements for mispairing and unequal exchange (i.e. in duplication), either becoming recombination break-point sites or excising and disrupting the gene by leaving “footprint” sequences or deletions (Bennetzen, 2000). “Footprint” sequences are adjacent, short, direct repeats of 4-12 nucleotides caused by transposon excision and the re-ligation and repair mechanisms (Bennetzen, 2000; Döring & Starlinger, 1986; Li & Graur, 1991). For example, the short direct repeats in RPP13 alleles could be transposon “footprint” sequences (Bittner-Eddy et al., 2000). Therefore the presence of transposons in an R gene could increase the frequency of duplication and recombination or alter R gene sequences through excision.

I.4.4: Insertion and deletion

R gene evolution involves insertion and deletion (indel). Indels of one or two base pairs arise through replication errors and longer indels are caused by either intragenic recombination or transposon excision (Li & Graur, 1991; Li, 1997). Many small indels are present in the NLOA sequence from the Hcr-9 NL cluster (Parniske & Jones, 1999). L1 and L8 sequences have a deletion with respect to L6 sequence (Ellis et al., 1999). The RPP4 sequence has a deletion with respect to the RPP5 sequence (van der Biezen et al., 2002). R1 is located on a 15 kb indel that distinguishes resistant from susceptible genotypes (Ballvora et al., 2002). Indels were found between sequences of RPS2 alleles from A. thaliana and Brassica species (Wroblewski et al., 2000). RPM1 is absent from the susceptible A. thaliana ecotype Su, and instead the null locus contains a 98 bp segment of unknown origin surrounded by RPM1 flanking sequences (Grant et al., 1998). Grant et al. (1998) concluded that RPM1 had been deleted in several independent events during the evolution of the Brassicaceae.
I.4.5: Positive, diversifying, adaptive selection of resistance gene sequences

Nucleotide substitutions (point mutations) of a group of related sequences are informative and can indicate evolutionary direction (McDonald & Kreitman, 1991). Redundancy in the genetic code means that a nucleotide substitution in a codon does not always change the amino acid in the protein sequence. Nucleotide substitutions that change an amino acid in the protein sequence are known as non-synonymous and those that don’t change the protein sequence are known as synonymous. The ratio (Ka/Ks) of the number of non-synonymous substitutions per potential codon (Ka) to the number of synonymous substitutions per potential codon (Ks) for a group of related sequences indicates whether the sequences are prone to nucleotide substitutions which a) alter the amino acid sequence (Ka/Ks > 1), b) conserve the amino acid sequence (Ka/Ks < 1), or c) are neutrally selective (Ka/Ks ~ 1). If the Ka/Ks ratio is less than one then functional constraints mean sequence conservation predominates. For Ka/Ks ratios greater than one divergence predominates and the sequences are considered to be under positive, diversifying or adaptive selection.

The LRR encoding regions from paralogues of I2 (Simons et al., 1998) and Xa21 (Song et al., 1997) are the most variable regions. The LRR encoding sequences were also the most variable region in the RPP5 gene clusters of A. thaliana ecotypes Ler and Col-0 and are subject to diversifying selection (Nöel et al., 1999). Diversifying selection was shown for 11 sequences in the Cf-0, Cf-4 and Cf-9 gene clusters; there were seven hypervariable regions between the sequences and all were located in the LRR encoding sequence (Parniske et al., 1997). The nucleotide substitutions between RPP8-Ler and RPP8-Col alleles in the LRR encoding domain (McDowell et al., 1998) and the coding sequences of RPP13-Nd and RPP13-Col alleles (Bittner-Eddy et al., 2000) showed these sequences were subject to diversifying selection. The LRR encoding sequences of Xa21 and Xa21D are subject to diversifying selection (Wang et al., 1998). Within the Cf-9 gene cluster the first 16 LRRs of all the Hcr9s were subject to diversifying selection (Parniske et al., 1997). The N terminal pre-LRR encoding region of RPP13 alleles had a Ka/Ks ratio of less than one (Bittner-Eddy et al., 2000). Therefore analysis of Ka/Ks ratios between R genes or within R gene clusters has shown that R genes are subject to diversifying
selection, which is particularly evident in the highly divergent, LRR encoding regions. Conversely, non-LRR encoding regions are conserved.

$K_a/K_s$ ratios for nucleotide substitutions between $Mla1$ and $Mla6$ or $Gpa2$ and $Rx$ sequences and within the $Dm3$ cluster were approximately one in the N terminal pre-LRR region and significantly greater than one in the C terminal LRR region, mainly due to non-conserved residues in the solvent-exposed $\beta$-strand, $\beta$-turn motif (Halterman et al., 2001; Meyers et al., 1998; Van der Vossen et al., 2000). The divergence between $L$, $RPP8$ and $RPP13$ allele sequences was especially high in sequence coding for the solvent-exposed $\beta$-strand $\beta$-turn motif of the LRRs (Bittner-Eddy et al., 2000; Ellis et al., 1999; McDowell et al., 1998); the remainder of the $RPP8$ allele LRRs had a $K_a/K_s$ ratio less than 1 (McDowell et al., 1998). Within the $RPP1$ cluster $K_a/K_s$ ratios indicated that the coding sequences are neutrally selected, although the solvent-exposed LRR regions are subject to divergent selection (Botella et al., 1998). $K_a/K_s$ ratios between orthologues of $P2$ show diversifying selection only in the solvent-exposed $\beta$-strand $\beta$-turn motif of the LRRs (Dodds et al., 2001). Therefore in the LRR encoding regions there is conservation in the sequence encoding the hydrophobic backbone, and divergence in the sequence encoding the hypervariable, hydrophilic, interacting residues.

Diversifying selection is rarely seen in genomes. Endo et al. (1996) searched various databases for groups of genes that were subject to diversifying selection and found 17 (0.47% of the total) gene groups. Of these 17 groups, nine were for sequences encoding surface proteins of parasites and viruses. The variable regions of mouse and human immunoglobulin genes are also subject to diversifying selection in the antigen recognition site, but not in other regions (Hughes & Nei, 1988; Tanaka & Nei, 1989). Therefore, the majority of sequences subject to diversifying selection are the elicitors and receptors of host-pathogen systems.

**I.4.6: Resistance gene recognition specificity evolution: the “boom and bust” cycle**

Rapid adaption to resistance in the pathogen – due to a greater number of individuals and hence a greater gene pool – necessitates rapid evolution of new resistance specificities in the host; such rapidly
evolving specificities must come from a reservoir of potential structures (Pryor & Ellis, 1993; Staskawicz et al., 1995). By comparison with the evolution of novel specificities at the self-incompatibility loci, Pryor & Ellis (1993) proposed that intragenic recombination between R gene orthologues and paralogues was the only feasible mechanism capable of generating novel R gene recognition specificities at the rate required to keep virulent pathogens at bay. For instance, intragenic recombination at the \( RpI \) (Sudupak et al., 1993; Richter et al., 1995) and the \( L \) (Islam & Shepherd, 1991) produced novel resistance specificities. Also, Crameri et al. (1998) simulated intragenic recombination by DNA shuffling and found that it accelerated directed evolution. However, Michelmore & Meyers (1998) concluded that intragenic recombination has been a rare event between paralogues. As argued by Hughes (1994), duplicated sequences are not free from functional constraints and must diverge to obtain new specificities. For instance, eosinophil cationic protein evolved by duplication – from eosinophil-derived neurotoxin – and diverged to acquire a novel anti-pathogen function (Zhang et al., 1998). Intragenic recombination between orthologous R gene sequences followed by diversifying selection is the basis for the “birth and death” model (Michelmore & Meyers, 1998) – a molecular interpretation of how novel R gene recognition specificities evolve. It is not clear how these novel R gene recognition specificities spread through the gene pool: Thompson & Burdon (1992) were unable to conclude if frequency-dependent or density-dependent selection could account for R allele polymorphism.

Stahl et al. (1999) analysed the polymorphism of flanking sequences at the \( RPM1 \) locus. They claim that flanking sequence polymorphism has persisted for so long that the \( RPM1 \) allele frequency can’t simply be moving towards fixation in the gene pool and is actually in dynamic equilibrium, probably due to the cost of resistance. However, the arrival of novel resistance specificities is most likely to be due to intragenic recombination, which does not involve new, polymorphic flanking sequences. There are significant flaws in the work by Stahl et al. (1999). Stahl et al. (1999) did not analyse the polymorphism of \( RPM1 \) sequences at the \( RPM1 \) locus; not all the \( RPM1 \) loci containing \( RPM1 \) sequence are necessarily functional. The weakest aspect of the work by Stahl et al. (1999) was that they did not understand the concepts behind molecular evolution and applied the wrong analytical tools. The polymorphism between \( RPM1 \) flanking sequences was between different accessions of \( A. thaliana \). An accession is a geographically localised population and, by definition, a gene pool. Therefore, the polymorphism analysed by Stahl et al. (1999) was actually between gene pools, which
do not interbreed. Hence, because Stahl et al. (1999) analysed the frequency of RPM1 flanking sequences in separate populations (which do not interbreed), it is hardly surprising that the RPM1 allele frequency had not become fixed (100%) across these gene pools. The conclusions of Stahl et al. (1999) should be rejected.

I.4.7: Resistance gene relationships

R gene clusters appear to be related. Nucleotide sequences of M and L6 are 86% identical (Anderson et al., 1997), implying relatedness. Mi-1.1 and Mi-1.2 sequences are 97% identical (Milligan et al., 1998). The intron boundary 41bp at the 3’ end of the termination codon and the polyadenylation signal are absolutely conserved between Cf-2 and Cf-5 (Dixon et al., 1998). Lipoxygenase sequence flanks the Cf-4 and Cf-9 gene clusters NI, MW and SC, which led Parniske et al. (1997) and Parniske & Jones (1999) to propose that the Cf-4 and Cf-9 gene clusters were derived from a common ancestor. Indeed, the translocation of sequences between non-homologous chromosomes could move an R gene sequence to new locations (Pickersky, 1990), potentially forming new R gene clusters. Therefore, relatedness between certain R gene clusters suggests they belong to phylogenetically linked R gene families.

R gene families may be related too. The NBS encoding sequences from a wide variety of plant species formed two clades: one contained TIR encoding sequence and the other clade contained non-TIR encoding sequence (Meyers et al., 1999; Pan et al., 1999; Richly et al., 2002). Included in the TIR encoding sequences were L6, M, N, RPP1 and RPP5. The’non-TIR encoding sequences included Cre3, Dm3, I2, Mi, Pib, Prf, RPM1, RPP8, RPS5 and Xa1. NBS encoding sequences were present in the gymnosperms and angiosperms, but surprisingly, sequences encoding TIR and NBS domains were not found in the monocots (Meyers et al., 1999; Pan et al., 1999). The results of these two studies imply that the NBS encoding gene family is ancient but the TIR and NBS encoding family has been lost from the monocots after the dicot-monocot divergence. Hence, NBS encoding sequence R gene families appear to be related and form two superfamilies, one encoding TIRs and the other encoding non-TIR sequences.

I.5: Summary
The presence of R genes in host plants permits recognition of pathogens carrying the corresponding “host-specific” Avr gene by the host plant and induction of active defence responses (including the HR and SAR) in a gene-for-gene manner. Hence, the R gene must produce a protein (R protein) which performs two functions, a) specifically recognising the Avr protein and b) initiating the induction of active defence responses (which they can do cell-autonomously). The elicitor-receptor model predicts that the R protein is an Avr protein specific receptor at the site of elicitor recognition, which is dependent on the modus operandi of the pathogen. However, most R proteins do not interact directly with the Avr protein elicitor, although they do determine the specificity of elicitor recognition, possibly as part of a recognition complex. Except for Pto the specificity of elicitor recognition is dependent on the LRR domain, a highly variable, versatile and adaptive binding domain. R proteins are sometimes membrane-associated and composed of a just few structural motifs, including PK, transcription factor and TIR-like motif signal transduction domains. There are no dispensable regions within R proteins, indicating that all the motifs are necessary for function. There is emerging evidence that R protein signal transduction requires the oligomerisation and phosphorylation of the R protein and involves lipases, ubiquitin-mediated proteolytic degradation and Apaf-1/CED-4 activated apoptotic caspases. R protein signal transduction pathways can be defined by EDS1, NDR1, SGT1 and RAR1 components (Figure I.13).

**Powdery mildew disease resistance genes in**

*Arabidopsis thaliana* (L.) Heyhn.

Powdery mildews are economically important crop pathogens. A greater understanding of resistance to powdery mildew diseases could present new, possibly non-chemical, opportunities for crop protection.
Plant disease resistance (R) genes for powdery mildew pathogens of Arabidopsis have been identified. One R gene locus, RPW8, was able to confer resistance to all the known Arabidopsis powdery mildew pathogens (Xiao et al., 2001). The RPW8 locus is particularly interesting because the predicted polypeptides of its two functional genes, RPW8.1 and RPW8.2, are unique.

II.1: Powdery mildews

Powdery mildews (the Erysiphales) are obligate, ascomycete biotrophs of the Fungi Imperfecti (Agrios, 1988; Braun, 1987; Hudson, 1986; Yarwood, 1978). The powdery mildews acquired their name due to the characteristic appearance of abundant hyphae and conidiophores on the leaf surfaces of infected plants. The sexual stages of powdery mildews have not been observed on A. thaliana (Crute et al., 1994; Hall, 1994). The perithecia of E. cruciferarum seldom developed on a range of species tested by Junell (1967). Yarwood (1978) defined powdery mildews, where the perfect (sexual) stage is either not present or difficult to see, as

“fungi with white superficial hyphae on the aerial parts of living plants, with large one-celled conidia produced terminally on isolated aerial unbranched conidiophores and with haustoria in the epidermal cells of their hosts”.

If the perfect stage is not present powdery mildew infections spread via asexual, airborne spores (Agrios, 1988; Hudson, 1986). These spores land on the aerial surfaces of plants, especially the leaves. Spores on leaf surfaces of susceptible plants germinate, form appressoria, penetrate the cuticle and develop haustoria in the epidermal cells. Within a few days hyphal growth produces an extensive, superficial mycelium on the leaf surface. Conidiophores, emerge from the mycelia, perpendicular to the surface of infected organs, and produce conidia (the cylindrical asexual spores). The asexual spores are released into air currents passing over the leaf surface to repeat the cycle.

Powdery mildews are probably the most common cause of plant disease, affecting many economically important plants world-wide (Agrios, 1988; Spencer, 1978; Crute et al., 1994). Crop quality and yield
can be significantly reduced by a combination of the following: a direct loss of yield, suppression or distortion of growth and consequently indirect yield loss, spoilage of fruit, or disfigurement of ornamentals (Bent, 1978). Cereals, such as wheat, barley and oats, are severely affected by powdery mildew disease (Colhoun, 1971), although vegetables, ornamentals, fruit and forest trees are also affected (Agrios, 1988; Dixon, 1981).

At present powdery mildew diseases are primarily controlled with fungicides (Agrios, 1988; Bent, 1978; McGrath, 2001; McGrath & Shishkoff, 2001; Wolfe, 1984). Despite the effectiveness of these fungicides, they are toxic (Renner & Nguyen, 1982; Rogers et al., 1986; Suwalsky et al., 2000), expensive, may require frequent applications and powdery mildews can evolve insensitivity to them (Bent, 1978; McGrath, 2001; McGrath & Shishkoff, 2001; Wolfe, 1984). Resistant crops have been successfully used to control mildews (Bent, 1978; Wolfe, 1984) and could limit the need for fungicides (Bingham, 1981; Russel, 1978; Wolfe, 1984). Plant breeders have introgressed resistant traits from donor cultivars into closely related, agronomically superior cultivars and then selected for agronomically superior and resistant traits whilst diluting out the donor genome (Johnson, 1981; Winter & Kahl, 1995), albeit not always durably (Russel, 1978; Wolfe, 1984). However, plant breeding is costly, labour intensive and time consuming (Winter & Kahl, 1995) and a species of interest may not be sexually compatible with the resistant cultivar. Knott & Dvořák (1976) proposed transferring R genes from resistant species to sexually incompatible and susceptible crop species. Therefore, by transferring R genes for powdery mildew resistance into (i.e. genetically modifying) susceptible crop plant species it may be possible to generate durable, non-chemical resistance to powdery mildews.

II.2: *Arabidopsis* as a model plant for studying plant pathogen interactions

*Arabidopsis thaliana* was chosen as a likely source of R genes, to isolate and characterise, with the intention of applying the knowledge to crop protection.

*Arabidopsis thaliana* is the model flowering plant (Meyerowitz & Somerville, 1994). Several features have led to its ubiquity in plant biological research. It was attractive to geneticists because it is small, with a short generation time (approximately 4 weeks) and produces copious
amounts of seeds (more than $10^4$ per plant). *A. thaliana* was ideal for molecular biological investigation (Holub & Beynon, 1997; Leutwiler *et al*., 1984; Pruitt & Meyerowitz, 1986; The *Arabidopsis* Genome Initiative, 2000), which simplifies gene isolation and cloning because it is: comprehensively studied using classical genetics, has accession stocks, a high efficiency of stable transformation (Schmidt & Willmitzer, 1988). Consequently high density genome maps in yeast and bacterial artificial chromosomes (YACs and BACs) and an expressed sequences tag (EST) database are available for the small (125Mb), fully sequenced, predominantly single copy, diploid genome. Resistance in *A. thaliana*, to a wide variety of bacterial, viral and fungal pathogens, is similar to other flowering plants (Buell, 1998; Dangl, 1993; Glazebrook *et al*., 1997; Kunkel, 1996; Mauch-Mani & Slusarenko, 1993). Hence, *Arabidopsis thaliana*, the model flowering plant, has been adopted as the model plant pathology system (Buell, 1998; Dangl, 1993; Glazebrook *et al*., 1997; Kunkel, 1996; Mauch-Mani & Slusarenko, 1993).

### II.3: Resistance to powdery mildews in *Arabidopsis thaliana*

The symptoms of powdery mildew disease caused by *E. cichoracerum* and *E. cruciferarum* on susceptible *A. thaliana* plants differ. *E. cichoracerum* forms dense white mycelia, apparent at 5 days post-inoculation (dpi), which coalesce to cover the entire leaf surface of ecotypes Col-0 and Ler and produces abundant conidiophores by 7-8 dpi (Adam & Somerville, 1996; Adam *et al*., 1999; Schiff *et al*., 2001; Wilson *et al*., 2001; Xiao *et al*., 1997). In contrast, to the unaided eye *E. cruciferarum* infected Col-0, Ler or Wei-0 accession leaves appear symptomless. *E. cruciferarum* hyphal growth was thin and inconspicuous, and conidiophore formation at 5 dpi was sparse (Adam *et al*., 1999; Hall, 1994; Koch & Slusarenko, 1990; Xiao *et al*., 1997).

Adam & Somerville (1996) found five independent plant disease resistance loci, on chromosomes II, III, IV and V (Table 2), that limited *E. cichoracerum* isolate UCSC1 growth on *A. thaliana* accessions Kas-1, Wa-1, Te-0, Stw-0 & Su-0. These disease resistance loci were named *recognition of powdery mildew* 1-5 (*RPW1-RPW5*).

In a screen of 50 *A. thaliana* accessions, Hall (1994) discovered that accession Ms-0 was very resistant
to *E. cruciferarum* isolate UEA1 – preventing sporulation and reducing hyphal growth. Accession Ms-0 was found by Xiao *et al.* (1997) to be very resistant to *E. cichoracerum* isolate UCSC1 as well. The growth of isolate UCSC1 on accession Ms-0 leaves was arrested after the emergence of several germ tubes from the conidium, with HR cell death in the mesophyll cells at the site of infection, under the germinating conidium. The germination rate of isolate UEA1 on leaves of accession Ms-0 was less than 50%, and growth was arrested after the formation of the primary appressorium, with collapse of the underlying host cell, and occasionally the HR cell death of one or two host mesophyll cells below the germinated conidia.

Fig. II.1: *A. thaliana* plants inoculated with *E. cichoracerum* isolate UCSC1. Accession Ms-0, a), was resistant and accession Col-0, b), was susceptible. Images courtesy of S. Xiao.

Genetic dissection of resistance in accession Ms-0 found that it was due to a single locus (*RPW8*) for *E. cichoracerum* isolate UCSC1, and two independently segregating loci (*RPW6* and *RPW7*) for full resistance to *E. cruciferarum* isolate UEA1 (Xiao *et al.*, 1997). Individually, *RPW6* and *RPW7* conferred intermediate resistance (Xiao *et al.*, 1997). Xiao *et al.* (1997) mapped *RPW6*, *RPW7* and *RPW8*, with SSLP and CAPS, and located them to loci in the *A. thaliana* genome not previously identified as *RPW1-RPW5* by Adam & Somerville (1996). *RPW6* was located on chromosome V
RPW7 and RPW8 were mapped to the same location in an 8.5 cM interval between CAPS markers CDC2A & AFC1 on chromosome III (Table II.1 & Figure II.2). Resistance to isolate UCSC1 co-segregated with RPW7, which also conferred resistance to isolate UEA1. It was thought that RPW7 and RPW8 were the same gene, thereafter referred to as RPW8, with dual resistance to *E. cichoracerum* isolate UCSC1 and *E. cruciferarum* isolate UEA1.

Quantitative trait loci (QTL) analysis has been used to dissect polygenic resistance to *E. cichoracerum* isolate UCSC1 in *A. thaliana* accessions Kas-1 (Wilson et al., 2001) and Wa-1 (Schiff et al., 2001). In accession Kas-1 Wilson *et al.* (2001) identified three independent, additively-acting QTLs, RPW10, RPW11 and RPW12. In accession Wa-1 Schiff *et al.* (2001) identified two independent, additively-acting QTLs, RPW13 and RPW14. RPW10 and RPW13 QTLs both mapped to CAPS marker R30025 on chromosome III (Table II.1 & Figure II.2). RPW11, RPW12 and RPW14 QTLs mapped to locations on chromosomes V, II, and I respectively (Table II.1 & Figure II.2). The RPW10 and RPW13 QTLs are the strongest contributors to resistance in accessions Kas-1 and Wa-1 respectively.

Comparison of map positions for the RPW loci suggests that some, like RPW7 and RPW8, are actually the same locus, especially considering that the confidence intervals for RPW10-RPW14 QTLs are approximately 10 cM (Wilson *et al.*, 2001; Schiff *et al.*, 2001). RPW1 and RPW12 were both cloned from the same accession challenged with the same pathogen (Table II.1). The confidence interval of the RPW12 QTL (Wilson, *et al.*, 2001) lies across the mapped location of RPW1 (Adam & Somerville, 1996) – interval nga1126-nga168 (Table II.1 & Figure II.2). Mapping of RPW10 and RPW13 QTLs located them both to CAPS marker R30025 on chromosome III (Wilson *et al.*, 2001; Schiff *et al.*, 2001). The RPW10 QTL was mapped to a 6 cM (500 kb) confidence interval around CAPS marker R30025 (Wilson *et al.*, 2001). The RPW13 QTL was mapped to an approximately 1000kb interval between CAPS markers CIC3D2L and R30025, and to a 6 cM confidence interval around CAPS marker R30025 (Schiff *et al.*, 2001). The stated positions of RPW8 (Xiao *et al.*, 1997), RPW10 (Wilson *et al.*, 2001) and RPW13 (Schiff *et al.*, 2001) are less than 10 cM apart. This separation is within the average accuracies (confidence intervals) for these QTLs (Wilson *et al.*, 2001) and phenotypically all three loci confer strong resistance to isolate UCSC1 (Schiff *et al.*, 2001; Wilson *et al.*, 2001; Xiao *et al.*, 1997). Markers known to be either side of the RPW8 locus were found to encompass the RPW10 and RPW13 QTLs; and a YAC containing the RPW8 locus was homologous to
RPW10 and RPW13 QTLs (J.G. Turner, pers. comm.; S.C. Somerville, pers. comm.). Peptide sequences of Kas-1 and Wa-1 ecotype orthologues of RPW8.1 and RPW8.2 were identical to RPW8.1 and RPW8.2 from accession Ms-0 (Xiao et al., 2001). Therefore, RPW1 and RPW12 are the same locus; and RPW10 & RPW13 are RPW8 alleles from accessions Kas-1 and Wa-1 respectively.

II.4: RPW8

II.4.1: Map-based cloning of RPW8

RPW8 was fine-mapped by Xiao et al. (2001) to an interval between CAPS markers CDC2A and g19397 (Figure II.3). YAC, BAC and cosmid constructs further narrowed the region containing the RPW8 locus to cosmid B6 (Figure II.3). Susceptible, Col-0 accession plants transformed with cosmid B6 were resistant to E. cichoracerum isolate UCSC1 (Xiao et al., 2001).

Sequence analysis of cosmid B6 with NIX suggested that there are three open reading frames (ORFs, Figure II.3ii). A. thaliana Col-0 accession plants were transformed with cosmid B6 subclones containing genomic sequences with combinations of the three ORFs. The transformants were screened with E. cichoracerum isolate UCSC1 and only those that contained genomic sequences with either of two ORFs were resistant (Figure II.3ii, Xiao et al., 2001). To confirm that RPW8.1 and RPW8.2 could independently confer resistance, Col-0 accession plants were transformed with RPW8.1 or RPW8.2 cDNAs driven by the CaMV-35S promoter (35S::RPW8.1 and 35S::RPW8.2). Accession Col-0 plants transformed with 35S::RPW8.1 and 35S::RPW8.2 were resistant to E. cichoracerum isolate UCSC1. Accession Col-0 plants transformed with 35S::SKP-2, the other of the two ORFs were not resistant to isolate UCSC1. It was concluded that the RPW8 locus has two ORFs, RPW8.1 and RPW8.2 (Table II.2), which can independently impart a resistant phenotype on previously susceptible Col-0 accession plants.

The amino acid sequences of the RPW8 predicted polypeptides (Table II.2) are small, basic and unlike any other known protein (Xiao et al., 2001), although they are predicted to contain recognisable motifs. Putative coiled-coil (CC) and trans-membrane (TM) domains were found in RPW8.1 and RPW8.2 predicted polypeptides by Xiao et al. (2001).
**Table II.2: RPW8.1 & RPW8.2 ORF sequences (lengths in nucleotides), protein molecular weights (Mr) and Isolectric points (Ie) predicted from cosmid B6 genomic insert sequence data (GenBank accession number AF273059).**

<table>
<thead>
<tr>
<th>Transcript</th>
<th>Intron</th>
<th>Coding Sequence</th>
<th>Mr</th>
<th>Ie</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPW8.1</td>
<td>711</td>
<td>197</td>
<td>444</td>
<td>17,000</td>
</tr>
<tr>
<td>RPW8.2</td>
<td>798</td>
<td>128</td>
<td>522</td>
<td>19,973</td>
</tr>
</tbody>
</table>

**II.4.2: RPW8-mediated broad spectrum disease resistance**

Col-0 accession plants transformed with 35S::RPW8.1, 35S::RPW8.2, cosmid B6, subclone EP3.7 or subclone XE3.8 were resistant to a broad spectrum of geographically distant powdery mildew pathogens: 15 isolates of *E. cichoracearum* (including UCSC1), *E. cruciferarum* isolate UEA1, *E. orontii* isolate MGH, *O. lycopersicon* isolate Oxford (Xiao et al., 2001; Figure II.4). These are all the known powdery mildew pathogens of *A. thaliana* (Adam & Somerville, 1996; Koch & Slusarenko, 1990; Plotnikova et al., 1998). However, Col-0 ecotype plants transformed with RPW8 were not resistant to the downy mildew fungus *Peronospora parasitica* isolate Noco2, Cauliflower mosaic virus, or the bacterium *Pseudomonas syringae pv. tomato* DC3000 (Xiao et al., 2001; Figure II.4).

*RPW8*-mediated resistance is effective in *A. thaliana* against powdery mildews that pathogenise a range of other hosts. *E. cichoracearum* can infect begonia, chrysanthemum, cosmos, dahlia, flax, lettuce, phlox, zinnia and the cucurbit (*Cucurbitaceae*) family (Agrios, 1988; Lebeda & Kríštová, 1994; Lebeda & Kríštová, 1996); isolate UCSC1 can infect squash cv. Kuta (Plotnikova et al., 1998). *E. cruciferarum* pathogenises the *Cruciferaceae*, which includes crop plants such as cabbages, cauliflowers, brussel sprouts, turnips and swede (Junell, 1967). *E. orontii* isolate MGH was isolated from *A. thaliana* by Plotnikova et al. (1998) and also pathogenised *Capsella bursa-pastoris* and squash cv. Kuta, but not other cucurbits, radish or arugula. *O. lycopersicon* isolate Oxford mainly pathogenises members of the *Cucurbitaceae* and *Solanaceae* (Fletcher & Smewin, 1988; Whipps et al., 1998). Therefore RPW8 confers resistance to a broad spectrum of powdery mildew pathogens that also cause disease on a wide range of economically important crop species.
Structural predictions of RPW8.1 and RPW8.2 predicted polypeptides

RPW8 predicted polypeptides (Figure III.1) are small, basic and unlike any other known protein (Xiao et al., 2001). However, Xiao et al. (2001) found two intriguing pieces of structural information: a) putative CC and TM domains, and b) similarity (51% and 62%, respectively) and identity (22% and 33%, respectively) between RPW8.1 and RPW8.2 predicted polypeptides with the 130 N-terminal residues of BAB08633. BAB08633 is an 815 nucleotides long A. thaliana sequence, registered as disease resistance protein-like in GenBank (accession No. AB010700.1) and has putative NBS and LRR motifs (Xiao et al., 2001).
RPW8.1
MPIGELAIGAVLGVGAQAIYDRFRKARDISFVHRLCATILSIEPFLVQIDKRSKVEGSLREVINELTC
FLEAYVFVEAYPKLRRRQVLKRYIKAIEHTIERTLSIIIVDFQVDDIKEIKAKISEMDTKLAE
VISCASKIRA

RPW8.2
MIAEVAAGGALGLSLHEAVKRAKDRSVTTRFILHRLEATIDTSITPLVVQIDKFSEEMEDSRSKVN
KRLKLLENANESLVEENAESRNNVRKFRYMDIKEFEAKLRWVVDVQVQNADKELKAKMSEIS
TKLDKIMFPKFEIHIGWCSGNRRAIRFTFCSDDS

Since Xiao et al. (2001) structurally analysed the predicted polypeptides of RPW8, protein sequences have been added to the databases and prediction algorithms have improved. RPW8 predicted polypeptides were assessed for the presence of CCs, TM helices, protein targeting signals and post-translational modification sites. Ab initio tertiary structure predictions for RPW8 predicted polypeptides were also attempted.

III.1: Coiled-coils predictions

There are two algorithms capable of predicting CCs from a polypeptide sequence, COILS (Lupas, 1991; Lupas, 1996) and PAIRCOIL (Berger et al., 1995). COILS and PAIRCOIL both compare the CC forming potential (probability) of a residue, within its local sequence, with a probability matrix. Probability matrices (e.g. COILS MTK or MTIDK) are constructed from databases of known CC forming polypeptide sequences. The potential to form a CC is determined for each heptad position. Hence, each residue in a polypeptide will be predicted a heptad position and a CC forming, probability depending on the residues around it in the polypeptide sequence.

COILS matrices MTK and MTIDK (weighted and unweighted) predicted solvent exposed, left-handed CCs (Lupas, 1996) – P > 0.5 – in RPW8.1 and RPW8.2 predicted polypeptides (Table III.1). Probabilities from weighted and unweighted matrix databases were within 20% of each other, suggesting the COILS predicted heptad repeats in RPW8.1 and RPW8.2 are genuine, rather than a cluster of highly hydrophilic and charged residues (Lupas, 1996). PAIRCOIL predicted a region of elevated CC forming potential in RPW8.1 and RPW8.2 predicted polypeptides with the same heptad repeat as predicted by COILS. However, the CC forming probabilities in RPW8.1 and RPW8.2 predicted polypeptides were lower than the PAIRCOIL threshold, 0.5. COILS (estimated with the 21
residue window) and PAIRCOIL disagree where the ends of the CCs lie. However, of the two algorithms, COILS is reported to be the more accurate for short heptad repeats (Lupas, 1997) like those in RPW8 proteins. Many CC predictions were proven correct when the crystal structures of their respective proteins were determined (Lupas, 1997). Thus, RPW8.1 and RPW8.2 predicted polypeptides have the potential to form CCs.

Table III.1: Predictions of CC forming regions in RPW8.1 and RPW8.2 predicted polypeptides by COILS (weighted and unweighted MTIDK & MTK matrices) and PAIRCOIL algorithms.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Matrix</th>
<th>Weighted/Unweighted</th>
<th>Region of polypeptide</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPW8.1</td>
<td>COILS</td>
<td>MTIDK</td>
<td>W</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>120-140</td>
</tr>
<tr>
<td></td>
<td></td>
<td>U</td>
<td>120-140</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MTK</td>
<td>W</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>120-142</td>
</tr>
<tr>
<td></td>
<td></td>
<td>U</td>
<td>120-142</td>
</tr>
<tr>
<td></td>
<td>PAIRCOIL</td>
<td>PAIRCOIL</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>119-148</td>
</tr>
<tr>
<td>RPW8.2</td>
<td>COILS</td>
<td>MTIDK</td>
<td>W</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>66-97</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>121-145</td>
</tr>
<tr>
<td></td>
<td></td>
<td>U</td>
<td>66-97</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>121-145</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MTK</td>
<td>W</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>66-95</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>121-145</td>
</tr>
<tr>
<td></td>
<td></td>
<td>U</td>
<td>68-93</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>121-145</td>
</tr>
<tr>
<td></td>
<td>PAIRCOIL</td>
<td>PAIRCOIL</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>62-97</td>
</tr>
</tbody>
</table>

Even though RPW8.1 and RPW8.2 predicted polypeptides have been assigned CCs, this does not necessarily mean that the polypeptide will form a CC. COILS and PAIRCOIL predictions indicate CC forming potential (Lupas, 1997), which is dependent on the environmental conditions surrounding the protein as well as the polypeptide sequence (Anfinsen, 1973). For example, the region of influenza haemagglutinin HA2 with the greatest CC potential does not form a CC in the native protein at neutral pH, but it does in an acidic environment – pH4.8 (Carr & Kim, 1993).

III.2: Trans-membrane helix predictions

A hydrophobicity plot can identify a putative TM helix (Alberts et al., 1994; Stryer, 1995) providing it is not amipathic (Engelman et al., 1986). Hydrophobicity plots graphically represent the free energy of transfer to water (in kcal/mol) of an α-helix by calculating the sum of hydrophobicities for residues in the window around the reference amino acid. If a sufficiently long (approximately 20 residue) section of the peptide sequence prefers to be in a hydrocarbon, rather than an aqueous environment, it is likely to be a TM helix.
The White-Wimley (WW) experiment-based whole-residue hydrophathy scale (White & Wimley, 1999) performs better than the GES (Engelman et al., 1986) or the KD (Kyte & Doolittle, 1984) scales at predicting the TM helices of membrane proteins in the MPTopo database (Jayasinghe et al., 2001). TM helices are most accurately and clearly predicted with the augmented WW octanol-interface (Oct-IF) scale compared with the prediction criteria of the GES and KD scales. Jayasinghe et al. (2001) assigned this greater accuracy to the inclusion of the backbone dehydration cost (~1.2 kcal/mol) in the augmented WW scale. Therefore the Oct-IF scale is the preferred algorithm for detecting TM helices in RPW8 predicted polypeptides.

The MPEx1.50 algorithm is based on the Oct-IF scale. MPEx1.50 predicted TM helices in RPW8.1 and RPW8.2 predicted polypeptides (Figure III.2). The predicted TM helix in the RPW8.1 predicted polypeptide (Figure III.2a) is centred on residue 39T, with sequence SFVHRLCATILSIEPFLVQ. The RPW8.2 predicted polypeptide is predicted to have a TM helix (Figure III.2b) centred on residue 44D, with sequence ILHRLEATIDSITPLVVQI.

Möller et al. (2001) have extensively evaluated membrane topology prediction algorithms, especially TM helix detection. TMHMM2.0 (Sonnhammer et al., 1998) was the best algorithm for recognising true TM helices in known membrane proteins and had the lowest proportion of false positive predictions in soluble proteins; MEMSAT2 (Jones et al., 1994) was second best. Neither TMHMM2.0 nor MEMSAT2.0 algorithms predicted TM helices in RPW8.1 and RPW8.2 predicted polypeptides.

Some TM predictions are false positives. Signal peptides can be incorrectly predicted as TM helices (Möller et al., 2001). Hydrophobic regions, possibly occupying the interior of the proteins, known as interfacial (IF) α-helices can be mistakenly predicted to be TM helices (White & Wimley, 1999). Möller et al. (2001) suggest that SignalP, an algorithm for predicting the presence of a secretory signal peptide in a polypeptide sequence, should be used in parallel with algorithms for TM helix predictions. Hence, predictions of TM helices at the N termini of RPW8.1 and RPW8.2 predicted polypeptides by Xiao et al. (2001) might have been regions of hydrophobicity or secretory signal peptides.

III.3: Predicted subcellular targeting of the RPW8 predicted polypeptides
Proteins can be targeted to different cellular or extracellular compartments and locations. It is possible to imply the cellular location of a protein from the polypeptide sequence. Proteins in different cellular compartments differ in amino acid composition (Nielsen, 1999) and it is possible to predict the subcellular localisation of a protein with amino acid composition data (Nakai, 2000). Also, polypeptide sequence motifs or consensi are associated with compartmental targeting.

Secretory signal peptides have three distinct regions (Figure III.3): a positively charged N terminal \( n \) region, a central hydrophobic \( h \) region and polar \( c \) region (von Heijne, 1985). The \( c \) region is approximately five residues long and encodes the cleavage site (von Heijne, 1985). Signal anchors have an \( n \) region and a longer \( h \) region (approximately as long as a TM helix), but no \( c \) region (Nielsen et al., 1999).

Fig. III.3: Representation of a secreted polypeptide’s primary structure. The secretory signal peptide is located at the N terminus of a secreted polypeptide sequence.

SignalP predicts the presence of secretory signal peptides or signal anchors in a polypeptide sequence. SignalP2.0 has two algorithms, SignalP-NN (SignalP1.1, Nielsen et al., 1997) and SignalP-HMM (Nielsen & Krogh, 1998). SignalP-NN is better than SignalP-HMM at determining the presence of a secretory signal peptide, whereas SignalP-HMM is better at discriminating between secretory signal peptides and signal anchors (Nielsen & Krogh, 1998). SignalP-NN did not predict a secretory signal peptide in the RPW8.1 predicted polypeptide, but did in the RPW8.2 predicted polypeptide. However, SignalP-NN did not strongly predict a cleavage site for a secretory signal peptide in the predicted
polypeptide sequence of RPW8.2. SignalP-HMM analysis of the N terminus of RPW8.2 showed a signal peptide sequence with a c region, consistent with the primary structure of a secretory signal peptide but not a signal anchor. Therefore the predicted polypeptide sequence of RPW8.2 may contain a secretory signal peptide with a cleavage site between the 25th and 26th amino acids (KRA-KD) of the pre-protein.

Proteins in the ER lumen without vacuolar targeting signals will be secreted (Chrispeels, 1991; Denecke et al., 1990; Gomord & Faye, 1996; Hadlington & Denecke, 2000). Proteins targeted to reside in the ER lumen have either ER retention motifs (Chrispeels, 1991; Gomord & Faye, 1996; Hadlington & Denecke, 2000; Neuhaus & Rogers, 1998) or associate with other ER proteins (Pagny et al., 1999). Functional plant ER retention motifs HDEL, KDEL or RDEL (Denecke et al., 1992) were not found in the sequence of RPW8.1 or RPW8.2 predicted polypeptides. Therefore unless RPW8.2 associates with other ER resident proteins or has signals targeting it to the vacuole, it will be secreted into the apoplast.

Soluble vacuolar proteins are targeted to the vacuole via the ER (Gomord & Faye, 1996; Hadlington & Denecke, 2000; Nakamura & Matsuoka, 1993; Neuhaus, 1996; Neuhaus & Rogers, 1998). Access to the vacuole, via the ER lumen, necessitates a secretory signal peptide in the pre-propeptide sequence. Three vacuolar targeting signals are known, an internal determinant, a cleavable N terminal propeptide (NTPP) or a C terminal propeptide (CTPP). No consensus for the internal determinants is known. There is no conserved motif in CTPPs, but a consensus of an acidic amino acid preceded by three hydrophobic amino acids is sufficient for CTPP vacuolar targeting. NTPPs contain the conserved NPIR motif. Neither an NPIR motif or peptide sequences consistent with the CTPP consensus were found in the RPW8 predicted polypeptides.

Cleavable N-terminal pre-sequences target globally distinct, nuclear-encoded peptides to mitochondria and plastid compartments. The mitochondrial targeting peptide (mTP) targets the preprotein for translocation to the mitochondrial matrix (Kiebler et al., 1993). The chloroplast transit peptide (cTP) is an N-terminal sequence ranging from 20-120 residues: it targets the preprotein for translocation to the stroma (Soll & Tien, 1998). Post-translocationally, mTP and cTP N-terminal presequences are cleaved from the preprotein to produce a mature protein in the mitochondrial matrix or stroma respectively.
ChloroP1.1 (Emanuelsson et al., 1999), MitoProtII1.0 (Claros, 1995; Claros & Vincens, 1996) and TargetP1.01 (Emanuelsson, 2000) are algorithms capable of predicting mitochondrial and plastid targeting. RPW8.1 and RPW8.2 predicted polypeptide sequences were predicted by TargetP1.01 and MitoProtIII1.0 not to have an mTP and by ChloroP1.1 and TargetP1.01 not to have a cTP. Therefore RPW8.1 and RPW8.2 predicted polypeptides are not expected to be translocated to mitochondria or plastids.

PSORT (Nakai & Kinehisa, 1992; Nakai & Horton, 1999) is an integrated system of algorithms that detects sorting signals in polypeptide sequences and predicts their subcellular location. PSORT is useful because it can search for sorting sequences and subcellular locations with algorithms not available elsewhere (Nielsen et al., 1999), despite not being particularly accurate (Nakai & Horton, 1999). PSORT algorithms did not predict any targeting signals listed in Table III.2 for RPW8.1 or RPW8.2 predicted polypeptides.

Table III.2: Sorting signals, with motif sequences, that PSORT algorithms did not predict to be within the RPW8 predicted polypeptides. The motif sequences are written in the standard amino acid letter code (Alberts et al., 1994), except a or X, which refer to aliphatic (a) and any (X) residue.

<table>
<thead>
<tr>
<th>Targeting signals assessed by PSORT</th>
<th>Motif sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>plasma membrane endocytosis motif</td>
<td>NPYX or YXRF</td>
</tr>
<tr>
<td>proposed Golgi consensus motif</td>
<td>(S/T)(E/Q)(R/K)</td>
</tr>
<tr>
<td>myristoylation site</td>
<td>MG at the N terminus</td>
</tr>
<tr>
<td>lipid modification motif</td>
<td>CaaX</td>
</tr>
<tr>
<td>nuclear localisation consensi</td>
<td>mono &amp; bipartite basic residue patterns</td>
</tr>
<tr>
<td>peroxisomal sorting motif</td>
<td>SKL</td>
</tr>
</tbody>
</table>

III.4: Predicted post-translational modification of the RPW8 predicted polypeptides

Post-translational modifications regulate protein activity. Protein kinases regulate cell signalling via phosphorylation at serine, threonine and tyrosine residues (Blom et al., 1999). Modification of serine and threonine residues, by O-linked glycosylation with N-acetylglucosamine (O-β-GlcNAc), may be as abundant as phosphorylation in nuclear and cytoplasmic proteins (Hart, 1997). Multimer formations of many O-β-GlcNAc modified proteins are dependent on modification with O-β-GlcNAc (Hart, 1997).
O-β-GlcNAc and phosphorylation, which may have opposing functions, compete for modification sites, thus preventing modification by the other moiety (Hart, 1997). This is known as yin-yang regulation. NetPhos2.0 (Blom et al., 1999) and YinOYang1.2 predict phosphorylation and O-β-GlcNAc acceptor sites (characteristic of the close proximity of proline and valine residues upstream and serine residues downstream) respectively, in eukaryotic proteins. NetPhos2.0 and YinOYang1.2 predictions for RPW8 predicted polypeptides are listed in Table III.3.

Table III.3: NetPhos2.0 predicted serine, threonine and tyrosine phosphorylation sites and YinOYang1.2 predicted O-β-GlcNAc sites in the RPW8 predicted polypeptides.

<table>
<thead>
<tr>
<th>Potential phosphorylation sites</th>
<th>Potential O-β-GlcNAc sites</th>
<th>Potential YinOYang sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPW8.1</td>
<td>S53</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S58</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T68</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Y75</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T101</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S130</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>S141</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S144</td>
</tr>
<tr>
<td>RPW8.2</td>
<td>S29</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S57</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S63</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T64</td>
<td>T64</td>
</tr>
<tr>
<td></td>
<td>S65</td>
<td></td>
</tr>
</tbody>
</table>
Post-translational modification can tether a soluble protein to a membrane. Glycosylphosphatidylinositol (GPI) lipid attached to an ER luminal protein anchors it to the internal face of the ER membrane (Eisenhaber et al., 1999). After the C terminal propeptide is cleaved off, GPI is attached to the mature protein at the C terminus on the ω-site – a conserved pattern of 15-40 residues (Eisenhaber et al., 1998). GPI anchored proteins are then destined to be tethered to the extracellular face of the plasma membrane (Eisenhaber et al., 1998; Eisenhaber et al., 1999). GPI sites were not predicted by big-Π predictor (Eisenhaber et al., 2000) in RPW8 predicted polypeptides.

Post-translational sulphination of tyrosine residues is a determinant of several protein-protein interactions and occurs in the Golgi; hence tyrosine sulphination modifies secreted and membrane-bound proteins in eukaryotes (Kehoe & Bertozzi, 2000). Tyrosine sulphination sites were not predicted in RPW8 predicted polypeptides by Sulfinator.

III.5: RPW8 predicted polypeptide tertiary structure predictions

Protein tertiary structure prediction methods fall into three categories: comparative modelling, threading (fold recognition) and ab initio methods (Defay & Cohen, 1995; Moult, 1999). Comparative modelling uses a structurally determined, homologous protein as a template for the protein of interest’s tertiary structure. If structurally determined homologues are not available then, threading – modelling a protein of interest’s secondary and tertiary structure using regions of structurally determined proteins that have similar amino acid characteristics – can be used as a basis. The secondary and tertiary structures of a protein of interest, which doesn’t have any structurally determined homologues or similar sequences, must be structurally predicted de novo (from polypeptide sequence alone), using ab initio methods.
BLASTP (Altschul, 1997) is an extensively used polypeptide sequence similarity algorithm, able to cope with gapped alignments. The BLOSUM-62 matrix is considered the best algorithm for general purposes and for identifying sequences with weak similarity (Altschul, 1991). Statistically significant alignments of structural homologues, identified by BLASTP, can be converted to a position-specific-score-matrix and used to re-screen the database, known as PSI-BLAST. PSI-BLAST is more sensitive to weak, but biologically relevant similarities. Hence, BLASTP, with the BLOSUM-62 matrix, was used to screen the EMBLplant (plant DNA sequences), EMBLotherest (EST), GenBank, NRL-3D (proteins of known structure), SWISS-PROT (protein) and trembl (translated EMBL) databases.

In the GenBank database, RPW8.1 and RPW8.2 predicted polypeptides were homologues of each other, as well as the predicted polypeptides of HR1, HR2, HR3 and a 200 nucleotide long *A. thaliana* cDNA clone (GenBank accession No. AF361849.1). HR1-HR3 sequences were submitted by Xiao et al. (2001) and are considered to be members of the RPW8 gene cluster. The homologous sequences (HR1-HR3, RPW8.1 and RPW8.2) were converted into a position-specific-score-matrix and PSI-BLAST iterated. Two further weakly homologous *A. thaliana* sequences, BAB08633.1 and BAB10890.1 (713 nucleotides long), were identified. BAB08633.1 was identified by Xiao et al. (2001). Regions of various myosin heavy chains from the SWISS-PROT database are homologues of the RPW8.2 predicted polypeptide, probably because myosin heavy chains contain CCs (Lupas, 1996). HR1, HR2 and HR3 predicted polypeptides were identified as RPW8 predicted polypeptide homologues in the trembl database, but no other interesting sequences were found. No homologues of RPW8.1 or RPW8.2 predicted polypeptides were identified in the NRL-3D or EMBLplant databases. In the EMBLotherest database the RPW8.2 predicted polypeptide was homologous to the 5’ mRNA sequence of an insect herbivory cDNA clone from *Medicago truncatula* (EMBL No. NF120B061N1F1057). RPW8.1 and RPW8.2 predicted polypeptides do not have any suitable structurally determined homologues or similar polypeptides in the databases for comparative modelling or threading. Therefore tertiary structures of RPW8.1 and RPW8.2 predicted polypeptides must be produced using *ab initio* methods.

ROSETTA (Simons *et al.*, 1997) is a knowledge based *ab initio* tertiary structural prediction algorithm that performed well at CASPIII (3rd Critical Assessment of Structure Prediction), the latest of the
community-wide blind prediction experimental meetings (Orengo et al., 1999; Simons et al., 1999). Ab initio structures are generated by ROSETTA for polypeptides of up to 150 residues (Baker & Sali, 2001; Bonneau et al., 2001; Simons et al., 2001) in two stages. First, local interactions of 3 and 9 residue fragments of the target sequence are compared to the fragment library of similar, structurally determined 3 and 9 residue fragments. Second, all modelled fragment combinations are assembled using a Monte Carlo annealing simulation to produce 3D structures with the least free energy, depending on hydrophobic burial, electrostatic interactions, disulphide bonds and steric α-helical and β-strand packing factors. According to Simons et al. (2001) ROSETTA can generate structures for large proteins – 120-150 residues – with accuracies of under a 7Å root mean square deviation.

Functional insights can be gained from low resolution, ab initio tertiary structure predictions (Baker & Sali, 2001; Bonneau et al., 2001). For example, Bonneau et al. (2001) structurally matched Bacteriocin AS-48 (a cyclic lysin) with 1nkl (NK-Lysin), a functionally similar protein, when the sequence identity between the two lysins was only 4%. Hence, searches of structural databases with ROSETTA, low resolution RPW8.1 and RPW8.2 predicted polypeptide tertiary structure predictions could identify proteins with similar functions, despite the lack of structurally determined homologues.

Tertiary structure predictions of the RPW8.2 (minus the SP) and RPW8.1 predicted polypeptides by ROSETTA are in Figure III.4.

III.6: Conclusion

RPW8.1 and RPW8.2 predicted polypeptides are predicted to have CC but not TM domains (Figure III.5). They are also predicted to be modified by phosphorylation or O-β-GlcNAc (Figure III.5), although not sulphonated. RPW8 predicted polypeptides were not predicted to be localised to the nucleus, mitochondrion, plastid or peroxisome; or anchored to a membrane. The predicted polypeptide of RPW8.1 was not predicted to have a secretory signal peptide, suggesting that it is cytoplasmically localised. The RPW8.2 predicted polypeptide was predicted to have a secretory signal peptide (Figure III.5) with a cleavage site between the 25th and 26th amino acids. ER retention, Golgi, or vacuolar targeting motifs were not found in the predicted polypeptide of RPW8.2, therefore implying that it is
IV.1: Introduction

Highly purified RPW8.1 or RPW8.2 protein has several potential uses. Co-immunoprecipitated with purified RPW8 proteins interacting proteins or complexes from plant extracts could be identified in vitro. More importantly, RPW8.1 and RPW8.2 tertiary structures could be determined with either X-Ray crystallography or nuclear magnetic resonance (NMR) techniques using purified RPW8 proteins.

Purifying RPW8 proteins from *A. thaliana* extracts would probably not be economical. According to S. Xiao (pers. comm.) it was difficult to detect *RPW8.1* and *RPW8.2* transcripts on Northern blots, meaning *RPW8* mRNA is not abundant in planta. This implies that RPW8.1 and RPW8.2 are poorly expressed, producing only small quantities in planta. Indeed, the high affinity, anti-RPS2 antibody was barely able to detect RPS2 on Western blots of *A. thaliana* extracts. Also, the expression of RPS2 and RPM1 in *A. thaliana* produced proteins that were only a small proportion of the total cellular protein. Leister et al. (1996) attempted to translate RPS2 mRNA in vitro, but the RPS2 protein was barley detectable on Western blots. Hence, successful extraction and purification of sufficient RPW8 protein will require using an organism based expression system.

Heterologous production of eukaryotic proteins in *E. coli* is simple, efficient, high yielding and inexpensive (Frommer & Ninnemann, 1995; Gelissen, et al., 1992; Georgiou & Valax, 1996; Giga-Hama & Kumagai, 1999; Gold, 1990; Hockney, 1994; Marston 1986). *E. coli* can produce sufficient protein for co-immunoprecipitation, X-Ray crystallography and NMR, faster and easier than using yeast, insect, mammalian or plant heterologous expression systems (QIAexpressionist). Also, Pi-ta (Jia
et al., 2000), Pto (Loh & Martin, 1995; Sessa et al., 1998; Sessa et al., 2000) have been heterologously expressed in *E. coli* and Pto as fusion protein was an active protein kinase. Pto and its signal transduction components Pti1, Pti4 and Pti5 were heterologously expressed in *E. coli* for *in vitro* assays (Gu et al., 2000; Sessa et al., 1998; Sessa et al., 2000; Zhou et al., 1995). Hence, heterologous expression of RPW8 proteins in an *E. coli* system, for extraction and purification, was chosen.

The QE, *E. coli*, heterologous expression system (QIAGEN<sup>TM</sup>) expresses a sequence, inserted into an *E. coli* expression vector, fused to a six histidine, monoclonal antibody, epitope tag (His<sub>6</sub>) when induced in *E. coli* (QIAexpressionist). The His<sub>6</sub> tagged proteins can then be extracted from *E. coli* and purified in a single, conformation independent nickel-nitrilotriacetic acid (Ni-NTA) metal affinity chromatography matrix step with mild elution conditions. The QE, *E. coli*, heterologous expression system is also capable of producing large quantities of active and stable protein. Waeber et al. (1993) used nickel chelate affinity chromatography to purify large quantities (1.3 mg) of C terminally His<sub>6</sub> tagged IIBC<sup>Glc</sup> from 5 g (wet weight) of cells. The activity or stability of IIBC<sup>Glc</sup> was not affected by a C-terminal His<sub>6</sub> tag (Waeber et al., 1993). Tagging a protein can also increase its stability in *E. coli* (Hockney, 1994). Therefore, RPW8 fusion proteins should not be affected by the His<sub>6</sub> epitope tag.

*RPW8.1* and *RPW8.2* cDNAs were cloned into the pQE-11 *E. coli* expression vector. *E. coli* was transformed with *RPW8.1-*pQE-11 or *RPW8.2-*pQE-11 and tested for the expression of RPW8 fusion proteins in time course experiments.

**IV.2: Methods**

*RPW8.1* and *RPW8.2* cDNAs were amplified from Ms-Pc cDNA (1 µl) in duplicate 20 µl PCRs. Each PCR, in thin-walled, 0.2 ml PCR tubes (Hybaid), contained 1.25U of Pfu Turbo<sup>TM</sup> (Stratagene<sup>TM</sup>) in Mg<sup>2+</sup> Tris buffer (BOEHRINGER MANNHEIM), 0.2 mM of dNTP master mix (Bioline) and 0.2 µM of each primer. MSC2 QE F (5’ – GGG CCG GGA TCC TCA TGC CGA TTG GTG AGC TTG CG –
3') and MSC2 QE R (5' – CCC CCG GTC GAC TCA AGC TCT TAT TTT ACT ACA AGC – 3') primers (MWG Biotech AG) were used to amplify *RPW8.1* cDNA. MSC3 QE F (5' – GGG CCG GGA TCC TCA TGA TTG CTG AGG TTG CCG CA – 3') and MSC3 QE R (5' – CCC CCG GTC GAC TCA AGA ATC ATC ACT GCA GAA CGT – 3') primers (MWG Biotech AG) were used to amplify *RPW8.2* cDNA. The negative controls contained no DNA; the positive controls included cosmid B6 (Fig. II.3) DNA (courtesy of Dr S. Xiao). The thermal cycle conditions (Figure IV.2) were generated in a Hybaid Omnigene Thermal Cycler.

Fig. IV.1: Pfu Turbo PCR thermal cycle conditions for amplifying RPW8.1 and RPW8.2.

Each 20 µl PCR was mixed with 4 µl of 6X orange G loading dye and electrophoresed on a 1% agarose gel at 120V. Duplicate PCR fragment bands of approximately 500 bp were excised from the gel, over a glass plate, with an EtOH sterilized scalpel (size confirmed against 500 ng of 1 kb DNA ladder marker, GIBCOBRL™). The duplicate PCR fragments were extracted from the gel slices using the QIAGEN™ QIAquick kit. Gel slices were placed in an Eppendorf tube with 3 µl of buffer OG for every µg of gel and the tube incubated at 50°C for 10 minutes (inverted every 2 minutes). The dissolved gel slice solution was then centrifuged at 14,000 xg for 60 seconds in a QIAquick spin column. Buffer PE (0.75 ml) was added to the spin columns, which were then re- centrifuged at 14,000 xg for 60 seconds. The QIAquick spin column was placed in an Eppendorf tube, 50 µl of TE was added to the middle of the QIAquick spin column and the column incubated at room temperature for 60 seconds. The TE incubated QIAquick spin column in an Eppendorf tube was then centrifuged at 14,000 xg for 60 seconds to elute the DNA in the TE into the Eppendorf tube. Elutes (50 µl) of duplicate PCR fragments were combined and a 5 µl aliquot was mixed with 1 µl of 6X orange G loading dye and electrophoresed (against 500 ng of 1kb DNA ladder marker) on a 1% agarose gel at 120V. Elute DNA concentrations were estimated at approximately 5 ng µl⁻¹ for *RPW8.1* cDNA and at approximately 1 ng µl⁻¹ for *RPW8.2* with 500 ng of λ.HindIII marker (GIBCOBRL™). *RPW8.1* and *RPW8.2* cDNAs were re-amplified using 1 µl of combined elute per PCR, repeating the procedure above, to ensure sufficient DNA for ligation. Re-amplified *RPW8.1* and *RPW8.2* cDNA concentrations...
were estimated at 16 ng µl⁻¹ and 10 ng µl⁻¹ respectively.

Total combined elutes (amplified and re-amplified) and 3 µg of pQE-11 (QIAGEN™) were independently digested by 20U of SalI (GIBCOBRL™) in REACT2 buffer (GIBCOBRL™) for 2 hours at 37°C. The digests were inactivated in a water-bath at 65°C for 15 minutes, centrifuged at 14,000 xg for 60 seconds to remove condensation from the microfuge lid, and incubated at room temperature for 30 minutes. 5M NaCl and 20U of BamHI (BOEHRINGER MANNHEIM) were added to the inactivated digests before incubating for 2 hours at 37°C. Digests were inactivated in a water-bath at 85°C for 15 minutes, centrifuged at 14,000 xg for 60 seconds and
incubated at room temperature for 30 minutes.

Digested pQE-11 was 5' de-phosphorylated, preventing re-ligation of the cleaved ends, by adding 2U of Shrimp Alkaline Phosphatase (Amersham Life Science) to the digest and incubating for 2 hours at 37°C. The Shrimp Alkaline Phosphatase was inactivated at 85°C for 15 minutes, centrifuged at 14,000 xg for 60 seconds and incubated at room temperature for 30 minutes. Aliquots of the pQE-11 digest (1 µl) and pQE-11 stock (0.2 µl) were diluted to 5 µl with sterile Milli-Q H₂O, mixed with 1 µl of 6X orange G loading dye and electrophoresed (against 500 ng of 1kb DNA ladder and 500 ng of λHindIII marker) on a 1% agarose gel at 120V. Complete digestion of pQE-11, to 50 ng µl⁻¹ of a linear, approximately 3.4kb fragment, was confirmed.

Optimal ligation substrate concentrations (i and j values) depend on the ligation volume and the size and number of DNA molecules involved. For an i = 2j reaction, in a 5 µl ligation the i and j values (Hadfield, 1986) are 134 ng µl⁻¹ of the 0.5 kb inserts (650ng of RPW8.1 or RPW8.2 cDNA) and 76 ng µl⁻¹ of the 3.4 kb vector (360ng of pQE-11).

Insert and vector were combined in the correct quantities and EtOH precipitated at -20°C for 2 hours after adding 1/10X volume of 3M NaAc and 2X volume of EtOH. The precipitated DNA was centrifuged at 14,000 xg for 15 minutes and the supernatant discarded. The DNA pellet was washed with 150 µl of 70% EtOH (v/v), re-centrifuged at 14,000 xg for 2 minutes; then the supernatant was discarded. Incubating the DNA pellet in the opened micro-fuge tube for 1 minute at room temperature dried the DNA before it was re-suspended in 3.5 µl of 10mM Tris·Cl (pH8.0) and incubated for 30 minutes at room temperature. Re-suspended DNA was mixed with 1 µl of 5X BRL ligase buffer (GIBCOBRL™) and a 0.45 µl (10%) aliquot was removed and diluted to 8 µl with sterile Milli-Q H₂O. T4 DNA ligase HC (2.25U, GIBCOBRL™) was added to the buffered, re-
suspended DNA and the ligation incubated at 12°C for 12 hours. A 0.45 µl (10%) aliquot was removed from the post-T4 DNA ligation mix and diluted to 8 µl with sterile Milli-Q H₂O. Pre- and post-T4 DNA ligase diluted aliquots were mixed with 6X orange G loading dye and electrophoresed on a 1% agarose gel at 120V to confirm ligation.

RPW8.1-pQE-11 ligated DNA was prepared for electroporation by EtOH precipitation, as in section 4.4.1. of the Cell Porator® protocol; RPW8.2-pQE-11 ligated DNA was diluted with an equal volume (5 µl) of 0.5X TE. Aliquots (20 µl) of electrocompetent, M15 E. coli cells were transformed with 1 µl of prepared, ligated DNA in a Cell Porator® (Life Technologies™). The negative control contained no DNA. Electroporated cells were then transferred to 1 ml of LB in 1.5 ml Eppendorf tubes and grown for 1 hour at 37°C in a shaker at 250 rpm. The electroporated cells were pelleted by centrifuging at 14,000 x g for 5 minutes and re-suspended in 0.2 ml of LB. The re-suspended electroporated cells were diluted 10 fold and a 0.2 ml aliquot was plated on to LB-Amp-Kan agar plates (200 µg ml⁻¹ filter sterilised ampicillin, DUCHEFA; 25 µg ml⁻¹ filter sterilised kanamycin, DUCHEFA) and grown at 37°C overnight. Colonies of RPW8.1-pQE-11 and RPW8.2-pQE-11 were picked with P2 pipette tips, inoculated into 10 ml of LB-Amp-Kan, and cultured overnight in a 37°C shaker at 250 rpm.

RPW8.1-pQE-11 cultures were put on ice for 1 hour and then 1 ml of culture was centrifuged at 5,000 xg for 5 minutes at 4°C. The supernatant was decanted, then the pellet was re-suspended in 100 µl of ice-cold Solution A and incubated on ice for 30 minutes. Solution B (200 µl) was added and the suspension was incubated for 5 minutes at room temperature. Then 150 µl of ice-cold 5M KAc/AcH was added and the suspension was incubated for 10 minutes on ice. The suspension was centrifuged for 5 minutes at 14,000 xg and the supernatant was transferred to a clean 1.5 ml Eppendorf tube. The supernatant was mixed with 400 µl of phenol:chloroform:amyl alcohol in the ratio 25:24:1 (SIGMA®), shaken, and centrifuged at 14,000 xg for 2 minutes. The upper, aqueous phase was removed, mixed
with 2X volumes of EtOH and incubated at room temperature for 2 minutes. After centrifugation at 14,000 xg for 5 minutes, the supernatant was discarded and the plasmid DNA pellet was dried by incubation in the opened Eppendorf tube for 1 minute at room temperature, before it was re-suspended in 50 µl of TE+RNase (10 mg ml⁻¹ RNase). A 5 µl portion of re-suspended plasmid (10%) was mixed with 1 µl of 6X orange G loading dye and separated on a 1.2% agarose gel at 120V against 500 ng of 1 kb DNA ladder marker.

Re-suspended plasmid DNA (1 µl) was amplified in 20 µl PCRs. Each PCR, in thin-walled 0.2 ml PCR tubes, contained 1.25U of Taq (BOEHRINGER MANNHEIM) in Mg²⁺ Tris buffer (BOEHRINGER MANNHEIM), 0.2 mM of dNTP master mix and 0.2 µM of primers, QE PF (5’–CCC GAA AAG TGC CAC CTG – 3’) and MSC2 QE R. The negative control contained no DNA and the positive control was cosmid B6 (Figure II.3). The thermal cycle conditions (Figure IV.2) were generated in a Hybaid Omnimgene Thermal Cycler. After amplification each PCR was mixed with 4 µl of 6X orange G loading dye and electrophoresed on a 1.2% agarose gel at 120V against 500 ng of 1 kb DNA ladder marker.

Fig. IV.2: Taq PCR thermal cycle conditions to test for the presence and orientation of the *RPW8.1* cDNA insert in pQE-11.

All 10 ml of *RPW8.2*-pQE-11 cell culture and the remaining 9 ml of *RPW8.1*-pQE-11 construct cell culture were centrifuged at 14,000 xg for 5 minutes at 4°C. Plasmid DNA was extracted from the cell pellet by the QIAGEN™ mini-prep technique. The supernatants were discarded and the pellets were re-suspended in 0.4 ml of buffer P1. The re-suspended pellets were transferred to a 1.5 ml Eppendorf tube, mixed gently by inverting with 0.4 ml of buffer P2 and incubated at room temperature for 5 minutes. After 5 minutes, chilled buffer P3 (0.3 ml) was added and the re-suspended pellets were incubated on ice for at least 10 minutes. The re-suspended pellets were then centrifuged at 14,000 xg for 10 minutes and the supernatant was transferred through a layer of miracloth (CALBIOCHEM®) to QIAGEN Tip20 columns, already equilibrated with 1 ml of buffer QBT. The columns were then washed four times with 1 ml of buffer QC. The DNA was eluted to a 1.5 ml Eppendorf tube from each
column with 0.8 ml of buffer QF and the DNA precipitated with 0.56 ml of room temperature isopropanol. The precipitating elutes were immediately centrifuged for 30 minutes at 14,000 xg. The supernatants were discarded and the DNA pellets washed with 150 µl of 70% EtOH (v/v) before they were centrifuged for 2 minutes at 14,000 xg and the supernatants discarded. The DNA pellets were dried in the Eppendorf tubes by incubating the tubes at room temperature with their caps open for 1 minute. Dried DNA pellets were re-suspended in 20 µl of TE and incubated at room temperature for 30 minutes. A 1 µl portion of the re-suspended DNA was diluted to 5 µl with sterile Milli-Q H_2O, mixed with 1 µl of 6X orange G loading dye and electrophoresed at 120V (against 500 ng of 1 kb DNA ladder and 500 ng of λHindIII markers) on a 1.2% agarose gel for RPW8.1 and a 1% agarose gel for RPW8.2 to estimate plasmid DNA concentration at 100 ng µl^{-1}.

The presence of RPW8.1 or RPW8.2 cDNA inserts in QIAGEN™ mini-prep plasmid DNA was checked by digesting 1 µl with 20U of SalI in a 20 µl reaction, buffered by REACT2, for 2 hours at 37°C. SalI was inactivated by incubating the digest in a water-bath at 65°C for 15 minutes, then centrifuged at 14,000 xg for 60 seconds and incubated at room temperature for 30 minutes. NaCl (5M) and 20U of BamHI were added to the inactivated RPW8.1-pQE-11 plasmid digest before incubating for 2 hours at 37°C. RPW8.2-pQE-11 plasmid DNA was EtOH precipitated for at -20°C for 2 hours after adding a 1/10X volume of 3M NaAc and 2X volumes of EtOH. Precipitated plasmid DNA was centrifuged at 14,000 xg for 15 minutes and the pellet was aspirated. The plasmid DNA pellet was washed with 150 µl of 70% EtOH (v/v), centrifuged at 14,000 xg for 2 minutes, and the supernatant was discarded. Incubating the plasmid DNA pellet in the opened micro-fuge tube for 1 minute at room temperature dried the DNA. The dried, plasmid DNA pellet was re-suspended with 20U of BamHI in REACT3 to a volume of 20 µl and incubated at 37°C for 2 hours. Digests were inactivated in a water-bath at 85°C for 15 minutes, centrifuged at 14,000 xg for 60 seconds and incubated at room temperature for 30 minutes. Digested QIAGEN™ mini-prep plasmid DNA was mixed with 4 µl of 6X orange G loading dye and half (12 µl) was electrophoresed (against 500 ng of 1 kb DNA ladder and 500 ng of λHindIII markers) on a 1.2% agarose gel at 120V.
RPW8.1- pQE-11 and RPW8.2-pQE-11 plasmid DNAs (2 µl of QIAGEN™ mini-prep DNA per reaction) were sequenced in 10 µl reactions for both primers (1.6 pmol) with 2 µl of PR mix (ABI). Primers used to sequence RPW8.1-pQE-11 plasmids were QE PF and QE TR (5’– GTT CTG AGG TCA TTA CTG G – 3’). The thermal cycle conditions are shown in Figure IV.3. Sequencing reactions were EtOH precipitated by vortexing the reactions with 8 µl of sterile Milli-Q H₂O and 32 µl of 95% EtOH (v/v) and then incubated at room temperature for at least 2 hours. The precipitated sequencing reactions are then centrifuged at 14,000 x g for 20 minutes, the supernatant was discarded and the DNA pellet washed in and vortexed with 250 µl of 70% EtOH (v/v). Washed DNA pellets were centrifuged at 14,000 xg for 10 minutes, the supernatants discarded and then the DNA pellets dried by heating the tubes, with their lids open, on a block at 90°C for 1 minute. Dried sequencing DNA was then sent to the John Innes Centre (Norwich Research Park, Norwich, England) for processing. Sequence chromatogram files in UNIX format were converted to text and compared to RPW8.1 and RPW8.2 cDNA sequences, using ted and xgap programs respectively.

Fig. IV.3: The thermal cycle conditions used to sequence RPW8.1-pQE-11 and RPW8.2- pQE-11 plasmid DNAs, with the ABI Big Dye Terminator™ system.

IV.2.1: RPW8.1 expression test

QIAGEN™ mini-prep plasmid DNA (1 µl) was diluted with 5 µl of 0.5X TE and mixed with 20 µl of electrocompetent M15 E. coli cells. The M15 E. coli, plasmid DNA mixture was electroporated as above. Glycerol stocks of M15 E. coli cells transformed with pQE-11 were used as a negative, vector control. Duplicate 20 ml Universal Containers (SARSTEDT), each containing 2 ml of LB-Amp-Kan, were inoculated with colonies picked from LB-Amp-Kan plates for both the RPW8.1-pQE-11 and the pQE-11 vector control, and incubated overnight at 37°C in a shaker at 250rpm. Pre-warmed LB-Amp-Kan (10 ml) in 20 ml Universal Containers were inoculated with 0.5 ml of the combined duplicate, overnight cultures and incubated in a 37°C shaker at 250 rpm until an optical density (OD₆₀₀) of 0.6 +/- 0.1 was achieved. The optical density of the cultures was measured in 3 ml plastic cuvettes (SARSTEDT) with a PYE UNICAM SP6-550 UV/VIS spectrophotometer. The 10 ml culture was then
split into two 20 ml Universal Containers and 0.1mM of filter-sterilised IPTG (dioxan free isopropyl–
β–D–thiogalactopyranoside, Melford Laboratories Ltd) was added to one of the cultures. At 0, 1, 3 and
6 hours after IPTG addition 1 ml from each of the four cultures was removed to plastic cuvettes and the
$\text{OD}_{600}$ was measured; then the sample was transferred to a 1.5 ml Eppendorf tube and centrifuged at
14,000 xg for 2 minutes. The supernatant was decanted and the cell pellet was stored at -20°C until the
sample could be sonicated. After storage the cell pellet was re-suspended in 100 µl of STE buffer and
sonicated, on ice, with six 10 second bursts, using a JENCONS SCIENTIFIC Ltd Ultrasonic Processor.
The sonicate was boiled with 33 µl of 4X SDS-PAGE loading dye for 5 minutes and then centrifuged
for 10 minutes at 14,000 xg. Samples (10 µl), were loaded on to 12% SDS-PAGE gels at 80V and run
at 120V against 5 µl of Low Molecular Weight (LMW) markers (Amersham).

**IV.2.2: The solubility of expressed RPW8.1**

Glycerol stocks of *RPW8.1*-pQE-11 and pQE-11 transformed M15 *E. coli* cells were streaked, with a
platinum wire loop, on to LB-Amp-Kan plates and incubated at room temperature for 2 days. Duplicate
overnight cultures were grown and treated as in IV.2.1 until after sonication. The sonicate
was centrifuged at 14,000 xg for 10 minutes at 4°C, the supernatant transferred to a 0.5 ml micro-fuge
tube and the pellet re-suspended in 100 µl of STE buffer. Supernatant and pellet fractions were mixed
with 33 µl of 4X SDS-PAGE loading dye, boiled for 5 minutes, then centrifuged at 14,000 xg for 10
minutes. Pellet (10 µl) and supernatant (25 µl) samples were loaded on to 15% SDS-PAGE gels at 80V
and run at 120V against 5 µl LMW markers.

**IV.2.3: The effects of temperature and inducer concentration on the solubility of RPW8.1**

Glycerol stocks of *RPW8.1*-pQE-11 and pQE-11 transformed M15 *E. coli* cells were streaked with a
platinum wire loop on to LB-Amp-Kan plates and incubated overnight at 37°C. Duplicate 2 ml
samples of LB-Amp-Kan in 20 ml Universal Containers (SARSTEDT) were inoculated with colonies
picked from LB-Amp-Kan plates for both *RPW8.1*-pQE-11 and the pQE-11 vector control, and incubated overnight in a shaker at 250rpm and 37°C. Five samples (10 ml) of LB-Amp-Kan in 20 ml Universal Containers (one for each of the 5 treatments, Table IV.1) were inoculated with 0.5 ml of the combined duplicate, overnight culture and incubated in a 37°C shaker at 250rpm until an OD$_{600}$ of 0.6 +/- 0.1 was achieved. At OD$_{600} = 0.6 +/- 0.1$, cultures were incubated on shakers at 250rpm, at their respective temperatures for 1 hour before IPTG induction. Two samples (of 1 ml each) were taken at 0, 1, 3 and 6 hours after addition of IPTG for all cultures; and additionally at 12, 18, 24 and 30 hours after IPTG addition for temperature tests 1 and 2. A 1 ml sample of culture was transferred to plastic cuvettes for OD$_{600}$ measurement and a 1 ml sample was transferred to a 1.5 ml Eppendorf tube and centrifuged at 14,000 x g for 2 minutes. The supernatant was decanted and the cell pellet was stored at -20°C until sonication. The cell pellet was re-suspended in 100 µl STE buffer and sonicated, on ice, with six 10 second bursts using a JENCONS SCIENTIFIC Ltd Ultrasonic Processor. The sonicate was boiled with 33 µl of 4X SDS-PAGE loading dye for 5 minutes and then centrifuged for 10 minutes at 14,000 xg. Samples (10 µl), were loaded on to 15% SDS-PAGE gels at 80V and run at 120V against 5 µl LMW markers.

Table IV.1: Temperature and inducer (IPTG) concentration conditions for each of the five separate treatment tests (-C, +C, inducer concentration test, temperature test 1 and temperature test 2) to assess the effect of temperature and inducer concentration on the solubility of RPW8.1.

<table>
<thead>
<tr>
<th></th>
<th>-C</th>
<th>+C</th>
<th>Inducer concentration test</th>
<th>Temperature test 1</th>
<th>Temperature test 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>37°C</td>
<td>37°C</td>
<td>37°C</td>
<td>22°C</td>
<td>10°C</td>
</tr>
<tr>
<td>Inducer Concentration</td>
<td>0</td>
<td>0.1mM</td>
<td>0.01mM</td>
<td>0.1mM</td>
<td>0.1mM</td>
</tr>
</tbody>
</table>

**IV.2.4: RPW8.2 expression and solubility test**

QIAGEN™ mini-prep plasmid DNA (1 µl), diluted with 5 µl of 0.5X TE mixed with 20 µl of electrocompetent M15 *E. coli* cells and electroporated as above. Glycerol stocks of pQE-11
transformed M15 *E. coli* cells were streaked with an X-Ray sterilised Bioloop® on to LB-Amp-Kan plates and incubated overnight at 37°C. Duplicate 10 ml of LB-Amp-Kan in 20 ml Universal Containers were inoculated with colonies picked from LB-Amp-Kan plates for both *RPW8.2*-pQE-11 and the pQE-11 vector control, and incubated overnight in a shaker at 250rpm and 37°C. Duplicate samples (250 ml) of LB-Amp-Kan were each inoculated with 5 ml of overnight culture and grown, on a shaker at 250 rpm and 37°C until OD$_{600}$ = 0.6 +/- 0.1 was achieved. At that point 0.1mM of filter sterilised IPTG was added to one of the duplicate 250 ml cultures. Two samples were taken at 0, 1, 3 and 6 hours after induction with IPTG for all cultures. A 1 ml sample of culture was transferred to plastic cuvettes for OD$_{600}$ measurement and a 50 ml sample was transferred to a 50 ml centrifuge tube (Nalge NUNC™ International) and centrifuged at 5,000 x g for 5 minutes at 4°C. The supernatants were discarded and the pellets stored at -20°C. The pellets were re-suspended in 20 ml of STE buffer on ice and sonicated with six 10 second bursts using a DAWE Soniprobe. A 1 ml portion of the sonicate was removed for analysis and centrifuged in a 1.5 ml Eppendorf tube at 14,000 x g for 10 minutes at 4°C; the remainder was stored at -80°C. A 200 µl portion of the supernatant was transferred to a 0.5 ml micro-fuge tube. The remaining supernatant was discarded and the pellet re-suspended in 200 µl of STE buffer. Supernatant and pellet fractions were mixed with 66 µl of 4X SDS-PAGE loading dye, boiled for 5 minutes, and centrifuged for 10 minutes at 14,000 x g. Pellet (10 µl) and supernatant (25 µl) fractions were loaded on to 15% SDS-PAGE gels at 80V and run at 120V against 5 µl of LMW markers

Sonicates (stored at -80°C) of *RPW8.2*-pQE-11 transformed M15 *E. coli* at 0 and 6 hours after induction with and without 0.1mM filter sterilised IPTG, were defrosted on ice. A 2 ml portion of the defrosted sonicate was centrifuged at 14,000 x g for 10 minutes at 4°C in a 15 ml centrifuge tube (Nalge NUNC™ International). The pellets were re-suspended in 100 µl of STE buffer and the supernatants were transferred to pre-rinsed Centricon-3® filters – 2 ml of STE buffer centrifuged
through the membranes for 2 hours at 7,500 xg and reverse centrifuged into a retentate cap for 2 minutes at 1,000 xg. Supernatants and filters were centrifuged at 7,500 xg through the membranes into the reservoirs for 6 hours at 4°C. The reservoirs were removed and retentate caps were attached to the filters before they were inverted and centrifuged at 1,000 xg for 2 minutes at 4°C to collect the retentates (concentrated supernatants). The volumes of the retentates were measured and retentate and re-suspended pellet samples were mixed with 4X SDS-PAGE loading dye, boiled for 5 minutes and centrifuged at 14,000 xg for 10 minutes. Pellet (10 µl, or 2.5 µl of RPW8.2 +IPTG at 6 hours) and retentate (25 µl, or 5 µl for RPW8.2-pQE-11 + IPTG at 6 hours) samples were loaded on to 15% SDS-PAGE gels at 80V and run at 120V against 5 µl of LMW markers. Gels were scanned, between two acetate overhead projector sheets using a ScanJet 6300C (Hewlett Packard) with the HP PrecisionScan Pro programme, by C. Taylor.

IV.3: Results

IV.3.1: RPW8.1 expression test

An E. coli culture transformed with RPW8.1-pQE-11 was induced (with 0.1 mM IPTG) to express heterologously RPW8.1 fusion proteins. Samples of total E. coli protein were extracted from portions of the culture, taken at various time points and separated on the SDS-PAGE gels in Figure IV.4. A band of approximately 19.6 kDa (the predicted size of the RPW8.1 fusion protein) was only present in induced RPW8.1-pQE-11 transformed, but not pQE-11 transformed, M15 E. coli samples. The increase in optical density (and therefore culture growth) of RPW8.1-pQE-11 transformed M15 E. coli and induced with 0.1 mM IPTG, was slower than either pQE-11 transformed M15 E. coli or non-induced RPW8.1-pQE-11 transformed M15 E. coli (Figure IV.5). Therefore RPW8.1 inhibited E. coli culture growth and was mildly toxic to E. coli.

IV.3.2: The solubility of expressed RPW8.1

RPW8.1-pQE-11 transformed M15 E. coli cells, induced to heterologously express the RPW8.1 fusion protein as in IV.3.1, were sonicated so that the soluble (supernatant) and insoluble (pellet) fractions
could be separated and tested independently for the presence of the RPW8.1 fusion protein. Figure IV.6 shows that the RPW8.1 fusion protein is only visible in the insoluble, membranous pellet fraction. The optical density data for this experiment is displayed in Figure IV.7.

**IV.3.3: The effects of temperature and inducer concentration on the solubility of RPW8.1 fusion protein**

The solubility of heterologously expressed proteins in *E. coli* can be optimised by decreasing the rate of protein synthesis (*i.e.* either by reducing the inducer concentration or lowering the growth temperature) to decrease the concentration of aggregating protein (Georgiou & Valax, 1999; Georgiou & Valax, 1996; Hockney, 1994; QIAexpressionist). Lower growth temperatures have increased the soluble yields of a range of heterologously expressed proteins in *E. coli* (King *et al.*, 1996; Oka *et al.*, 1985; Schein & Noteborn, 1988). King *et al.* (1996) suggest that partially folded intermediates are thermolabile and aggregate if the temperature is elevated, based on the formation of inclusion bodies (IBs) as growth temperatures are elevated. According to Jaenicke & Rudolph (1986) decreasing the temperature alters the folding kinetics, decreasing the driving force for self-association. Therefore, reducing the inducer concentration or lowering the *E. coli* growth temperature could solubilise RPW8 fusion proteins.

*E. coli* transformed with *RPW8.1*-pQE-11 was grown at different temperatures for one hour before (and after) induction with various concentrations of inducer. Samples of total *E. coli* protein, including the heterologously expressed RPW8.1 fusion protein, were separated on the SDS-PAGE gels in Figure IV.8. The optical density data for this experiment (Figure IV.9) shows decreased inhibition of *E. coli* cultures if a) the concentration of inducer was decreased, or b) the temperature was lower than 37°C. However, at temperatures lower than 37°C the growth of *E. coli* cultures was significantly reduced (Figure IV.9) and the RPW8.1 fusion protein was still only visible (Figure IV.8) in insoluble, membranous pellet fractions. Therefore, decreasing either the concentration of inducer or temperature did not detectably increase the solubility of RPW8.1 fusion protein in *E. coli*. 
IV.3.4: RPW8.2 expression and solubility test

*E. coli* transformed with RPW8.2-pQE-11 was induced (with 0.1 mM IPTG) to heterologously express the RPW8.2 fusion protein. Samples of total *E. coli* protein were extracted from sonicated portions of the culture, taken at various time points and, divided into pellet and supernatant fractions. SDS-PAGE gels for supernatant (soluble) and pellet (insoluble) fractions (Figure IV.10) show that the RPW8.2 fusion protein was detectable only in the insoluble membranous pellet. Optical density data (Figure IV.11) indicates that *E. coli* culture growth was strongly inhibited by heterologous expression of the RPW8.2 fusion protein and therefore RPW8.2 was toxic to *E. coli*.

RPW8.1 and RPW8.2 toxicity in *E. coli* implies that RPW8.1 and RPW8.2 fusion proteins are present in the supernatant (QIAexpress handbook, QIAGEN™), albeit at a low concentration because they were not detected on SDS-PAGE gels. If true, it is theoretically possible to detect RPW8.1 and RPW8.2 fusion proteins by concentrating the soluble, supernatant fraction. Despite 5.7 fold concentration of the supernatant fraction of the RPW8.2-pQE-11 transformed M15 *E. coli* at 6 hours post-induction sample, the RPW8.2 fusion protein was still undetectable on SDS-PAGE gels (Figure IV.12).

IV.4: Discussion

Heterologously expressed RPW8.1 and RPW8.2 fusion proteins were only detected in the insoluble, membranous pellet fraction of *E. coli* extracts, not in the soluble, supernatant fraction. However, RPW8.1 and RPW8.2 fusion protein toxicity suggests that there was some soluble protein in *E. coli*. Yet concentrating the six hour post-induction supernatant fraction 5.7 fold was not sufficient to detect the RPW8.2 fusion protein. Also, reduced inducer concentration or lowered *E. coli* growth temperature did not increase the solubility of the RPW8.1 fusion protein, probably because the RPW8.1 fusion protein was mildly toxic. Heterologous expression of RPW8.1 minus the 20 N-terminal residues and RPW8.2 minus the 22 N-terminal residues, fusion proteins in various strains of *E. coli* were also mildly toxic and toxic, respectively and insoluble (O. Kolada & S. Nettleship, pers. comm.). Therefore, RPW8.1 and RPW8.2 fusion proteins are insoluble and, to different degrees, toxic to *E. coli*; there was
probably some soluble protein, albeit at a low concentration.

The detection limit of SDS-PAGE gels stained with Coomassie Brilliant Blue (depending on the extent of destaining) is 0.1 µg protein per band (Sambrook, et al., 1989). Therefore if the amount of RPW8.1 or RPW8.2 fusion protein loaded on to the gel is undetectable even if the protein was concentrated means there was probably not enough soluble RPW8.1 or RPW8.2 fusion protein for structural determination with X-Ray crystallography or NMR, which require mg quantities (A. Leach, pers. comm.; K. Wüthrich, pers. comm.).

Insoluble RPW8.1 and RPW8.2 fusion proteins are likely to be in inclusion bodies (IBs). A number of eukaryotic proteins expressed were insoluble and aggregated into IBs when heterologously expressed in E. coli (Gellissen et al., 1992; Giga-Hama & Kumagai, 1999; Georgiou & Valax, 1996; King et al., 1996; Kohno et al., 1990; Marston, 1986; Schein & Noteborn, 1988; Stader & Silhavy, 1990); especially if they were small and complex proteins (Hockney, 1994) like RPW8.1 and RPW8.2. Proteins aggregate in vivo because partially folded intermediates accumulate at energy-demanding points in the folding pathway and self-associate through external hydrophobic surfaces (Georgiou & Valax, 1996; King, 1986; King et al., 1996). RPW8.1 and RPW8.2 were predicted to have CC motifs (Xiao et al., 2001; III.1); domains which could cause RPW8.1 or RPW8.2 fusion protein homodimerisation and aggregation in E. coli.

The formation of IBs can be beneficial. Proteins in IBs are relatively pure, yet misfolded and (from isolated IBs) can be denatured and refolded into active protein, although this process is complex and expensive (Cammarata & Schmidt, 1992; Georgiou & Valax, 1999; Georgiou & Valax, 1996; Hockney, 1994; King, 1986; Kohno et al., 1990; Marston, 1986; Paulsen et al., 1990). For example, Shi et al. (1997) recovered fully active, His$_6$ tagged protein – 10 mg per litre of culture – by solubilising the IBs, gradually renaturing the protein while it was immobilised on a nickel chelate affinity matrix and then eluting active protein. Sequestering proteins in IBs also protects them from degradation and would prevent any toxic activities of the native protein (Shi et al., 1997). Therefore, if RPW8 fusion proteins form IBs and it is possible to extract and purify RPW8.1 and RPW8.2 from E. coli IBs. However, Dr V. Bamford extracted RPW8.1 and RPW8.2 fusion proteins from E. coli under denaturing conditions but was not able to renature the proteins (V. Bamford pers. comm.).
It might be possible to successfully express RPW8.1 and RPW8.2 fusion proteins in *E. coli*, despite their toxicity, by secreting the proteins to the outer cell membrane, periplasm, or extracellular medium (Hockney, 1994; Marston, 1986; Pines & Inouye, 1999; Stader & Silhavy, 1990). Secreted proteins are less likely to be insoluble or toxic to *E. coli* (Gold, 1990; Pines & Inouye, 1999) and can easily be purified to more than 90% purity (Stader & Silhavy, 1990). However, secreted proteins make up less than 1% of the total protein in *E. coli* (Marston, 1986). Increasing the solubility of RPW8.1 and RPW8.2 fusion proteins by changing host strain, expression vector, fusing RPW8.1 or RPW8.2 to highly soluble bacterial proteins or co-expressing with chaperones were not considered because of the toxicity of RPW8.1 and RPW8.2 (Georgoiou & Valax, 1996; Georgoiou & Valax 1999; Hockney, 1994; Marston, 1986).

A prokaryotic heterologous expression system may not be suitable for the eukaryotic RPW8 proteins. Heterologously expressed RPS2, like RPW8.1 and RPW8.2 fusion proteins, was insoluble and toxic in a bacterium (M. Axtell pers. comm.). A eukaryotic expression system (e.g. yeast or plant) might provide a simpler, successful and more relevant route to soluble RPW8.1 and RPW8.2 proteins. Eukaryotic proteins produced heterologously in *E. coli* are not necessarily identical to the authentic protein produced in its natural organism because a) unnatural N terminal methionine residues are added and, b) there are no post-translational modifications, as in eukaryotes, e.g. acetylation, amidation and glycosylation (Gold, 1990; Marston, 1986). Expression of the chloroplast triose phosphate translocator was toxic to *E. coli*, forming IBs from which only non-functional protein could be extracted, but it was successfully expressed in yeast (Loddenkötter *et al*., 1993). Bacteriorhodopsin was functionally expressed in the membranes of *Schizosaccharomyces pombe* by Hildebrandt *et al.* (1989), whereas *E. coli* expression was low and required isolation and reconstitution into lipid vesicles (Dunn, *et al*, 1987; Karnik *et al*., 1987). Hence, the heterologous expression of RPW8.1 and RPW8.2 fusion proteins in yeast or expressed in planta are possible systems for producing sufficient protein for structural determination.
Expression of RPW8.1 and RPW8.2 in *Schizosaccharomyces pombe*

V.1: Introduction

Heterologous expression of RPW8.1 and RPW8.2 proteins in the eukaryote, *S. pombe*, was intended as a parallel option to expressing RPW8.1 and RPW8.2 in the prokaryote, *E. coli*. A eukaryotic expression system might provide a simpler, more successful route to soluble, authentic RPW8.1 and RPW8.2 proteins (IV.4). Emr (1990) considers heterologous expression in yeast to be the next best option after *E. coli* for proteins incompatible with bacterial expression. Since RPW8.1 and RPW8.2 fusion proteins were insoluble and toxic in *E. coli* (IV.3) it was important to assess the ability of *S. pombe* to express RPW8.1 and RPW8.2 fusion proteins. Considering RPW8.1 and RPW8.2 had been heterologously expressed in *S. cerevisiae* (S. Xiao pers. comm.) expression in yeast appeared to be the most suitable system for producing RPW8.1 and RPW8.2 proteins for structure and biochemical tests.

Yeasts are amenable to molecular biological research and have advantages over *E. coli* as an heterologous expression system. Yeasts, compared to higher eukaryotes, are simple, unicellular microorganisms that can grow rapidly to high cell densities on simple media when carrying plasmids and can be manipulated as easily as *E. coli* (Egel *et al.*, 1980; Emr, 1990; Gellissen *et al.*, 1992; Giga-Hama & Kumagai, 1999; Goeddel, 1990; Mitchison, 1970). Also, yeasts share elements of cellular processes with higher eukaryotes, such as cell compartmentation, intracellular transport, vesicular trafficking, cell
cycle control, signal transduction, chromatin structure and RNA splicing and post-translational modifications, e.g. folding, glycosylation, phosphorylation and acetylation (Emr, 1990; Frommer & Ninnemann, 1995; Gellissen et al., 1992; Goeddel, 1990; Mitchison, 1970; Moreno et al., 1991; Tohda et al., 1994).

The cell biology, biochemistry, classical and molecular genetics of *S. pombe* is the most intensively studied of the yeasts and techniques for its manipulation are used extensively (Brake, 1990; Giga-Hama & Kumagai, 1999; Moreno et al., 1991). *S. pombe*, the fission yeast, is genetically and physiologically well characterised (Tohda et al., 1994) and is considered to be a model system (Moreno et al., 1991). Despite the advantages of *S. pombe* as an organism for heterologous expression, only a few functional proteins have been expressed heterologously in *S. pombe* (Tohda et al., 1994), including examples of proteins from *A. thaliana* and *Chlorella* (Li et al., 1998; Loddenkötter et al., 1993; Sauer et al., 1990; Sauer et al., 1990b).

The amount of protein produced by heterologous expression in yeast and *E. coli* is comparable; *S. cerevisiae* has also expressed large amounts of soluble or secreted proteins (Emr, 1990). Eukaryotic proteins, like human lipocortin I or chloroplast triose phosphate translocator, have been successfully expressed in yeast but not in *E. coli*. Human lipocortin I (which was poorly expressed in *E. coli* and lacked the post-translational modifications of the native protein) expressed in *S. pombe* produced 150 mg of purified, functional protein from 10 g of wet cell paste (Giga-Hama et al., 1994). Functional and tagged chloroplast triose phosphate translocator, which could not be expressed in *E. coli*, was expressed and purified by Loddenkötter et al. (1993) from the ER of *S. pombe*. Therefore yeasts are widely used as an heterologous expression system (Brake, 1990), especially for plant proteins (Frommer & Ninnemann, 1995).

The ESP heterologous, *S. pombe* expression system (Stratagene™) uses a high copy number extrachromosomal vector to express proteins fused at the N terminus to FLAG and glutathione-S-transferase (GST) monoclonal antibody tags. Heterologous expression is controlled by the highly expressed *nmt1* (no message in thiamine 1) promoter, whose activity is very effectively inhibited by thiamine (Vitamin B₁) repression (Forsburg, 1993; Maundrell, 1990; Maundrell, 1993; Zhao et al., 1998). Fusion
proteins, heterologously expressed in *S. pombe* can be purified to as much as 15 µg protein ml⁻¹ of culture with a single-step, conformation independent purification using GST affinity resin columns (Smith & Johnson, 1988). Various higher eukaryote fusion proteins have been expressed and purified from *S. pombe* (ESP™ Instruction Manual) including functional STP1 (Sauer *et al.*, 1990).

*RPW8.1* and *RPW8.2* cDNAs were cloned into the ESP expression vector, pESP-2. *S. pombe* was transformed with *RPW8.1*-pESP-2 and *RPW8.2*-pESP-2 and then the expression of RPW8.1 and RPW8.2 fusion proteins, respectively, was de-repressed. At various time points portions of the cultures were taken and analysed for soluble and insoluble RPW8.1 and RPW8.2 fusion protein.

**V.2: Methods**

Ten-fold diluted Ms-Pc cDNA (1 µl) was amplified in duplicate 20 µl PCRs with 1.25U of Pfu Turbo™ (Stratagene™) in Mg²⁺ Tris buffer (BOEHINGER MANNHEIM), 0.2 mM of dNTP master mix (Bioline) and 0.2 µM of each primer, in thin-walled, 0.2 ml PCR tubes (Hybaid). MSC2 ESP F (5’ – GGG CGG ATG CAT ATG CCG ATT GGT GAG CTT GCG – 3’) and MSC2 ESP R (5’ – CCC CCG ATG CAT TCA AGC TCT TAT TTT ACT ACA AGC – 3’) primers (MWG Biotech AG) were used to amplify *RPW8.1* cDNA. MSC3 ESP F (5’ – GGG CCG ATG CAT ATG ATT GCT GAG GTT GCC GCA – 3’) and MSC3 ESP R (5’ – CCC CCG ATG CAT TCA AGA ATC ATC ACT GCA GAA CGT – 3’) primers were used to amplify *RPW8.2* cDNA. Negative controls contained no DNA; the positive controls were cosmid B6 (Figure II.3). The thermal cycle conditions (Figure V.1) were generated in a Hybaid Omnigene Thermal Cycler.

Fig. V.1: Pfu Turbo PCR thermal cycle conditions used to amplify *RPW8.1* and *RPW8.2* cDNAs.

Each 20 µl PCR was mixed with 4 µl of 6X orange G loading dye and electrophoresed on a 1% or 1.2% (*RPW8.1* and *RPW8.2*, respectively) agarose gel at 120V. Duplicate PCR fragment bands of approximately 500 bp were excised from the gel, over a glass plate, with an EtOH sterilized scalpel (size confirmed against 5 µl of 1 kb DNA ladder marker, GIBCOBRL™). The duplicate PCR fragments were extracted from the gel slices using the QIAGEN™ QIAquick kit as in IV.2. Elutes
(50 µl) of duplicate PCR fragments were combined and a 5 µl aliquot was mixed with 1 µl of 6X orange G loading dye and electrophoresed (against 5 µl of 1 kb DNA ladder marker) on a 1% agarose gel at 120V; the DNA concentration of the RPW8.1 and RPW8.2 cDNA fragments were estimated as approximately 5 ng µl⁻¹ of and 1 ng µl⁻¹, respectively, with 500 ng of λHindIII marker (GIBCOBRL™). This procedure was repeated to ensure there was sufficient DNA for ligation. RPW8.1 and RPW8.2 cDNA fragments were re-amplified using 1 µl per PCR of combined elute. Re-amplified RPW8.1 and RPW8.2 cDNA fragment concentrations (45 or 35 cycles of PCR, respectively) were estimated on a 1% agarose gel as 16 ng µl⁻¹ and 2 ng µl⁻¹, respectively.

Total combined elutes (amplified and re-amplified) and 30 µl of pESP-2 (3 µg, Stratagene™) were independently digested by 20U of NsiI (Stratagene™) in Universal buffer (Stratagene™) for 4 hours at 37°C. The digests were inactivated in a waterbath at 65°C for 15 minutes, centrifuged at 14,000 rpm for 60 seconds and incubated at room temperature for 30 minutes.

Digested pESP-2 was 5’ de-phosphorylated by adding 2U of Shrimp Alkaline Phosphatase (Amersham Life Science) to the digest and incubating for 2 hours at 37°C. Shrimp Alkaline Phosphatase was inactivated at 85°C for 15 minutes, centrifuged at 14,000rpm for 60 seconds and incubated at room temperature for 30 minutes. Aliquots of the pESP-2 digest (1 µl) and pESP-2 stock (1 µl) were diluted to 5 µl with sterile Milli-Q H₂O, mixed with 1 µl of 6X orange G loading dye and electrophoresed (against 500 ng of 1 kb DNA ladder and 500 ng of λHindIII marker) on a 1% agarose gel² at 120V. Vector pESP-2 had been completely digested to produce 80 ng µl⁻¹ of a linear 9.7 kb fragment.

Optimal ligation substrate concentrations (i & j values) are dependent on the ligation volume and the size and number of DNA molecules involved. For a 5 µl ligation the i and j values (for an i = 2j reaction) are approximately 134 ng µl⁻¹ of insert and 62 ng µl⁻¹ of vector (Hadfield, 1986). This is
equivalent to 670 ng of *RPW8.1* or *RPW8.2* cDNA and 250 ng of pESP-2.

Insert and vector were combined in the correct quantities and EtOH precipitated at 

-20°C for 2 hours after adding 1/10X volumes of 3M NaAc and 2X volumes of EtOH. The precipitated DNA was centrifuged at 14,000rpm for 15 minutes and the supernatant discarded. The DNA pellet was washed with 150 µl of 70% EtOH (v/v), re-centrifuged at 14,000rpm for 2 minutes, and the supernatant was discarded. Incubating the DNA pellet in the opened micro-fuge tube for 1 minute at room temperature dried the DNA before it was re-suspended in 3.5 µl of 10mM Tris·Cl (pH8.0) and incubated for 30 minutes at room temperature. Re-suspended DNA was mixed with 1 µl of 5X BRL ligase buffer (GIBCOBRL™) and a 0.45 µl (10%) aliquot was removed and diluted to 8 µl with sterile Milli-Q H₂O. T4 DNA ligase HC (2.25U, GIBCOBRL™) was then added to the buffered, re-suspended DNA and the ligation was incubated at 12°C for 12 hours. A 0.45 µl (10%) aliquot was removed from the post-T4 DNA ligase solution and diluted to 8 µl with sterile Milli-Q H₂O. Pre and post-T4 DNA ligase diluted aliquots were mixed with 6X orange G loading dye and electrophoresed on a 1% agarose gel at 120V to confirm ligation.

*RPW8.1*-pESP-2 and *RPW8.2*-pESP-2 ligated DNA was prepared for electroporation by EtOH precipitation and was re-suspended in 10 µl of 0.5X TE, as in section 4.4.1 of the Cell Porator® protocol. Aliquots (20 µl) of electrocompetent, DH10B, *E. coli* cells were transformed with 1 µl of prepared, ligated DNA in a Cell Porator® (Life Technologies™). The negative control contained no DNA. Electroporated cells were then transferred to 1 ml of LB in 1.5 ml Eppendorf tubes and grown for 1 hour in a 37°C shaker at 250 rpm. The electrophoresed cells were pelleted by centrifuging at 14,000 rpm for 5 minutes and re-suspended in 200 µl of LB. Aliquots of the re-suspended electroporated cells, 100 µl and 200 µl of a 10 fold dilution, were plated on to LB-Amp-Kan agar plates (200 µg ml⁻¹ filter sterilised ampicillin, DUCHEFA; 25 µg ml⁻¹ filter sterilised kanamycin,
DUCHEFA) and grown at 37°C overnight.

Colonies of RPW8.1-pESP-2 and RPW8.2-pESP-2 were picked with P2 pipette tips, inoculated into 10 ml of LB-Amp and cultured overnight in a 37°C shaker at 250 rpm. Cultures were put on ice for 1 hour and then 1 ml of each culture was centrifuged at 5,000 x g for 5 minutes at 4°C. The supernatant was decanted and the pellet was re-suspended in 100 µl of ice cold Solution A. The pellet was then incubated on ice for 30 minutes. After incubation, solution B (200 µl) was added and the suspension incubated for 5 minutes at room temperature. Then 150 µl of ice-cold 5M KAc/AcH was added and the suspension was incubated for 10 minutes on ice. The suspension was centrifuged for 5 minutes at 14,000rpm and the supernatant was transferred to a clean 1.5 ml Eppendorf tube. The supernatant was mixed with 400 µl of 25:24:1 phenol:chloroform:isoamyl alcohol (SIGMA®), shaken and centrifuged at 14,000rpm for 2 minutes. The upper, aqueous phase was removed, mixed with 2X volumes of EtOH and incubated at room temperature for 2 minutes. After centrifugation at 14,000rpm for 5 minutes, the supernatant was discarded and the plasmid DNA pellet dried by incubating the opened Eppendorf tube for 1 minute at room temperature, before it was re-suspended in 50 µl of TE+RNAse (10 mg ml⁻¹ RNAse). A 5 µl portion of re-suspended plasmid DNA (10%) was mixed with 1 µl of 6X orange G loading dye and separated on a 1.2% agarose gel at 120V against 5 µl of 1 kb DNA ladder marker to ensure that the plasmid DNA had been extracted.

Re-suspended plasmid DNA (1 µl) was PCR amplified in 20 µl reactions, with 1.25U Taq (BOEHRINGER MANNHEIM) in Mg²⁺ Tris buffer, 0.2mM of dNTP master mix and 0.2µM of primers. Primer ESPY F (5’– GTA CTT GAA ATC CAG CAA GTA TAT AGC – 3’) was used with either MSC2 ESP R or MSC3 ESP R to amplify RPW8.1 and RPW8.2 cDNAs, respectively, in thin-walled, 0.2 ml PCR tubes. The negative control contained no DNA and the positive control was cosmid B6 (Figure II.3). The thermal cycle conditions (Figure V.2) were generated in a Hybaid Omnimgene Thermal Cycler. Each PCR was mixed with 4 µl of 6X orange G loading dye and electrophoresed on 1.2% agarose gel at 120V against 500 ng of 1 kb DNA ladder marker. Plasmids which produced a PCR fragment of approximately 500 bps were considered to contain RPW8.1 or
RPW8.2 cDNAs, ligated into the pESP-2 vector in the correct orientation.

Fig. V.2: Taq PCR thermal cycle conditions used to amplify RPW8.1 and RPW8.2 cDNA to test for insert presence and orientation in the pESP-2 expression vector.

All of the remaining \textit{RPW8.1}-pESP-2 and \textit{RPW8.2}-pESP-2 cell cultures (9 ml) were centrifuged at 14,000 rpm for 5 minutes at 4\textdegree C. Plasmid DNA was extracted from the cell pellet by the QIAGEN\textsuperscript{TM} mini-prep technique. The supernatants were discarded and the pellets were re-suspended in 0.4 ml of buffer P1. The re-suspended pellets were transferred to a 1.5 ml Eppendorf tube, mixed gently by inverting with 0.4 ml of buffer P2 and incubated at room temperature for 5 minutes. After 5 minutes chilled buffer P3 (0.3 ml) was added and the re-suspended pellets were incubated on ice for at least 10 minutes. The re-suspended pellets were then centrifuged at 14,000 \textit{xg} for 10 minutes and the supernatant was transferred through a layer of miracloth (CALBIOCHEM\textsuperscript{®}) to QIAGEN Tip20 columns, already equilibrated with 1 ml of buffer QBT. The columns were washed four times with 1 ml of buffer QC. The DNA was then eluted to a 1.5 ml Eppendorf tube from each column with 0.8 ml of buffer QF and the DNA precipitated with 0.56 ml of room temperature isopropanol. The precipitating elutes were immediately centrifuged for 30 minutes at 14,000 \textit{xg}. The supernatants were discarded and the DNA pellets washed with 150 \mu l of 70% EtOH (v/v) before they were centrifuged for 2 minutes at 14,000 \textit{xg} and the supernatants discarded. The DNA pellets were dried in the Eppendorf tubes by incubating the tubes at room temperature with their caps open for 1 minute. Dried DNA pellets were re-suspended in 20 \mu l of TE and incubated at room temperature for 30 minutes. A portion (1 \mu l) of the re-suspended DNA was diluted to 5 \mu l with sterile Milli-Q H\textsubscript{2}O, mixed with 1 \mu l of 6X orange G loading dye and electrophoresed (against 500 ng of 1 kb DNA ladder and 500 ng of \textit{λ}-\textit{HindIII} markers) on a 1.2% agarose gel at 120V to estimate the plasmid DNA concentration, which was approximately 100 ng \mu l\textsuperscript{-1}.

Insert presence in QIAGEN\textsuperscript{TM} mini-prep plasmid DNA was checked by digesting 1 \mu l with 20U of \textit{NsiI} in a 20 \mu l reaction, buffered by Universal buffer, overnight at 37\textdegree C. Digests were inactivated by
incubating in a water-bath at 65°C for 15 minutes, then centrifuged at 14,000 rpm for 60 seconds and incubated at room temperature for 30 minutes. Inactivated digests were mixed with 4 µl of 6X orange G loading dye and half (12 µl) was electrophoresed (against 500 ng of 1 kb DNA ladder and 500 ng of λ-HindIII markers) on a 1.2% agarose gel at 120 V. Digestion with NsiI released a fragment of approximately 500 bp, corresponding to the size of RPW8.1 and RPW8.2 cDNAs, from RPW8.1-pESP-2 and RPW8.2-pESP-2 plasmids of approximately 10 kb.

QIAGEN™ mini-prep plasmid, RPW8.1-pESP-2 and RPW8.2-pESP-2 DNAs (2 µl per reaction) were sequenced in 10 µl reactions for both (1.6 pmol) primers with 2 µl of PR mix (ABI). The primers used were ESPY F and ESPY R (5’ – GTT TTA GCA TTA TAC GTC GAA CTT ACC CGA AGG – 3’). The thermal cycle conditions are shown in Figure V.3. Sequencing reactions were EtOH precipitated by vortexing the reactions with 8 µl of sterile Milli-Q H₂O and 32 µl of 95% EtOH (v/v) and then incubated at room temperature for at least 2 hours. The precipitated sequencing reactions are then centrifuged at 14,000 xg for 20 minutes, the supernatant was discarded and the DNA pellet washed in and vortexed with 250 µl of 70% EtOH (v/v). Washed DNA pellets were centrifuged at 14,000 xg for 10 minutes, the supernatants discarded and then the DNA pellets dried by heating the tubes, with their lids open, on a block at 90°C for 1 minute. Dried sequencing DNA was then sent to the John Innes Centre (Norwich Research Park, Norwich, England) or the Bioanalytical Laboratory (School of Biological Sciences, UEA) for processing. Sequence chromatogram files in UNIX format were converted to text and compared to RPW8.1 and RPW8.2 cDNA sequences, using ted and xgap programs.

Fig.V.3: The thermal cycle conditions used to sequence RPW8.1-pESP-2 and RPW8.2-pESP-2 with the ABI Big Dye Terminator™ system.

A glycerol stock of S. pombe was streaked, with a platinum wire loop, on YES agar plates and incubated for 3 days at 30°C. Two 2-3 mm colonies were inoculated to duplicate 50 ml centrifuge tubes (Nalge NUNC™) with 5 ml of YES and sterile foam bungs. The YES cultures were then grown
overnight on a 30°C shaker at 250 rpm. A 0.5 ml portion from each overnight culture was inoculated into separate 60 ml of YES and grown on a 30°C shaker at 250rpm until OD$_{600}$ = 0.5+/-.1. Both 60 ml cultures were transferred to two 50 ml centrifuge tubes and centrifuged at 1,000 xg for 5 minutes at room temperature. The supernatants were discarded and the pellets washed with Milli-Q H$_2$O.

Washed pellets were re-centrifuged at 1,000 xg for 5 minutes at room temperature and the supernatant discarded. All the pellets were re-suspended in a total of 600 µl of TE/LiAc. Aliquots (100 µl) of re-suspended S. pombe cells were vortexed with 1 µl of construct DNA (pESP-2, RPW8.1-pESP-2 or RPW8.2-pESP-2) and 0.1 mg of Herring testes carrier DNA (GIBCOBRL$^{\text{TM}}$). The negative control was no DNA. PEG/LiAc (0.6 ml) was added to each transformation and vortexed before incubating on a 30°C shaker at 200rpm for 30 minutes. Transformations were inverted with 70 µl of dimethyl sulfoxide (DMSO, SIGMA$^{\text{®}}$) and incubate in a water-bath for 15 minutes at 42°C, then put on ice for 2 minutes. Transformations were centrifuged at 14,000rpm for 60 seconds, the supernatant discarded, and the pellet re-suspended in 100 µl TE and plated on to EMM-agar (Mitchinson, 1970; Moreno et al., 1991) with 5µM thiamine plates and incubated for 7 days at 30°C.

A colony from each transformation plate (pESP-2, RPW8.1-pESP-2 or RPW8.2-pESP-2) was picked with a sterile P2 tip into 5 ml of YES in a 50ml centrifuge tube with a sterile foam bung and cultured overnight on a 30°C shaker at 250rpm. Each culture was used to inoculate 10 ml of YES in a 50 ml centrifuge tube with a sterile foam bung to OD$_{600}$ = 0.2-0.4 and grown on a 30°C shaker at 250rpm until OD$_{600}$ = 0.7-1.0. Cultures were centrifuged at 12,000 xg for 10 minutes at room temperature, the supernatant discarded, and the pellet was washed with 20 ml of Milli-Q H$_2$O. Washed pellets were re-centrifuged at 12,000 xg for 10 minutes at room temperature, the supernatant discarded, and the pellets were re-suspended in 20 ml of EMM. The re-suspended pellets were split equally into two 50 ml centrifuge tubes with sterile foam bungs (10 ml in each). Thiamine – at a final concentration of 5µM – was added to one of the centrifuge tubes. All the tubes were then incubated on a 30°C shaker at 250
rpm. Two 1 ml samples were taken at 0, 6, 12, 18 and 24 hour time points for a) OD$_{600}$ measurement and b) protein analysis. Protein analysis samples were put into 1.5 ml Eppendorf tubes and centrifuged at 14,000rpm for 10 minutes at room temperature, the supernatant was discarded and the cell pellets were frozen at -80°C. When all the samples had been collected the cell pellets were thawed on ice.

Approximately 0.25 g of acid-washed glass beads (425-600 microns, SIGMA®) and 100 µl of ice-cold PBST, with 2µM Benzamidine and 5µM DTT, were added to the thawed cell pellets and the cells were lysed by vortexing for 5 minutes at 4°C. The lysate was centrifuged at 12,000 x g for 5 minutes at 4°C and approximately 100 µl of lysate supernatant was transferred to a 0.5 ml Eppendorf tube. Ice-cold PBST (100µl), with 2µM Benzamidine (SIGMA®) and 5µM dithiothreitol (SIGMA®), was added to lysate pellets. Lysate pellet and supernatant fractions were mixed with 33 µl of 4X SDS-PAGE loading dye and boiled for 5 minutes. Boiled samples were centrifuged at 14,000 rpm for 10 minutes and then 25 µl was loaded (at 80V) on to 12% SDS-PAGE gels and run at 120V against 5 µl of low molecular weight (LMW) markers (Amersham). Proteins were visualised on the SDS-PAGE gels with Brilliant Blue. Gels were photographed by S. Davies.

**V.3: Results**

*S. pombe* cultures transformed with *RPW8.1*-pESP-2 or *RPW8.2*-pESP-2 were de-repressed (by washing thiamine from the growth medium) in order to express RPW8.1::GST and RPW8.2::GST fusion proteins. Portions of the *S. pombe* cultures were taken at various time points, lysed and separated into pellet (insoluble) and supernatant (soluble) fractions. RPW8.1::GST and RPW8.2::GST fusion proteins were only detected in the de-repressed lysate pellet fractions of *S. pombe*, whereas GST was present in the de-repressed lysate pellet and supernatant fractions (Figure V.4). Therefore, heterologously expressed RPW8.1::GST and RPW8.2::GST fusion proteins were insoluble in *S. pombe*.

The growth of de-repressed, but not repressed, *S. pombe* cultures carrying *RPW8.2*-pESP-2 was inhibited after 6 hours post de-repression (Figure V.5), although RPW8.2::GST was not detected until
24 hours post de-repression (data not shown). There was no difference between the growth of repressed and de-repressed *S. pombe* cultures carrying RPW8.1-pESP-2 (Figure V.6) and RPW8.1::GST was detectable at 12 hours (data not shown). Transcription from the *nmt1* promoter was only detectable 10 hours after de-repression and was not maximal until 16 hours after de-repression (Maundrell, 1990; Tommasino & Maundrell, 1991) and according to Maundrell (1993) the induction kinetics of the *nmt1* promoter are true for many coding sequences. Also, the *nmt1* promoter is not completely repressed in EMM + thiamine; Forsburg (1993) detected low levels of β-galactosidase activity produced from the *nmt1* promoter in a repressed culture of *S. pombe*. Therefore RPW8.2::GST was probably toxic to *S. pombe* at concentrations below the 0.1 µg detection limit of the SDS-PAGE gels (Sambrook *et al*., 1989) and before activity of the *nmt1* promoter is detectable.

**V.4: Discussion**

Heterologously expressed RPW8.1::GST and RPW8.2::GST were both insoluble and RPW8.2::GST was probably toxic to *S. pombe*. Despite this fully active protein could be obtained from yeast by secretion (Smith *et al*., 1985). Cytoplasmically expressed prochymosin was insoluble and inactive, although prochymosin fused to the yeast invertase secretion signal was secreted into the media and was fully active. Secreted proteins in yeast are also compartmentalised and removed from the bulk of yeast proteins preventing potential toxicities (Moir & Davidow, 1991). However, less than a tenth of prochymosin produced by yeast was secreted, although two mutants *ssc1* and *ssc2* did increase the amount secreted by approximately 10 fold (Smith *et al*., 1985). Using *S. cerevisiae* secretion mutants to improve protein yield efficiency is a common strategy (Chisholm *et al*., 1990; Moir & Davidow, 1991) and can generate yields of 10 mg to 50 mg per litre of culture (Moir & Davidow, 1991). According to Gellissen *et al.* (1992) it is possible to secrete proteins as efficiently in other yeasts as in *S. cerevisiae*. Hence secreting heterologously expressed RPW8.1 and RPW8.2 from yeast would be less efficient, but might produce soluble and fully active protein, whilst also avoiding any associated toxicity.
RPW8.1 and RPW8.2 proteins could be produced from plant expression systems. The advantage of expressing of RPW8.1 and RPW8.2 \textit{in planta} is that the tagged RPW8 fusion proteins can be functionally tested to confirm correct folding and viability. Cf9 and RPM1 c-myc epitope fusion proteins have been extracted from plants (Boyee \textit{et al}., 1998; Piedras \textit{et al}., 2000). The c-myc epitope tag is small, meaning that the size of RPW8-c-myc fusion proteins would be amenable to NMR as well as X-Ray crystallography, if sufficient protein was extractable. Hence, RPW8.1 and RPW8.2 c-myc fusion proteins might be obtainable in sufficient quantities for plant expression systems for NMR and X-Ray crystallography as an alternative to \textit{E. coli} or \textit{S. pombe} heterologous expression systems.

Allelic polymorphism of \textit{RPW8.1} and \textit{RPW8.2} from accessions of \textit{Arabidopsis thaliana}
VI.1: Introduction

Very little is known about the structure-function relationships of RPW8.1 and RPW8.2. To assess the effect of structural features on the function of RPW8.1 and RPW8.2 it will be necessary to mutate these features and functionally test the mutant proteins. As RPW8.1 and RPW8.2 are unique proteins (III), structural inferences cannot be obtained from homologous polypeptides or protein structures. To fully assess the functional effect of RPW8.1 and RPW8.2 structural features would require an extensive mutagenic study, without \textit{a priori} assumptions. However, candidate structural features, likely to be essential for \textit{RPW8} function, can be identified by analysing \textit{RPW8.1} and \textit{RPW8.2} allele polymorphism. Therefore analysing \textit{RPW8.1} and \textit{RPW8.2} allele polymorphism would be less time-consuming and labour intensive in comparison with an extensive mutagenic study and should indicate which structural features should be assessed with mutagenesis.

Structure-function relationships can be identified from multiple sequence alignments (Livingstone & Barton, 1996). Caicedo \textit{et al}. (1999) sequenced RPS2 alleles from 17 resistant and susceptible ecotypes to search for structural features associated with resistance. Similarly, structural features that correspond with \textit{RPW8} function could be identified by comparing the predicted polypeptides of \textit{RPW8.1} and \textit{RPW8.2} alleles from resistant and susceptible accessions of \textit{A. thaliana}. Also, the analysis of \textit{RPW8.1} and \textit{RPW8.2} allele polymorphism, between resistant and susceptible accessions, can indicate evolutionary mechanisms and patterns, such as regions of conservation and diversity.

Chasman & Adams (2001) developed a model to predict the effect of amino acid polymorphisms (AAPs) on protein structure and function. The model was tested on unbiased mutation data for lysozyme and the \textit{lac} repressor. The AAPs which most significantly effected protein structure and function were either a “buried charge” or “unusual” (\textit{i.e.} not present in the phylogenetic profile) amino acid changes, particularly “unusual amino acid by class”. The significant effect of “buried charge” amino acid changes is consistent with the data of Sunyaev \textit{et al}. (2000) who showed that human genetic diseases were caused by AAPs, mostly at sites with low solvent accessibility.

The presence of hybridising restriction fragments on a Southern blot indicates the presence of a
homologous sequence in the DNA sample (Southern, 1975). For instance, homologues of *Pto* were detected in *L. hirsutum* on a Southern blot probed with *Pto*; one sequence (*LhirPto*) was cloned and had *Pto*-specific function (Riely & Martin, 2001). Southern blotting also found that *RPP13, RPM1* and *RPS2* were single copy sequences (Bittner-Eddy et al., 2000; Grant et al., 1995; Wroblewski et al., 2000) and that *Cf-9, Cre3, HRT/RPP8, M, N, P2, Pilb, Pto, Rp1, RPP1, RPP5, RPS2, Sw-5, Xa1* and *Xa21* loci were complex and composed of multiple homologous sequences (Anderson et al., 1997; Botella et al., 1998; Brommonschenkel et al., 2000; Collins et al., 1999; Cooley et al., 2000; Dodds et al., 2001; Jones et al., 1994; Lagudah et al., 1997; Martin et al., 1993; McDowell et al., 1998; Mindrinos et al., 1994; Parker et al., 1997; Rommens et al., 1995; Song et al., 1994; Wang et al., 1999; Whitham et al., 1994; Yoshimura et al., 1998).

The size of a hybridising restriction fragment is equal in length to the interval between restriction enzyme sites. Loss or gain of restriction enzyme sites during evolution results in longer or shorter restriction fragments, respectively. Thus, there is much similarity in the restriction fragment length patterns of two DNA sequences that are closely related (Nei & Kumar, 2000). Therefore, restriction fragment length polymorphisms (RFLPs) are indicators (albeit crude compared to sequencing) of the degree of polymorphism at nuclear DNA loci.

*RPW8.1* and *RPW8.2* alleles from different ecotypes were sequenced. Twenty eight accessions were selected (Table VI.1) because their resistance phenotypes to infection by *E. cichoracearum* isolate UCSC1 and *E. cruciferarum* isolate UEA1 were either highly resistant to both or highly susceptible to both (Adam et al., 1999). These extremes of resistance phenotype were chosen because the scoring is more reliable and it is more likely that the effect of the *RPWS* locus will be discernable from experimental variation. A Southern blot confirmed the presence or absence of *RPW8.1* or *RPW8.2* allele sequences in the genomic DNA of the accessions. The sequences of *RPW8.1* and *RPW8.2* alleles from each accession were analysed for evolutionary patterns, which might indicate functional constraints. The predicted polypeptides of *RPW8.1* and *RPW8.2* alleles were then analysed for structural features that correlated with the resistance phenotype of their accession.
Table VI.1: The sampling location, country of origin and seed stock numbers of the selected *A. thaliana* accessions. All the seed stocks used were single seed lines except where indicated by BULK. NASC = Nottingham *Arabidopsis* Stock Centre, ABRC = *Arabidopsis* Biological Resource Centre.

<table>
<thead>
<tr>
<th>Accession</th>
<th>Location/Country</th>
<th>NASC Stock No.</th>
<th>ABRC Stock No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bs-1</td>
<td>Basel, Switzerland</td>
<td>N996</td>
<td>N6627</td>
</tr>
<tr>
<td>Bu-8</td>
<td>Burghaun, Rhön, Germany</td>
<td>N1021</td>
<td>N6639</td>
</tr>
<tr>
<td>C24</td>
<td>N906 [BULK]</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Cit-0</td>
<td>Citou, Aude, France</td>
<td>N1081</td>
<td>-</td>
</tr>
<tr>
<td>Co-1</td>
<td>Coimbra, Portugal</td>
<td>N1085</td>
<td>N6669</td>
</tr>
<tr>
<td>Di-1</td>
<td>Dijon, France</td>
<td>N1109</td>
<td>N6681</td>
</tr>
<tr>
<td>Dra-2</td>
<td>Drahonin, Czechoslovakia</td>
<td>N1121</td>
<td>N6687</td>
</tr>
<tr>
<td>Ei-5</td>
<td>Eifel, Germany</td>
<td>-</td>
<td>N6691</td>
</tr>
<tr>
<td>Fl-1</td>
<td>/Finland</td>
<td>N1161</td>
<td>N6706</td>
</tr>
<tr>
<td>Fr-6</td>
<td>Frankfurt, Ginnheim, Germany</td>
<td>N1177</td>
<td>N6712</td>
</tr>
<tr>
<td>Gd-1</td>
<td>Gudon, Germany</td>
<td>N1185</td>
<td>N6716</td>
</tr>
<tr>
<td>Gü-1</td>
<td>Glückingen, Germany</td>
<td>N1215</td>
<td>N6731</td>
</tr>
<tr>
<td>Gy-0</td>
<td>La Miniere, France</td>
<td>N1217</td>
<td>N6732</td>
</tr>
<tr>
<td>Je-0</td>
<td>Jena, Germany</td>
<td>N1247</td>
<td>N6742</td>
</tr>
<tr>
<td>Kas-1</td>
<td>Kashmir, India</td>
<td>N1265</td>
<td>N6751</td>
</tr>
<tr>
<td>Kl-0</td>
<td>Koeln, Germany</td>
<td>N1275</td>
<td>N6756</td>
</tr>
<tr>
<td>Ksk-1</td>
<td>Keswick, England</td>
<td>N1678</td>
<td>-</td>
</tr>
<tr>
<td>Nw-0</td>
<td>Neuweihau, Germany</td>
<td>N1409</td>
<td>N6811</td>
</tr>
<tr>
<td>Nok-0</td>
<td>Noordwijk, Netherlands</td>
<td>N1399</td>
<td>N6807</td>
</tr>
<tr>
<td>Petergof</td>
<td>Petergof, Russia</td>
<td>N926 [BULK]</td>
<td>-</td>
</tr>
<tr>
<td>Pla-2</td>
<td>Playa de Aro, Spain</td>
<td>N1463</td>
<td>N6916</td>
</tr>
<tr>
<td>Sah-0</td>
<td>Sierra Alhambra, Spain</td>
<td>N1501</td>
<td>N6917</td>
</tr>
<tr>
<td>Shahdara</td>
<td>Pamiro-Alay, Tajikistan</td>
<td>-</td>
<td>N6180</td>
</tr>
<tr>
<td>Sorbo</td>
<td>Khurmatou, Tajikistan</td>
<td>N931 [BULK]</td>
<td>-</td>
</tr>
<tr>
<td>Ts-7</td>
<td>Tossa de Mar, Spain</td>
<td>N1563</td>
<td>N6873</td>
</tr>
<tr>
<td>Uk-1</td>
<td>Umkirch, Germany</td>
<td>N1575</td>
<td>N6879</td>
</tr>
<tr>
<td>Wa-1</td>
<td>Warshau, Poland</td>
<td>N1587</td>
<td>N6885</td>
</tr>
<tr>
<td>Wt-2</td>
<td>Wietze, Germany</td>
<td>-</td>
<td>N6893</td>
</tr>
</tbody>
</table>
VI.2: Methods

VI.2.1: Sequencing genomic *RPW8* alleles

Approximately 10 seeds from each *A. thaliana* accession were sown on sterile soil in 3” plastic pots (WARD) for each of the 28 accessions and watered to field capacity with tap water. The pots were kept in a tray and covered with clingfilm to maintain the humidity and cold treated at 4°C for 60 hours. The tray was transferred to a long-day growth room (16h light, 22+/−1°C, 50% humidity) for 47 days. The clingfilm was punctured at 72 hours and the plants were watered every two days.

Two or three leaves or siliques were cut from *A. thaliana* plants of each accession with EtOH (BDH) cleaned scissors, washed in Milli-Q dH₂O and snap frozen in liquid N₂, in a 1.5 ml Eppendorf tube. Snap frozen tissue was crushed, on ice, with a water and EtOH cleaned micropestle (Eppendorf); then 700 µl of extraction buffer was added to the crushed tissue and the samples were ground to a homogenate. Next, the samples were incubated at 65°C for 10 minutes and centrifuged at 13,000 rpm for 7 minutes at room temperature. The supernatant (700 µl) was transferred to a 1.5 ml Eppendorf tube, inverted with 0.7 ml of 1:1 phenol:chloroform (v/v, SIGMA®) and centrifuged at 13,000 rpm for 8 minutes at room temperature. The upper, aqueous phase (600 µl) was removed to a 1.5 ml Eppendorf tube, mixed with an equal volume of chloroform (SIGMA®) and centrifuged at 12,000 rpm for 5 minutes at room temperature. The upper, aqueous phase was removed and added to 30 µl of 3M NaAc
(pH 5.5, made with DEPC H$_2$O) and 0.6X volumes of isopropanol (BDH) in a 1.5 ml Eppendorf tube. The contents of the Eppendorf tubes were mixed and centrifuged at 13,000 rpm for 30 minutes at room temperature. The supernatant was discarded and the DNA pellet was washed with 150 µl of 75% EtOH (v/v) and centrifuged at 14,000 rpm for 2 minutes at room temperature. The supernatant was discarded and the DNA pellet was dried by incubating it in the Eppendorf tube, with the cap open, for 1 minute at room temperature. The dried pellet was then re-suspended with 50 µl of TE by vortexing and incubating at room temperature for 30 minutes. An aliquot (5 µl) of the re-suspended DNA was mixed with 1µl of 6X orange G loading dye and electrophoresed on a 1% agarose gel at 120V. The concentration of re-suspended *A. thaliana* accession genomic DNA was estimated against 500 ng of 1 kb DNA ladder marker (GIBCOBRL$^{\text{TM}}$) and 500 ng of λ*HindIII* marker (GIBCOBRL$^{\text{TM}}$).

Genomic DNA (1 µl) for each accession was amplified in triplicate 20 µl HOT-START PCRs with 1.25U of BIOTAQ (Bioline), in Mg$^{2+}$ Tris buffer (BOEHRINGER MANNHEIM), 0.2mM of dNTP set (Bioline) and 0.2µM of each primer (MWG Biotech AG), in thin-walled, 0.2 ml PCR tubes (Hybaid). MC2 F (5′ – ATG CCG ATT GGT GAG CTT GCG ATA – 3′) and MC2 R (5′ – TCA AGC TCT TAT TTT ACT ACA AGC – 3′) primers were used to amplify *RPW8.1*. MC3 F (5′ – ATG ATT GCT GAG GTT GCC GCA – 3′) and MC3 R (5′ – TCA AGA ATC ATC ACT GCA GAA CGT – 3′) primers were used to amplify *RPW8.2*. Negative controls contained no DNA. The PCR positive control was cosmid B6 (Figure II.3) and the DNA positive controls were genomic DNA from accessions Ms-0 (*RPW8.1* and *RPW8.2*) and Col-gl (*RPW8.2* only). The thermal cycle conditions (Figure VI.1) were generated in a Hybaid Omnigene Thermal Cycler, and filter pipette tips (TipOne$^{\text{TM}}$) were used to prevent contamination. Each 20 µl PCR was mixed with 4 µl of 6X orange G loading dye and electrophoresed on a 1% agarose gel against 500 ng of 1 kb DNA ladder marker at 100V. Fig. VI.1: PCR thermal cycle conditions used to amplify *RPW8.1* and *RPW8.2* from *A. thaliana* ecotype genomic DNA (* denotes when BIOTAQ was added to the PCR tubes).

For each accession, triplicate PCR fragment bands of approximately 500 bps were excised from the gel (on a UV light box over a glass plate) with an EtOH sterilized scalpel. The triplicate PCR fragments
were pooled and extracted from the gel slices in 1.5 ml Eppendorf tubes using the QIAGEN™ QIAEXII kit. For every 1 mg of gel, 3 µl of QX1 was added to the Eppendorf tube with 10 µl of re-suspended QIAEXII beads, vortexed for 30 seconds and incubated for 15 minutes at 50°C (inverting every 2 minutes). The Eppendorf tubes were centrifuged for 1 minute at 14,000 rpm and the supernatant was discarded. The pellet was then re-suspended in 0.5 ml of QX1 and vortexed for 30 seconds. The Eppendorf tubes were centrifuged for 1 minute at 14,000 rpm, the supernatant was discarded, then the pellet was re-suspended in 0.5 ml of PE and vortexed for 30 seconds. The previous centrifugation and re-suspension step was repeated once. Then the Eppendorf tubes were centrifuged for 1 minute at 14,000 rpm, the supernatant was discarded and the QIAEXII bead pellet was then dried by incubating the tubes at room temperature with the caps open until the pellet had just turned white.

PCR fragment DNA was eluted from the dried QIAEXII bead pellet with 20 µl of 10mM Tris•Cl (pH8.0), vortexed for 30 seconds, then incubated for 10 minutes at 50°C. The elute and beads were vortexed for 30 seconds and centrifuged for 14,000 rpm for 1 minute before removing the elute supernatant to a clean 0.5 ml Eppendorf tube. A 1 µl aliquot of the re-suspended PCR fragment DNA (20 µl) was diluted to 5 µl, then mixed with 6X orange G loading dye and electrophoresed on a 1% agarose gel against 500 ng of 1 kb DNA ladder marker and 500 ng of λHindIII marker at 100V. The DNA concentration, on average, was estimated as 200 ng µl⁻¹.

PCR fragments (0.5 µl DNA per reaction) were sequenced in 10 µl reactions for both primer pairs (1.6 pmol) with 2 µl of PR mix (ABI Big Dye™). The primers used to sequence the PCR fragments were MC2 Seq F and K3LF (5’ – TAG GCG ATC CAA AAA CCA ATA GAA – 3’) for RPW8.1, and MC3 Seq F and MC3 Seq R (5’ – GGGAGTTTATGGAAATCACAT – 3’) for RPW8.2. The thermal cycle conditions are shown in Figure VI.2. Sequencing reactions were EtOH precipitated by vortexing the reactions with 8 µl of sterile Milli-Q H₂O and 32 µl of 95% EtOH (v/v) and then incubated at room temperature for at least 2 hours. The precipitated sequencing reactions are then centrifuged at 14,000 xg for 20 minutes, the supernatant was discarded and the DNA pellet washed in and vortexed with 250 µl of 70% EtOH (v/v). Washed DNA pellets were centrifuged at 14,000 xg for 10 minutes, the
supernatants discarded and then the DNA pellets dried by heating the tubes, with their lids open, on a block at 90°C for 1 minute. Dried sequencing DNA was then sent to the Bioanalytical Laboratory (Dept. of Biological Sciences, UEA). Sequence chromatogram files in UNIX format were compared to RPW8.1 and RPW8.2 genomic sequences from accession Ms-0 (provided by S. Xiao) using ted and xgap programs. Alignments of allele genomic sequences were constructed using CLUSTAL W (Thompson et al., 1994) from FASTA format genomic sequences. Polypeptide sequences for each allele were predicted by comparing the genomic sequence of each allele with the cDNA sequences of RPW8.1 and RPW8.2, aligned by CLUSTAL W (Thompson, et al., 1994) and pretty printed by BOXSHADE3.2.

Fig. VI.2: The thermal cycle conditions used to sequence RPW8.1 and RPW8.2 alleles with the ABI Big Dye Terminator™ system.

VI.2.2: Southern Blot

Seeds of A. thaliana accessions (Figure VI.5) were cold treated at 4°C for 28 days. Cold treated seeds were sterilised by washing them in 100% EtOH for 2 minutes, twice in 15% sodium hypochlorite (Fischer Scientific International) for 15 minutes, and 3 times in Milli-Q H₂O. Sterilised seeds from each accession were plated on to MS-Agar in 50 mm (Sterelin) petri dishes and grown in the greenhouse under short day conditions at 22°C. Seedlings were transferred to double autoclaved soil in a 6½” x 3¾” plastic tray (WARD) per ecotype, watered to field capacity, covered with clingfilm for 3 days, and put in the greenhouse.

Once the plants had flowered, approximately 1 g of aerial tissue from each accession was harvested to a chilled pestle and mortar, snap frozen in liquid N₂ and ground to a fine powder. Powdered plant tissue was spooned into a 50 ml centrifuge tube (Nalge NUNC™ Int.) and vortexed with 15 ml of extraction buffer at 65°C, 2 ml of 10% SDS (w/v) and 17 µl of β-mercaptoethanol. The extracts were then incubated at 65°C for 15 minutes (inverting every 2 minutes), inverted with 5 ml of 5M KAc and
left on ice for at least 30 minutes before the chilled extracts were centrifuged at 10,000 rpm for 30 minutes at 4°C. The supernatants were transferred to 50 ml centrifuge tubes through a layer of miracloth (CALBIOCHEM®) and inverted with 0.7 volumes of isopropanol and left for 30 minutes to precipitate the nucleic acids. Precipitated nucleic acids were pelleted by centrifuging at 10,000 rpm for 30 minutes at 4°C, the supernatants decanted, and the pellets dried by standing the inverted tubes on a paper towel. Nucleic acid pellets were re-dissolved in 0.7 ml of TE and 1 µl of QIAGEN RNAse, then incubated at 37°C for 20 minutes. Re-dissolved DNA was transferred to a 1.5 ml Eppendorf tube, inverted with 0.7 ml of 1:1 phenol:chloroform (v/v), and centrifuged at 14,000 rpm for 10 minutes. The upper, aqueous layer was transferred to a 1.5 ml Eppendorf tube, inverted with 0.7 ml of 1:1 phenol:chloroform (v/v), and centrifuged at 14,000 rpm for 10 minutes. The upper, aqueous layer was removed again to a 1.5 ml Eppendorf tube and centrifuged at 14,000 rpm for 2 minutes to pellet any solid matter. The supernatant was transferred to a 1.5 ml Eppendorf tube and the DNA was precipitated with a 1/20X volume of NaAc and 0.7X volume of isopropanol at -20°C overnight. Precipitated DNA was pelleted by centrifuging at 14,000 rpm for 30 minutes; the supernatant was discarded. The pellet was washed with 0.5 ml of 70% EtOH (v/v) and then re-pelleted by centrifuging at 14,000 rpm for 2 minutes. The supernatant was discarded and then the pellet was air dried at room temperature for 15 minutes. Dried DNA was re-suspended in 200 µl of TE and incubated at room temperature for 30 minutes at 4°C overnight. The DNA (2 µl) concentration was electrophoresed on a 1% TAE agarose gel at 100V against 100 ng, 200 ng and 500 ng of λHindIII marker to estimate the DNA concentration.

Two micrograms of DNA extract for each accession was mixed with REACT2 buffer and incubated for 2 hours at 4°C. Buffered DNA was incubated for 6 hours at 37°C with 100U of HindIII (ROCHE). TE was added to the digested DNA to a final volume of 0.7 ml, then inverted with 0.7 ml of 1:1 phenol:chloroform (v/v), and centrifuged at 14,000 rpm for 10 minutes. The upper, aqueous layer was transferred to a 1.5 ml Eppendorf tube and inverted with 0.7 ml of 1:1 phenol:chloroform (v/v), then centrifuged at 14,000 rpm for 10 minutes. The upper, aqueous layer was removed again to a 1.5 ml Eppendorf tube and precipitated with a 1/20X volume of NaAc and a 0.7 volume of isopropanol at
Precipitated digested DNA was pelleted by centrifuging at 14,000 rpm for 30 minutes, the supernatant was discarded and the pellet was washed with 0.5 ml of 70% EtOH (v/v). The digested DNA was re-pelleted by centrifuging at 14,000 rpm for 2 minutes, the supernatant discarded, the pellet re-centrifuged at 14,000 rpm for 60 seconds and the remaining excess fluid removed. Pellets of digested DNA were air dried by incubating at room temperature with the Eppendorf tube cap open for 60 seconds, re-suspended in 40 µl of TE and incubated for 30 minutes at room temperature. A 5 µl (12.5%) aliquot of the digested DNA was electrophoresed on a 1% TAE gel at 100V to assess the concentration and digestion of the DNA.

The remaining 35 µl of fully digested DNAs were mixed with 7 µl of LD and electrophoresed on a 20.6 cm long, 0.8% TAE gel at 80V until the blue dye front was 1 cm from the end of the gel. The gel was rinsed in RO.H2O, followed by a wash in 900 ml of denaturing solution on a shaker for 30 minutes. The denatured gel was rinsed in RO.H2O and washed in 900 ml of neutralising solution on a shaker for 30 minutes. The neutralised gel was rinsed in RO.H2O and then washed with 900 ml of 10X SSC on a shaker for 30 minutes. The edges of the gel (not containing any DNA) were cut away and the gel was placed, inverted, on two sheets of 3MM chromatography paper (Whatman®), taking care not to get any air pockets under the gel. The gel on the sheets of chromatography paper was supported on a glass plate above a tank of 10X SSC. Clingfilm was placed across the edges of the gel and the wells.

Hybond N+ membranes (Amersham) were then cut to size, the orientation was marked and they were placed on to the inverted gel. Any bubbles between the gel and the membranes were removed with a glass roller. Three layers of blotting paper, pre-soaked in 10X SSC, were laid on top of the membranes and any more bubbles removed. Then a box of Lotus professional tissues was placed on top of the blotting paper-membrane-gel stack with 0.5 kg on a tray set on top. The blot was left for 48 hours to allow complete DNA transfer. Transferred blots were autocrosslinked by 12 µJ joules in a UVStratalinker® 2400 (Stratagene®) and then dried for 2 hours at 80°C.

Blots were pre-wet with 6X SSC and blocked in 5 ml of pre-hybridisation solution for 3 days at 65°C.
RPW8.1 and RPW8.2 genomic DNAs were heated at 95°C for two minutes, then put on ice for 5 minutes. The denatured RPW8 genomic DNAs (25 ng of each) were then labelled with the Amersham Multiprime DNA labelling kit in a 50 µl reaction containing 10 µl of buffer, 5 µl of primer solution, 2 µl of dCTP (approximately 3,000 Ci/mmol) and 2 µl of Klenow fragment. The Multiprime DNA labelling reaction was then incubated at 37°C for 30 minutes, denatured at 95°C for 5 minutes and then added to the pre-hybridisation solution and hybridised overnight. The hybridisation solution was removed and the blots were washed twice in 2X SSC with 0.1% SDS (w/v) at 65°C for 15 minutes. Then the blots were washed twice in 0.5X SSC with 0.1% SDS (w/v) at 65°C for 15 minutes, contact dried on chromatography paper, and then wrapped in clingfilm. The clingfilm-wrapped blots were put into a Kodak phosphorimager with a Kodak phosphorimaging screen to detect the radioactivity, for 3 days. The phosphorimaged blots were developed by scanning the phosphorimaging screen in a Molecular Dynamics STORM 840.

VI.3: Results & Discussion

VI.3.1: RPW8 allele sequences

Figures VI.3 & VI.4 show that sequences of RPW8.1 and RPW8.2 alleles are highly conserved within the accessions. Variation between sequences of RPW8.1 and RPW8.2 alleles is due to insertions, deletions and single nucleotide polymorphisms – 24 and 31 in RPW8.1 and RPW8.2 alleles, respectively.
VI.3.1.1: Insertions and deletions

Indels in \textit{RPW8.1} and \textit{RPW8.2} alleles (Table VI.2) have significant effects on the sequences of \textit{RPW8.1} and \textit{RPW8.2} predicted polypeptides. Indels in \textit{RPW8.2} alleles from accessions Nok-0B, Sah-0B, Petergof, Sy-0 and Ws-0 lead to truncated \textit{RPW8.2} predicted polypeptides (Figure VI.8). Two of the indels in \textit{RPW8.1} alleles (10 and 63 bp) involve a direct repeat. The 63 bp direct repeat occurs in six of the \textit{RPW8.1} alleles (from accessions Can-0, Fl-1, Fr-6, Nd-0, Ts-7 and Uk-1) and causes the predicted polypeptide to be extended at the C terminus (VI.7).

Direct repeats can occur by transposon excision or intragenic recombination. Transposons, which were found in the \textit{RPW8} locus by Xiao \textit{et al.} (2002), produce a 4-12bp direct repeat, known as a footprint, when they excise (I.4.3.). This could explain the 10 bp repeat in the \textit{RPW8.1} allele of accession C24. The 63 bp direct repeat is too long to be a transposon footprint and was probably formed by a single intragenic recombination event between two \textit{RPW8.1} orthologues.

VI.3.1.2: Single nucleotide polymorphisms

Nucleotide substitutions in \textit{RPW8.1} and \textit{RPW8.2} alleles demonstrate that these genes are evolving rapidly. Polymorphic nucleotide sites were 4.19\% and 5.36\% of the \textit{RPW8.1} and \textit{RPW8.2} allele sequences, respectively, from 24 accessions (Xiao \textit{et al.}, 2002), including the allele sequences analysed here. This intraspecific polymorphism between \textit{RPW8} alleles is much greater than the divergence (1.26\%) between \textit{RPS2} allele sequences from 17 ecotypes (Caicedo \textit{et al.}, 1999). Ka/Ks ratios (I.4.5) of pairwise comparisons between \textit{RPW8.1} allele sequences or between \textit{RPW8.2} allele sequences indicated that the coding sequences of \textit{RPW8.1} and \textit{RPW8.2} are conserved (Xiao \textit{et al.}, 2002), although this analysis is still in progress (S. Xiao pers. comm.).

A pair of sequences that had diverged early would be expected to have accumulated more mutations per base pair in the introns than the exons of their allele sequences, unless the introns are functionally conserved. There are more mutations per base pair in \textit{RPW8.2} allele sequences than in \textit{RPW8.1} allele sequences (Table VI.3) even though \textit{RPW8.2} diverged later than \textit{RPW8.1} (Xiao \textit{et al.}, 2002). This
implies that greater functional constraints have been imposed on RPW8.1 than RPW8.2. Surprisingly, the introns of RPW8.2 allele sequences have accumulated significantly fewer mutations per base pair than either the exons of RPW8.2 alleles or the introns of RPW8.1 alleles (Table VI.3). Therefore, for unknown reasons, the RPW8.2 intron seems to be functionally conserved.

Table VI.3: The number of mutations (Muts) per base pair (to three significant figures) in intron and exon regions of RPW8.1 and RPW8.2 alleles, excluding indels.

<table>
<thead>
<tr>
<th></th>
<th>Intron</th>
<th>Exons</th>
<th>Genomic sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base s</td>
<td>Mut s</td>
<td>Muts bp-1</td>
</tr>
<tr>
<td><strong>RPW8.1</strong></td>
<td>197</td>
<td>7</td>
<td>0.0355</td>
</tr>
<tr>
<td><strong>RPW8.2</strong></td>
<td>128</td>
<td>1</td>
<td>0.0078</td>
</tr>
</tbody>
</table>

**VI.3.1.3: Duplication and deletion of RPW8.1 and RPW8.2 alleles**

The RPW8.2 allele external primer amplified two fragments of different size from the genomic DNA of ecotypes Co-1, Je-0, Ksk-1, Nok-0 and Sah-0. The two differently sized fragments (the smaller labelled A and the larger B) were extracted and sequenced separately. Both Co-1A and Co-1B external primer sequences matched the coding primer sequence, indicating that in accession Co-1 RPW8.2 had duplicated. Also, RPW8.2 alleles from accessions Fr-6, Nok-0A and Pla-2 could not be sequenced between base pairs 555-565, due to sequencing difficulties consistent with competing, homologous sequences in the sequencing reactions. Multiple RPW8.2 sequences could be present in the sequencing reaction for two reasons: either it was contaminated or duplicate RPW8.2 sequences, with PCR fragments of the same length, were present in these accessions. Therefore, RPW8.2 is duplicated in ecotype Co-1 and might be in accessions Fr-6, Nok-0A and Pla-2.

RPW8.1 and RPW8.2 allele sequences could not be amplified, despite repeated attempts, from accessions Bs-1, Bu-8, Gd-1 and Gü-1. RPW8.2, but not RPW8.1, allele sequences could be amplified from accessions Di-1, Dra-2 and Gy-0. RPW8.1 and RPW8.2 hybridising sequence was not detected in the genomic DNAs of accessions Gd-1, Gü-1 or Col-0 (Figure VI.5). The RPW8 locus from accession Col-0 has been sequenced and does not contain RPW8.1 or RPW8.2 sequences (Xiao et al., 2002).
Therefore \textit{RPW8.1} and \textit{RPW8.2} alleles are absent from accessions Gd-1 and Gü-1, and probably absent from accessions Bs-1 and Bu-8. \textit{RPW8.1} and \textit{RPW8.2} hybridising sequence was detected in the genomic DNAs of accessions Dra-2 and Gy-0 (Figure VI.5), although neither accession has an approximately 1.8 kb band that was present in the genomic DNAs of other accessions. Therefore, the 1.8 kb band corresponds to a restriction fragment that contains \textit{RPW8.1}, which is absent in the genomic DNA of accessions Dra-2, Gy-0 and probably Di-1. Therefore \textit{RPW8.1} has been deleted from the genomes of accessions Di-1, Dra-2 and Gy-0 and, \textit{RPW8.1} and \textit{RPW8.2} have been deleted from the genomes of Bs-1, Bu-8, Gd-1 and Gü-1 accessions.

\textbf{VI.3.1.4: \textit{RPW8} homologues}

\textit{RPW8.1} \& \textit{RPW8.2} hybridising sequences in the genomes of C24, Co-1, Dra-2, Gy-0, Kas-1, Kl-0, Ksk-1, Nw-0, Petergof, Shahdara, Ts-7, Wa-1 and Wt-2 accessions (Figure VI.5) shows that the \textit{RPW8} locus in these accessions is polymorphic, complex and comprised of multiple homologous sequences. Homologues of \textit{RPW8.1} or \textit{RPW8.2} were not detected in the genomes of accessions Gd-1 or Gü-1, indicating that the \textit{RPW8.1} and \textit{RPW8.2} are either absent or have been deleted from the genome of these accessions. The restriction fragment length patterns of \textit{RPW8.1} \& \textit{RPW8.2} hybridising sequences in the genomic DNA from accessions C24 and Co-1 were very similar, implying that C24 might have duplicate copies of \textit{RPW8.2} too.

According to the topology of the phylogenetic tree in Figure VI.6, \textit{RPW8.1} and \textit{RPW8.2} alleles from accessions Kas-1, Ms-0, Nw-0, Shahdara and Wa-1 are closely related. These accessions are phenotypically resistant to powdery mildew (Adam \textit{et al.}, 1999) and closely related, whereas susceptible alleles are divergent (Figure VI.6). Alleles of \textit{RPS2} from phenotypically susceptible \textit{A. thaliana} accessions are also more divergent than alleles from phenotypically resistant accessions (Caicedo \textit{et al.}, 1999). The RFLP banding pattern on the Southern blot (Figure VI.5) implies that the \textit{RPW8} locus in these accessions has subsequently diverged into three groups: i) Kas-1 and Ms-0, ii) Nw-0 and Wa-1 and iii) Shahdara.

\textbf{VI.3.2: \textit{RPW8} predicted polypeptides}
The similarity and identity between the predicted polypeptides of \textit{RPW8.1} and \textit{RPW8.2} alleles is high (Figures VI.7 & VI.8); CLUSTAL W (Thompson et al., 1994) calculated that the identities were all greater than 90%. Figures VI.7 and VI.8 were pretty printed with BOXSHADE3.2. BOXSHADE3.2 labels identical amino acids with white text on a black background. Non-identical amino acids are labelled with: a) black text on a white background if there is no similarity in physico-chemical properties, and b) white text on a grey background if there is similarity. Amino acid groups considered by BOXSHADE3.2 to be physico-chemically similar are listed in Table VI.4

<table>
<thead>
<tr>
<th>Physico-chemical similarity group</th>
<th>Amino acids</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aromatic</td>
<td>F, Y &amp; W</td>
</tr>
<tr>
<td>Non-Polar Aliphatic</td>
<td>I, L, V &amp; M</td>
</tr>
<tr>
<td>Basic</td>
<td>R &amp; K</td>
</tr>
<tr>
<td>Acidic</td>
<td>D &amp; E</td>
</tr>
<tr>
<td>Non-Polar</td>
<td>G &amp; A</td>
</tr>
<tr>
<td>Uncharged Polar</td>
<td>T &amp; S</td>
</tr>
<tr>
<td>Uncharged Polar Amide</td>
<td>N &amp; Q</td>
</tr>
</tbody>
</table>

Most of the variation between accessions in the predicted polypeptides of \textit{RPW8.1} or \textit{RPW8.2} alleles are single AAPs with non-identical physico-chemical properties (Figures VI.7 & VI.8). However, the predicted polypeptides of \textit{RPW8.2} alleles from five accessions (Nok-0B, Sah-0B, Petergof, Sy-0 and Ws-0) are truncated compared to the Ms-0 allele sequence. The RPW8.1 predicted polypeptides from six accessions (Uk-1, Fr-6, Can-0, Nd-0, Fl-1 and Ts7) have an inserted direct repeat. Accession Kl-0 has a RPW8.2 predicted polypeptide variant not seen for any of the other \textit{RPW8.2} alleles, but its sequence is incomplete because the predicted stop codon could not be found.
In the predicted polypeptide structures from susceptible ecotypes there are no clear motif or AAP candidate structural features that could be responsible for susceptibility. Fl-1, Fr-6, Ler, Nd-0, and Ws-0 accessions were phenotypically susceptible (Adam et al., 1999). The only structural motif that was found exclusively in the predicted polypeptides from alleles of phenotypically susceptible accessions was the RK--- motif at the C terminus of the RPW8.1 predicted polypeptide from accession Fr-6. This was not found in the predicted polypeptides of other RPW8.1 alleles from susceptible accessions.

Assuming that all the predicted polypeptides in Figures VI.7 & VI.8 are expressed in *A. thaliana*, a susceptible accession must have non-functional RPW8.1 and RPW8.2 proteins, because both RPW8.1 and RPW8.2 genes are functional (Xiao et al., 2001). If a single motif or AAP were to be responsible for abolishing RPW8.1 or RPW8.2 function, then those motifs or AAPs would be common to all the predicted polypeptides of RPW8.1 or RPW8.2 alleles from phenotypically susceptible accessions. The AAPs T64S and D116G were found in all predicted polypeptides of RPW8.2 alleles from susceptible ecotypes and S40I was found in all predicted polypeptides of RPW8.1 alleles from susceptible ecotypes. Therefore T64S and D116G in RPW8.2 and S40I in RPW8.1 are good candidate structural features that could be responsible for susceptibility. However, the S40I AAP was found in the RPW8.1 predicted polypeptide of accession Can-0, the T64S and D116G AAPs were found in the RPW8.2 predicted polypeptide of accession Can-0 and the D116G AAP was found in the RPW8.2 predicted polypeptide of accession Gy-0. Accessions Can-0 and Gy-0 were both phenotypically resistant (Adam et al., 1999) and accession Gy-0 has no RPW8.1 allele (VI.3.1.4), meaning that the RPW8.2 allele might be the only functional RPW8 gene in this accession.

Assumptions in the analysis above mean that S40I, T64S and D116G AAPs are still candidate structural features that could be responsible for susceptibility, despite being
present in the predicted polypeptide sequence of the phenotypically resistant accessions Can-0 and Gy-0. Multiple motifs or AAPs, perhaps including S40I, T64S and D116G, may be responsible for abolishing \textit{RPW8.1} or \textit{RPW8.2} function. \textit{RPW8.1} and \textit{RPW8.2} alleles might not be expressed in some \textit{A. thaliana} accessions, meaning that any predicted polypeptide sequences from these accessions would be misleading. Factors or genes other than \textit{RPW8} are probably responsible for the resistant phenotypes of some accessions. Preliminary genetic analysis by Xiao \textit{et al.} (2002) indicates that resistance to \textit{E. cichoracerum} isolate UCSC1 in accession Can-0 is not linked to the \textit{RPW8} locus. \textit{RPW1}, \textit{RPW2}, \textit{RPW10}, \textit{RPW11}, \textit{RPW12}, \textit{RPW13} and \textit{RPW14} loci confer different degrees of resistance to \textit{E. cichoracerum} isolate UCSC1 in accessions Kas-1 and Wa-1 (Adam & Somerville, 1996; Schiff \textit{et al.}, 2001; Wilson \textit{et al.}, 2001). Even variation in the chemical composition of epicuticular wax between \textit{A. thaliana} accessions (Rashotte \textit{et al.}, 1997) could significantly affect an accession’s resistance phenotype. Also, Caicedo \textit{et al.} (1999) found that the sequences of \textit{RPS2} alleles from \textit{A. thaliana} accessions UIE132 and Tsu-0 were identical although accession Tsu-0 was resistant and UIE132 was partially resistant. As the \textit{RPW8} locus probably does not contribute to resistance in accession Can-0, then substitution of T64S and D116G \textit{RPW8.2} and of S40I in \textit{RPW8.1} are most likely to be the structural features responsible for susceptibility.

Testing the ability of different \textit{RPW8.1} and \textit{RPW8.2} alleles to confer resistance in \textit{A. thaliana} is an essential step towards understanding the relationship between structure and function of \textit{RPW8.1} and \textit{RPW8.2}. Assessing the effect of all the AAPs on \textit{RPW8} function requires transforming the susceptible accession Col-0 with \textit{RPW8} allele polymorphs and testing the transformant’s resistance phenotype. Xiao \textit{et al.} (2002) have determined that the cDNA of \textit{RPW8.1} from accession Sy-0, but not the cDNAs of \textit{RPW8.2} from Sy-0 or \textit{RPW8.1} and \textit{RPW8.2} from Ler, was partially resistant to the broad spectrum of powdery mildew pathogens. Therefore \textit{RPW8.1} containing F45L and F77L substitutions (Figure VI. 9) only partially conferred \textit{RPW8}-mediated resistance. \textit{RPW8.1} containing F31V, H33N, L39I, E43R, F45L and A148T substitutions (Figure VI.9) did not confer \textit{RPW8.1}-mediated resistance. \textit{RPW8.2} either containing T64S, K70E and D116G substitutions (Figure VI.9) or truncated after K143 and containing V17F, T64S, D116G substitutions (Figure VI.9) did not confer \textit{RPW8}-mediated resistance. According to the model by Chasman & Adams (2001) the T64S, but not the D116G, substitution is unlikely to effect protein structure and function, implying that the T64 residue is involved in the regulation of protein function.
Classes of non-functional amino acid mutants

<table>
<thead>
<tr>
<th>Amino acid polymorphisms</th>
</tr>
</thead>
<tbody>
<tr>
<td>unusual amino acid by class</td>
</tr>
<tr>
<td>unusual amino acid by class and charged</td>
</tr>
<tr>
<td>K70E</td>
</tr>
<tr>
<td>Neither unusual amino acid by class or charged</td>
</tr>
<tr>
<td>L39I, T64S</td>
</tr>
</tbody>
</table>

Table VI.5: Amino acid polymorphisms, organised by class, in the predicted polypeptides of the non-functional RPW8.2 alleles from ecotype Sy-0 and RPW8.1 and RPW8.2 alleles from ecotype Ler, or the partially functional RPW8.1 allele from ecotype Sy-0.

Structural features were predicted by various algorithms in RPW8.1 and RPW8.2 predicted polypeptides (III). It is possible to use these algorithms to predict the effect of polymorphisms on the structure (and possibly function) of the predicted polypeptides for RPW8.1 and RPW8.2 alleles. Structural predictions for secretory signal peptides, CC motifs, serine, threonine and tyrosine phosphorylation sites and O-β-GlcNAc modification sites were repeated with selected predicted polypeptides of RPW8.1 and RPW8.2 alleles.

A secretory signal peptide was predicted by SignalP2.0 in the predicted polypeptide of RPW8.2 (III.3). RPW8.2 predicted polypeptides of alleles from accessions Gy-0 and Petergof contain all the AAPs capable of affecting a secretory signal peptide prediction. SignalP2.0 predicted secretory signal peptides and cleavage sites in RPW8.2 from ecotypes Gy-0 and Petergof were identical to Ms-0. Therefore AAPs in the predicted polypeptides of RPW8.2 alleles do not alter secretory signal peptide predictions.

RPW8.1 from ecotypes Fr-6 and Uk-1 were predicted by COILS (weighted MTIDK matrix) to have a CC motif which was identical to Ms-0 and did not include the inserted repeat, probably because the inserted repeat duplicated a region from Ms-0 (DQWD) which was outside the CC motif in Ms-0. A CC motif was not predicted by COILS in RPW8.1 from accession Ts-7, due to AAPs: E138 in the inserted repeat, I142S and E144K. COILS predicted that the CC1 motif in RPW8.2 from accessions Gy-0 and Ksk-1B was shortened, due to the dissimilar AAPs (V68F, E77V and L98Q).
RPW8.2 from accessions Petergof or Nok-0B are truncated. In accession Petergof, RPW8.2 is truncated by two amino acids, which shortens, by two amino acids, the CCII motif predicted by COILS. RPW8.2 from accession Nok-0B is shorter than RPW8.2 from accession Petergof and was correspondingly not predicted by COILS to have a CCII motif. COILS predicted that the CC motifs in RPW8.2 from accessions C24 and Ler were identical to Ms-0. Therefore, there is no effect of T64S, K70E, L111I, D116G or T161K mutations on COILS predictions of CC1 and CCII motifs in RPW8.2.

NetPhos2.0 was used to predict serine, threonine and tyrosine phosphorylation sites in RPW8.1 from accessions Fr-6, Ms-0, Pla-2, Ts-7 and Ws-0 and RPW8.2 from accessions Co-1A, Ksk1-B, Ler, Ms-0, Petergof and Sah-0B: these are shown in Table VI.6. AAPs caused the loss or gain of NetPhos2.0 predicted serine, threonine and tyrosine phosphorylation sites. Gain of a NetPhos2.0 predicted phosphorylation site at S30 in RPW8.1 from accession Ws-0 was due to either the F31V or H33N AAPs. Phosphorylation site predictions by NetPhos2.0 were gained in the inserted repeat of RPW8.1 from accessions Fr-6 and Ts-7. Phosphorylation site predictions were gained at S122 and S144 in RPW8.1 from accession Ts-7. The gain of predicted phosphorylation sites at S122 and S144 was due to AAPs I122S and A148T, respectively. Two NetPhos2.0 predicted phosphorylation sites are gained in RPW8.2 from accession Ksk-1B by T64S and surrounding AAPs and S45T. The T64S AAP in RPW8.2 causes the loss of the NetPhos2.0 predicted phosphorylation site at T64.
Table VI.6: NetPhos2.0 predicted phosphorylation sites (denoted by P) in the predicted polypeptides of a) RPW8.1 alleles from ecotypes Fr-6, Ms-0, Pla-2, Ts-7 and Ws-0 and b) RPW8.2 alleles from accessions Co-1A, Ksk-1B, Ler, Ms-0, Petergof and Sah-0B.

<table>
<thead>
<tr>
<th>Phosphorylation sites</th>
<th>Ms-0</th>
<th>Pla-2</th>
<th>Ws-0</th>
<th>Fr-6</th>
<th>Ts-7</th>
</tr>
</thead>
<tbody>
<tr>
<td>S30</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>S53</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>S58</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>T68</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>Y75</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>S122</td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td>P</td>
</tr>
<tr>
<td>P130</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>INSERTED REPEAT</td>
<td>P</td>
<td>P</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S144</td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table VI.7: Predicted O-β-GlcNAc sites (denoted by $\beta$GN) and yin-yang regulation sites (denoted by $\mathbb{K}$) in the predicted polypeptides of a) RPW8.1 alleles (from accessions Fr-6, Ms-0 and Ts-7) and b)

O-β-GlcNAc modification and yin-yang regulation sites were predicted by YinOYang1.2 in RPW8.1 from accessions Ms-0, Ts-7 and Fr-6) and RPW8.2 from accessions Co-1A, Ksk-1B, Ler, Ms-0, Petergof and Sah-0B. These predictions are listed in Table VI.7. No AAPs affected O-β-GlcNAc modification site predictions in RPW8.1 predicted polypeptides. O-β-GlcNAc modification site predictions at S144 or S171 and S174 in RPW8.1 from accession Fr-6 or RPW8.2 from accession Sah-0B were lost because the polypeptides are truncated. An additional O-β-GlcNAc modification site was predicted in RPW8.2 from accession Sah-0B at S135. The loss of predicted O-β-GlcNAc modification sites at T168 and T64S was due to T160K and surrounding AAPs, respectively. The O-β-GlcNAc modification and yin-yang regulation site at T64 was lost in RPW8.2 from accessions that contained the T64S AAP.
**VI.3.3: Conclusion**

*RPW8.1* and *RPW8.2* evolution has been mediated by nucleotide substitution, possibly transposon excision and an intragenic recombination event between *RPW8.1* orthologues, resulting in single nucleotide polymorphisms, indels and an inserted direct repeat, respectively. The percentages of polymorphic nucleotide sites indicate that *RPW8.1* and *RPW8.2* are diversifying rapidly, although most of the nucleotide polymorphisms are present in alleles from phenotypically susceptible accessions. There is very little polymorphism between the alleles of phenotypically resistant accessions and Ka/Ks ratios between these ecotype alleles indicate that *RPW8.1* and *RPW8.2* (especially the *RPW8.2* intron) are conserved.

The structural features that are the strongest candidates that could be responsible for susceptibility are T64S and D116G in RPW8.2 and S40I in RPW8.1. T64S and D116G do not affect the CC forming potential of RPW8.2, although T64S abolished the predicted phosphorylation, O-β-GlcNAc modification and yin-yang regulation site at T64 in RPW8.2.

*RPW8.1* and *RPW8.2* are unique proteins that confer broad spectrum powdery mildew resistance (Xiao...
et al., 2001), meaning that the structure and function of RPW8.1 and RPW8.2 are of great interest. To completely understand the structure-function relationships of RPW8.1 and RPW8.2, their 3D structures must be determined and analysed in conjunction with data on the functional ability of RPW8.1 and RPW8.2 mutants. Already, Xiao et al. (2002) have tested the functional ability of RPW8.1 and RPW8.2 from accessions Ler and Sy-0, but further tests of the RPW8.2 T64S and D116G substitutions and the RPW8.1 S40I substitution are needed.

Transient expression of genes from the $RPW8$ locus generated the hypersensitive response in $Nicotiniana benthamiana$

VII.1: Introduction

To test the conclusions in VI.3.3 a functional assay for RPW8.1 and RPW8.2 expression is needed. A functional assay could be used to assess the structural features identified in VI.3 as candidate structural features that could be responsible for susceptibility. $RPW8.1$ and $RPW8.2$ mutants could (and should) be functionally assessed by stable expression in $A. thaliana$ ecotype Col-0, although this would be time consuming, laborious, slow and necessitate careful scoring of transgenic $A. thaliana$ lines for powdery mildew resistance. Transient expression is a fast, simple, efficient and reproducible technique for in vivo assays of gene expression and function in planta. A reliable and characterised transient expression system would provide a basis for a fast, easily identifiable, qualitative screen of $RPW8.1$ and $RPW8.2$ mutants for further investigation.

Stable transgenic expression of $RPW8.1$ or $RPW8.2$ in $A. thaliana$ ecotype Col-0, either as cDNA or genomic DNA, confers resistance to a broad spectrum of powdery mildews (Xiao, et al., 2001). However, $RPW8.1$ or $RPW8.2$-mediated resistance was intermediate between the resistances of ecotype
Col-0 transformed with cosmid B6 and non-transformed ecotype Col-0 plants (S. Xiao pers. comm.). An intermediate resistance conferred by *RPW8.1* or *RPW8.2*, compared to the full resistance conferred by genomic DNA containing *RPW8.1* and *RPW8.2*, implies that either both proteins are necessary or they act with significant synergism. Transient expression could assess the independent and concerted functional contributions of *RPW8.1* and *RPW8.2* in planta.


Transient expression of *avrPto* or *avrBs2* in tobacco plants transformed with *Pto* (Thilmony *et al.*, 1995; Rommens *et al.*, 1995) or *Bs2* (Tai *et al.*, 1999), respectively, generated an Avr-dependent HR. *N. tabacum* transformed with *Pto* generated an *avrPto*-dependent HR in transient expression assays and was resistant to *P. syringae pv. tabaci* (Tang *et al.*, 1996). *N. benthamiana* and tomato transformed with *Bs2* both generated an *avrBs2*-dependent HR and were resistant to *X. campestris pv. vesicatoria* carrying *avrBs2* (Tai *et al.*, 1999). Therefore functional R genes cause the HR in transient expression assays.

There are several advantages of transient expression over stable expression. Gene analysis, using transient expression, can be conducted within two or three days of DNA delivery (Negrutiu *et al.*, 1990). Expression via transient systems in plant cells occurs several orders of magnitude more frequently than in stable expression systems (De Buck *et al.*, 2000; Janssen & Gardner, 1989; Maximova *et al.*, 1998). Each transiently transformed plant cell contains more copies of non-integrated, transiently expressed DNA than copies of integrated, stably expressed DNA; therefore transient expression in plant cells produces more protein (Maximova *et al.*, 1998) and is not subject to
the “positional effects” of chromosomal integration (Gelwin, 1998). Most importantly, transient expression is independent of the physiological and biochemical history of the plant, and any detrimental or lethal consequences of expression. For example, transformation of tomato roots with a sequence for the C-terminal LRR encoding region of Mi-1.2 fused to the N-terminal encoding region of Mi-1.1 produced no viable transformants. This fused Mi gene was transiently expressed in N. benthamiana and caused an Avr-independent HR (Hwang et al., 2000). Therefore functional tests of the fused Mi gene could only be conducted using transient (and not stable) expression assays.

Stable transformation of A. thaliana with: a) RPW8.1 or RPW8.2 cDNAs produced stunted plants (S. Xiao pers. comm.) and b) multiple copies of the 7.5 kb fragment developed spontaneous HRs on the hypocotyls (S. Xiao, pers. comm.). Therefore RPW8.1 and RPW8.2, like the fused Mi gene product, are detrimental to plants and lethal to plant cells, so would be best tested for function using transient expression assays.

Transient expression methods use DNA delivery techniques like particle-bombardment, viral vectors or Agrobacterium-mediated transfer, in plant systems ranging from protoplasts and suspension cell cultures to intact tissues. Protoplasts and suspension cell cultures are difficult to prepare and are not robust plant systems; whereas intact tissues (which are easy to prepare) have cells that have not been chemically treated and are in situ. Particle bombardment needs expensive equipment and is brutal, often killing transformed plant cells (Leister et al., 1996; Yang et al., 2000). This causes variation in the transformation efficiency and necessitates co-transformation with an internal control to standardise the results (Yang et al., 2000). Viral vectors limit the size of the DNA insert that can be delivered and high viral recombination frequencies mean that large inserts tend to be unstable (Scholthof et al., 1996). Yang et al. (2000) reported that Agrobacterium-mediated transient expression in fully expanded leaves on six-week old and uniformly grown N. tabacum plants had excellent transformation efficiencies; they could also perform reproducible, multiple tests on a single leaf or on different leaves. Abbink et al. (1998), Bendahmane et al. (1999), Bendahmane et al. (2000), Erickson et al. (1999), Rathjen et al. (1999), Riely & Martin (2001), Scofield et al. (1996), Tai et al. (1999), Tang et al. (1996), Thomas et al. (2000), Hwang et al. (2000), Van der Hoorn et al. (2000) Van der Hoorn et al. (2001), Van der Hoorn et al. (2001b) and Wulff et al. (2001) have functionally assayed R and Avr genes for the HR in tobacco leaves via Agrobacterium-mediated transient expression. The functional effect
of R or Avr genes in tobacco can be assessed with transient expression in two ways. Firstly, by co-expression of both R and Avr genes in the same leaf region (Hwang et al., 2000; Tai et al., 1999; Van der Hoorn et al., 2000; Riely & Martin, 2001; Van der Hoorn et al., 2001b). Secondly, by expressing an R or Avr gene in a plant already transformed with the corresponding Avr or R gene (Bendahmane et al., 1999; Bendahmane et al., 2000; Scofield et al., 1996; Tang et al., 1996). Hence, Agrobacterium-mediated transient expression in N. benthamiana is an easy, robust and simple to score system, potentially suitable for RPW8.1 and RPW8.2 functional assays.

Protein phosphorylation is associated with the HR (I.1.1). Protein phosphorylation can be prevented by staurosporine (Ōmura et al., 1997) or K-252a (Kase et al., 1986). Staurosporine and K-252a are structurally related (Ōmura et al., 1994; Rüegg & Burgess, 1989) and inhibit protein kinase C (phospholipid/Ca\(^{2+}\)-dependent protein kinase) and cyclic nucleotide-dependent protein kinases by competing with ATP (Kase et al., 1986; Kase et al., 1987; Mizuno et al., 1993; Tamaoki et al., 1986). Staurosporine at concentrations of 10 \(\mu\)M or 1 \(\mu\)M (Suzuki et al., 1999; Zhang et al., 2000) or K-252a at 1 \(\mu\)M (Zhang et al., 2000) prevented MAP kinase activation and the HR in N. tabacum suspension cell cultures. The generation of ROS in N. tabacum suspension cell cultures was also prevented by 10 \(\mu\)M (Suzuki et al., 1999) or 1.25 \(\mu\)M (Viard et al., 1994) staurosporine. The prevention of ROS generation in N. tabacum suspension cell cultures by staurosporine coincided with inhibited protein phosphorylation (Viard et al., 1994) of 18 proteins (Lecourieux-Ouaked et al., 2000). K-252a at concentrations of one and 10 \(\mu\)M prevented the generation of ROS in soyabean suspension cell cultures (Chandra & Low, 1995; Guo et al., 1998; Rajasekhar et al., 1999). Therefore, staurosporine and K-252a can inhibit protein phosphorylation and prevent ROS generation and the HR in suspension cell cultures. However, neither 2.2 \(\mu\)M staurosporine nor 2.2 \(\mu\)M K-252a prevented the HR in N. tabacum leaves (He et al., 1994).

Histochemical detection of CaMV-35S driven \(\beta\)-glucuronidase (GUS) in situ is an accepted marker of T-DNA transformation efficiency (Kapila et al., 1997; Rossi et al., 1993) and CaMV-35S driven transient expression. GUS is a stable homotetrameric enzyme, encoded by the \textit{uid A} (\textit{gus A}) locus from \textit{E. coli} that hydrolyses \(\beta\)-
glucuronides (Jefferson et al., 1986; Jefferson et al., 1987). GUS expression can be visualised with a variety of β-glucuronides, hydrolysed to fluorophores or chromophores (Jefferson, 1989). X-Gluc is a colourless GUS substrate that when hydrolysed forms an insoluble, blue-indigo precipitate at the site of enzyme activity (Pearson et al., 1961; Figure VII.1). GUS is frequently used in planta (Jefferson, 1989; Naleway, 1992) where it accumulates in the cytoplasm (Martin et al., 1992). It has a low or undetectable intrinsic activity in higher plants, including tobacco (Jefferson et al., 1986; Jefferson et al., 1987). N. tabacum, transformed by Jefferson et al. (1987) with CaMV-35S driven GUS, showed consistent high expression and accumulation of GUS in the leaves of independent transformants and this was not detrimental to the plants. Also, in situ histochemical staining of GUS in the N. tabacum plants with X-Gluc was able to detect differences in GUS activity between single cells (Jefferson et al., 1987). Thus, CaMV-35S driven GUS activity is easily detected with X-Gluc in tobacco leaves in situ and is a clear, non-detrimental indicator of transformation at the single cell level.

Fig. VII.1: The chemical reactions of GUS catalysed X-Gluc hydrolysis and oxidative dimerisation to form an insoluble, blue precipitate.

Agrobacterium carrying binary vectors containing genomic fragments of the RPW8 locus were agro-infiltrated into regions of N. benthamiana leaves and assessed for indicators of the HR (cell death and autofluorescence). Protein kinase inhibitors were tested for their effect on RPW8-induced cell death in N. benthamiana. Also, various other species were agro-infiltrated with Agrobacterium carrying these binary vectors to determine the taxonomic restrictions of RPW8 function.
VII.2: Methods

VII.2.1: Transient expression of the RPW8 locus in Nicotiana benthamiana

Agrobacterial cells (strain GV3101) were streaked on a Rif-Gent-LB-Agar plate with an X-Ray sterilized BIOLOOQ® and incubated for 2 days at 30°C. An individual colony was picked with a sterile P2 pipette tip, inoculated into 10 ml of Rif-Gent-Agro-LB and grown on a shaker at 250 rpm for 2 days at 30°C. Duplicate 50 ml of Rif-Gent-Agro-LB were each inoculated with 0.5 ml aliquots of the 10 ml culture and grown on a shaker at 250 rpm at 30°C until OD$_{600}$ = 0.8. Both 50 ml cultures were then split equally into 50 ml centrifuge tubes (Nalge NUNC International) and centrifuged at 4,000 rpm for 10 minutes at 4°C. The supernatant was discarded and each pellet was re-suspended in 20 ml of sterile 10% glycerol. The re-suspended pellets were re-centrifuged at 4,000 rpm for 15 minutes at 4°C. The supernatant was discarded, each pellet was re-suspended as before in 20 ml of sterile 10% glycerol, and then re-centrifuged at 4,000 rpm for 15 minutes at 4°C. The supernatant was discarded and the pellets were combined to a total volume of approximately 1 ml in sterile 10% glycerol. Combined pellets of electrocompetent cells were divided into aliquots of 40 µl, then snap frozen in liquid N$_2$, and stored at -80°C.

Aliquots of electrocompetent GV3101 were defrosted on ice. Approximately 100 ng of pBIN19+
(Bevan, 1984), pSLJ755I5\textsuperscript{103} (Jones et al., 1992), pKMB\textsuperscript{103} (Mylne & Botella, 1998), GUS-pBIN19+ (Vancanneyt et al., 1990), RPW8.1-pKMB\textsuperscript{103}, RPW8.2-pKMB\textsuperscript{103}, 3.7kb-pSLJ755I5\textsuperscript{103}, 3.8kb-pSLJ755I5\textsuperscript{103}, 6.2kb-pSLJ755I5\textsuperscript{103}, and 7.5kb-pBIN19+\textsuperscript{103}, binary vector DNAs were each mixed with 20 µl of electrocompetent cells and electroporated in a Cell Porator\textsuperscript{®} (Life Technologies).

Porated cells were incubated in 1 ml of Agro-LB on a shaker at 250 rpm at 30\textdegree{}C for 1 hour and then centrifuged at 14,000 rpm for 5 minutes. The supernatant was discarded and the pellets were re-suspended in 200 µl of Agro-LB. Half (100 µl) was plated on to Rif-Gent-Kan-Agro-LB or Rif-Gent-Tet-Agro-LB plates and incubated for 3 days at 30\textdegree{}C.

Twelve individual colonies of each binary vector were picked with sterile P2 pipette tips to fresh plates of the same media and antibiotics, and then to 0.2 ml thin walled PCR tubes (Hybaid). The plates were incubated for 3 days at 30\textdegree{}C. The PCR tubes were heated for 1 minute at maximum temperature in a Panasonic microwave oven before adding 20 µl of PCR mix. MC2 F (5′ – ATG CCG ATT GGT GAG CTT GCG ATA – 3′) and MC2 R (5′ – TCA AGC TCT TAT TTT ACT ACA AGC – 3′) primers (MWG Biotech AG) were used to amplify RPW8.1. MC3 F (5′ – ATG ATT GCT GAG GTT GCC GCA – 3′) and MC3 R (5′ – TCA AGA ATC ATC ACT GCA GAA CGT – 3′) primers were used to amplify RPW8.2. T3 (5′ – AAT TAA CCC TCA CTA AAG GGA ACA – 3′) and T7 (5′ – GTA ATA CGA CTC ACT ATA GGG CGA – 3′) primers were used to amplify across the multiple cloning site of pBIN19+ containing GUS. Negative controls contained no DNA; the positive controls were cosmid B6 (Figure II.3). The thermal cycle conditions (Figure VII.2) were generated in a Hybaid Omnigene Thermal Cycler.

Fig. VII.2: PCR thermal cycle conditions used to amplify genomic DNA from the RPW8 locus to confirm the presence and orientation of inserts in the binary vectors.

Each 20 µl PCR was mixed with 4 µl of 6X orange G loading dye and electrophoresed on a 1% agarose gel at 120V against 500 ng of 1 kb DNA ladder marker (GIBCOBRL\textsuperscript{TM}). Colonies with PCRs
containing a single, correctly sized band were inoculated into 10 ml of either Rif-Gent-Kan-Agro-LB or Rif-Gent-Tet-Agro-LB and grown on a shaker at 250 rpm for 3 days at 30°C. Glycerol stocks of the cultures containing the binary vectors were prepared by mixing 0.8 ml of culture with 0.2 ml of sterile glycerol in 1 ml cryovials, snap freezing them in liquid N\textsubscript{2} and then storing them at -80°C.

A 5 \( \mu l \) aliquot of each glycerol stock was inoculated into 10 ml of either Rif-Gent-Kan-Agro-LB or Rif-Gent-Tet-Agro-LB and grown on a shaker at 250 rpm for 3 days at 30°C. One millilitre of each culture was inoculated into 50 ml of either Rif-Gent-Kan-Agro-LB or Rif-Gent-Tet-Agro-LB and grown on a shaker at 250 rpm at 30°C for 36 hours. Half of each culture was centrifuged in a 50 ml centrifuge tube at 4,000 rpm for 10 minutes at 4°C. The supernatant was decanted and the cell pellet was re-suspended in agro-infiltration media to OD\textsubscript{600} = 1.5 +/-0.1. In protein kinase inhibitor tests, either 20 \( \mu l \) of 500\( \mu M \) staurosporine (CALBIOCHEM) in dimethyl sulphoxide (DMSO, SIGMA) or 20 \( \mu l \) of 500\( \mu M \) K-252a\textsuperscript{121} (CALBIOCHEM) in DMSO were added to 2 ml of Agrobacterium (carrying 6.2kb-pSLJ755I5 and 7.5kb-pBIN19+ binary vectors) re-suspended in agro-infiltration media.

\textit{N. benthamiana} plants were sown in the greenhouse at the University of East Anglia, and watered regularly by D. Alden. At six weeks old, numbered plants were selected randomly for treatment, and assigned numbered treatments at random, using the RAND function of a Ti-80 graphic calculator (Texas Instruments). Five plants were selected for each treatment. Each plant was infiltrated by “knicking” the abaxial surface of a fully expanded leaf with a razor blade (just sufficiently to break the cuticle) to allow a 2 ml sterile, needle-less syringe placed over the “knick” to force the agro-infiltrate into the intercellular spaces. Agro-infiltrated leaves were marked with a strip of Micropore tape on the petiole and their symptoms were recorded at 24 hour intervals post agro-infiltration. Leaf cell death symptoms were scored either by eye or with a PRIOR SM3A Binocular microscope, used to view the abaxial surface of \textit{N. benthamiana} leaves. The leaves were photographed at the John Innes Centre 6 days post agro-infiltration (dpinf).

\textbf{VII.2.2: Histochemical detection of transient expression in \textit{Nicotiana benthamiana} with $\beta$-}
Leaves and leaf discs (cut with an EtOH cleaned, 11 mm cork borer) from pBIN19+ and GUS-pBIN19+ agro-infiltrated regions of the same leaves were harvested at 3 and 6 days dpinf. The leaf tissue was fixed in 2% (EM GradeII) glutaraldehyde (v/v, SIGMA®), in 1 mM EDTA and 50 mM Na-Phosphate (pH7.0) for 3 minutes on ice in a glass petri dish. The leaf tissue was then washed three times in ice-cold 50 mM Na-phosphate (pH7.0) on ice in a glass petri dish, with blotting using 3MM chromatography paper (Whatmann®) between washes. GUS substrate solution was vacuum infiltrated into leaf discs or infiltrated into the intercellular spaces of non-agro-infiltrated areas on the abaxial surface of leaves using the “knicking” and syringe method (VII.2.1). Leaves and leaf discs infiltrated with GUS substrate solution were immersed in GUS substrate solution and incubated at 37°C for 24 hours. GUS staining was stopped and the leaf tissue was decolourised by immersing the leaf tissue in EtOH on a shaker overnight. GUS stained tissue was photographed in the Department of Environmental Sciences at the University of East Anglia.

VII.2.3: Epifluorescence microscopical detection of autofluorescence

Three leaves with 7.5kb-pBIN19+, pBIN19+, 6.2kb-SLJ755I5 and pSLJ755I5 agro-infiltrated regions were harvested and infiltrated with 20mM (w.r.t. citrate) Na-citrate-phosphate buffer (pH5.6) using the “knicking” and syringe method in VII.2.1. Three sections (10 mm²) of agro-infiltrated regions were cut from each leaf and mounted on glass microscope slides (76 mm x 26 mm x 1-1.2 mm, BLUE STAR), under glass cover-slips (No.1½, 22 mm², Chance Propper Ltd) in 20mM (w.r.t. citrate) Na-citrate-chosphate buffer pH5.6. Leaf sections were viewed with a NIKON eclipse E800 microscope with VFM epifluorescence attachment using the UV-2A filter set. Photographs of leaf sections were taken with a FDX-35 35 mm camera using Kodak Professional PORTRA 400vc film.

VII.2.4: Transient expression of the RPW8 locus in various species
Three to six plants of various species were grown at: IACR-Broom’s Barn (Chinese cabbage and sugar beet), the John Innes Centre (*Chen amaranthus, N. clevelandii, N. samsun*, turnip and potato), Nottcutts Garden Centre (lettuce, marrow and pepper) or the greenhouses at the University of East Anglia (*N. benthamiana, N. tabacum* and tomato). Three leaves of each plant were agro-infiltrated (in a single test) with *Agrobacterium* carrying 7.5kb-pBIN19+, pBIN19+, 6.2kb-pSLJ755I5 and pSLJ755I5 binary vectors, as in VII.2.1.

**VII.3: Results & Discussion**

**VII.3.1: RPW8.1 and RPW8.2 are necessary for the *RPW8*-induced, avirulence-independent hypersensitive response in *Nicotiana benthamiana***

Transient expression of T-DNAs containing genomic *RPW8.1* and *RPW8.2* (from 7.5kb-pBIN19+ and 6.2kb-pSLJ755I5 binary vectors) caused weak cell death in agro-infiltrated regions between three and five dpInf (Figure VII.3; Table VII.1). The symptoms of cell death were silvery patterns on the abaxial surface of the leaf, with or without necrotic lesion formation. The silvery patterns were identified (using the binocular microscope) as areas where mesophyll cells had collapsed, drawing the abaxial leaf surface into the leaf blade (data not shown). Mesophyll cell collapse caused the trichodermal (leaf hair) cells, which were previously extended perpendicular to the abaxial leaf surface, to topple. Toppled trichodermal cells reflected incident light, thus causing the abaxial leaf surface to appear silvery. Cell death was not seen in regions agro-infiltrated with media, *Agrobacterium* or any of the other binary vectors (Figure VII.3; Table VII.1). Histochemical staining of *N. benthamiana* leaf tissue for GUS activity proved that the T-DNAs were transiently expressed in the agro-infiltrated regions, *in planta* (Figure VII.4). Therefore, transiently expressed T-DNAs containing genomic *RPW8.1* and *RPW8.2*, but not *RPW8.1* or *RPW8.2* (either as cDNAs or genomic DNA), were capable of causing *RPW8*-induced, Avr-independent cell death in *N. benthamiana* leaves, implying that both *RPW8.1* and *RPW8.2* proteins are necessary for cell death.

<table>
<thead>
<tr>
<th>Binary vector</th>
<th>Number of tests (number of agro-infiltrations)</th>
<th>% Cell death</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table VII.1: Cell death induction in several independent transient expression tests (each of five agro-infiltrations) in *N. benthamiana* due to the expression of T-DNAs from the binary vectors listed below.
Autofluorescence is a natural and specific indicator of the HR (I.1.1). Epifluorescence microscopy detected (in three independent tests, of three agro-infiltrations and three leaf sections per agro-infiltration) patches of collapsed and autofluorescent *N. benthamiana* mesophyll cells in regions agro-infiltrated with binary vector T-DNAs containing genomic *RPW8.1* and *RPW8.2* (Figure VII.5). Collapsed and autofluorescent cells were not found in *N. benthamiana* leaf tissues, whether agro-infiltrated with any of the other binary vectors or not. Therefore the accumulation of autofluorescent material associated with *RPW8*-induced, Avr-independent cell death proves the cell death was the HR.

In three independent tests of at least 15 assays, the protein kinase inhibitors staurosporine (5µM) and K-252a (5µM) failed to prevent the HR caused by binary vectors containing genomic *RPW8.1* and *RPW8.2* T-DNAs in *N. benthamiana* (data not shown). Therefore, neither protein kinase C nor cyclic nucleotide-dependent protein kinases are essential for the *RPW8*-induced, Avr-independent HR in *N. benthamiana* leaves. However, inhibiting these protein kinases in suspension cell cultures might have prevented this HR (VII.1).

Unexpectedly, both *RPW8.1* and *RPW8.2* proteins were required for the *RPW8*-induced, Avr-independent HR in *N. benthamiana*. *RPW8.1* and *RPW8.2* (as genomic or cDNA sequences) are both sufficient for resistance, albeit partial (S. Xiao pers. comm.), to *E. cichoracerum* isolate UCSC1 in *A. thaliana* (Xiao et al., 2001). However, in *A. thaliana* accession Col-0, there are homologues of *RPW8.1* and *RPW8.2* at the *RPW8* locus (Xiao et al., 2002), *HR1, HR2* and *HR3*, which might be sufficiently functional to replace the missing *RPW8.1* or *RPW8.2* partner for partial resistance. No
RPW8 hybridising sequences were found in the genomic DNA from *N. benthamiana* (S. Xiao pers. comm.). Hence, the difference in functional requirements between resistance in *A. thaliana* and the HR in *N. benthamiana* could be explained by partially functional RPW8 homologues at the RPW8 locus of *A. thaliana* accession Col-0.

There have only been two other cases of Avr-independent HR, which were caused by the transient expression of mutant *Pto* and *Mi* genes. The *Pto* mutant that caused an Avr-independent HR was constitutively active (Rathjen et al., 1999). Domain swaps between *Mi-1.2* and its closely related, non-functional homologue, *Mi-1.1*, generated an Avr-independent HR in *N. benthamiana* if the fused Mi sequences were composed of the Mi-1.2 LRR encoding region, the Mi-1.1 NBS encoding region and either the NT1 (CC encoding) or the NT2 (CC encoding) sequences of *Mi-1.1* (Hwang et al., 2000). This suggests that in the Mi-1.2 protein both the Mi-1.2 CC domains were necessary to prevent an Avr-independent HR and probably function to autoinhibit the Mi-1.2 LRR generated HR, which is presumably released from this regulation by the Avr protein elicitor. The CC domains in RPW8.1 and RPW8.2 are probably also involved in an HR regulatory role, requiring the presence of RPW8.1 and RPW8.2 CC domains, but not as inhibitors of an LRR.

**VII.3.2: RPW8.1 and RPW8.2 are probably not sufficient for the RPW8-induced, Avr-independent hypersensitive response in *Nicotiana benthamiana***

Co-expression of T-DNAs containing either genomic RPW8.1 with genomic RPW8.2 or RPW8.1 cDNA with RPW8.2 cDNA in *N. benthamiana* leaves failed to cause cell death (Figure VII.3; Table VII.1). This implies that RPW8.1 and RPW8.2 proteins are not sufficient to cause RPW8-induced, Avr-independent HR. The absence of cell death was not due to a lack of transient expression or stable protein: the CaMV 35S promoter is active in tobacco leaf tissue (Odell et al., 1985; Figure VII.4) and RPW8.1 and RPW8.2 were both transiently expressed (from cDNAs on the pKMB binary vector T-DNA) as stable, GFP fusion proteins in *N. benthamiana* (IX.3). However, a low co-transformation efficiency (*i.e.* the percentage of leaf cells transformed with both (genomic or cDNA) RPW8.1 and RPW8.2) could have meant that only a few cells were responding; this might not have been detected.

Transient expression of T-DNAs containing (cDNA or genomic) RPW8.1 or RPW8.2 in several leaves
of six T$_2$, *N. benthamiana* lines stably transformed with the 7.5 kb fragment failed, in a single test, to cause cell death in regions agro-infiltrated with *Agrobacterium* carrying the 7.5kb-pBIN19+ and 6.2kb-pSLJ755I5 (Figure VII.6). Therefore, RPW8.1 and RPW8.2 proteins are probably not sufficient to cause cell death in *N. benthamiana*; although the *RPW8*-induced, Avr-independent HR could be dependent on the amount of RPW8.1 and RPW8.2 proteins. For instance, *A. thaliana* accession Col-0 stably transformed with the 7.5kb fragment developed spontaneous HRs on the cotyledons (Brown, 2001) and the incidence of HR development correlated with the number of 7.5 kb fragment hybridising sequences in the genomes of the transgenic *A. thaliana* (S. Xiao pers. comm.). Spontaneous HRs were not seen on the leaves of *N. benthamiana* plants transformed with the 7.5 kb fragment. Also, transient expression of the 3.8 kb fragment with an additional CaMV 35S promoter
generated an Avr-independent HR in *N. benthamiana* (A. Balmuth pers. comm.) and the same T-DNA, stably transformed in *A. thaliana* caused spontaneous lesions (S. Xiao pers. comm.).

To determine if RPW8.1 and RPW8.2 proteins are sufficient for cell death in *N. benthamiana*, RPW8.1 and RPW8.2 must be, a) transiently expressed by the CaMV 3S promoter (so similar amounts of RPW8.1 and RPW8.2 proteins are produced and the expression would be comparable to RPW8.1-pKMB and RPW8.2-pKMB) and b) from the same T-DNA (to prevent any problems with transformation efficiencies). In order to express RPW8.1 and RPW8.2 cDNAs from the same T-DNA, in a system comparable to the transient expression of RPW8.1 or RPW8.2 cDNAs from the pKMB binary vector T-DNAs RPW8.1 and RPW8.2 cDNAs will be separately cloned behind CaMV 3S promoters (Figure VII.7) in the pBTEX binary vector. Frederick *et al.* (1998) used the pBTEX binary vector system – which can produce a T-DNA containing two genes in opposing orientations, each with CaMV 3S promoter (Figure VII.7) – to simultaneously transiently express Pto and avrPto, generating an avrPto-dependent, Pto-mediated HR in *N. benthamiana*. Using the pBTEX binary vector system RPW8.1 and RPW8.2 cDNAs will be transiently expressed, via Agrobacterium-mediated transfer, from pBTEX T-DNA in regions of *N. benthamiana* leaves, to simultaneously express RPW8.1 and RPW8.2 proteins.

![Fig. VII.7: T-DNA of the pBTEX binary vector, containing RPW8.1 and RPW8.2 cDNAs, for the simultaneous transient expression of RPW8.1 and RPW8.2 proteins in *N. benthamiana*. NPT II is the plant selectable marker on the pBTEX T-DNA. LB represents the left T-DNA border and RB represents the right T-DNA border. T is the terminator.](image)

If RPW8.1 and RPW8.2 proteins are necessary, but not sufficient, for cell death in *N. benthamiana* then there must be another factor (Factor X) produced by the 7.5 kb and 6.2 kb, but not by the 3.7 kb or 3.8 kb, fragments. Hence, Factor X would have to be encoded by a gene (*RPW8fx*) within the sequence of the 7.5 kb and 6.2 kb fragments, but located at a position which would prevent Factor X expression from the 3.7 kb and 3.8 kb fragments. Figure VII.8 shows the output of a NIX analysis for the 6.2 kb
fragment. Within this 6.2 kb fragment there is a small sequence, $RPW8fx$, predicted to be a gene by Fgene (Figure VII.8). Figure VII.9 shows the positions of the 6.2 kb, 3.7 kb and 3.8 kb fragments as well as $RPW8fx$. The full $RPW8fx$ sequence is not present in the 3.7 kb fragment, although it is present on the 3.8 kb fragment. Despite the full $RPW8fx$ sequence being present on the 3.8 kb fragment, essential sequences of an $RPW8fx$ promoter could be located in the 6.2 kb fragment between 0 bp and 2,361 bp, which would explain why the 3.8 kb fragment did not produce Factor X. Also, $RPW8fx$ spans the intron of $RPW8.2$, which could explain why the $RPW8.2$ intron appears to be functionally conserved (VI.3.1.2).

Fig. VII.9: An alignment of the 3.7 kb and 3.8 kb fragments against the 6.2 kb fragment showing the positions of $RPW8.1$, $RPW8.2$ and $RPW8fx$. The numbers indicate distances in base pairs along the 6.2 kb fragment. Green boxes, thick black lines and arrows indicate exons, introns and the direction of transcription, respectively.

The predicted polypeptide of $RPW8fx$ (i.e. Factor X) is 33 residues long. As $RPW8.1$ is localised to the cytoplasm and nucleoplasm and $RPW8.2$ is probably apoplastic (IX.3), then a potential function for Factor X might be transmembrane signal transduction. A TM helix, a motif found in R gene predicted polypeptides that recognise biotrophic fungal pathogens (Figure I.4a), may be sufficient to transduce a signal across the plasma membrane. Indeed, 33 residues is just long enough for a TM helix (I.2.1.2). However, TMHMM2.0 did not predict a TM helix in the predicted polypeptide of $RPW8fx$.

VII.3.3: The restricted taxonomic functionality of $RPW8$

An $RPW8$-induced, Avr-independent HR, caused by binary vector T-DNAs containing genomic $RPW8.1$ and $RPW8.2$, was detected in $N. benthamiana$, $N. clevelandii$, $N. Samsun$ and $N. tabacum$, but not in members of the Cruciferaceae or in other members of the Solanaceae (Table VII.2). Some species e.g. sugar beet were difficult to agro-infiltrate (Table VII.2) and therefore the HR was not expected, although Van der Hoorn et al. (2000) detected (albeit at a low level) GUS staining in agro-infiltrated regions of sugar beet leaves. Hence, the $RPW8$-induced, Avr-independent HR is taxonomically restricted to Nicotiana sp.. $RPW8$ is also the first cruciferous R gene to be functional in a solanaceous species; Cf-4, a solanaceous R gene was functional in lettuce, but no cruciferous R genes have been shown to be functional in solanaceous species (I.3.2.1). Therefore, $RPW8$-mediated
resistance in *A. thaliana* is likely to involve a signalling pathway that includes components found in the *Solanaceae*.

Table VII.2: Species assayed by transient expression for the ability to mediate *RPW8*-induced, Avr-independent HR.

<table>
<thead>
<tr>
<th>Species</th>
<th>Agro-infiltration ability</th>
<th><em>RPW8</em>-induced HR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chen amaranthus</td>
<td>poor</td>
<td>X</td>
</tr>
<tr>
<td>Marrow</td>
<td>poor</td>
<td>X</td>
</tr>
<tr>
<td>Cabbage</td>
<td>poor</td>
<td>X</td>
</tr>
<tr>
<td>Chinese cabbage</td>
<td>poor</td>
<td>X</td>
</tr>
<tr>
<td>Sugar beet</td>
<td>poor</td>
<td>X</td>
</tr>
<tr>
<td>Pepper</td>
<td>good</td>
<td>X</td>
</tr>
<tr>
<td>Tomato</td>
<td>good</td>
<td>X</td>
</tr>
<tr>
<td>Turnip</td>
<td>good</td>
<td>X</td>
</tr>
<tr>
<td>Potato</td>
<td>good</td>
<td>X</td>
</tr>
<tr>
<td>Lettuce</td>
<td>good</td>
<td>X</td>
</tr>
<tr>
<td><em>N. Samsun</em></td>
<td>good</td>
<td>✓</td>
</tr>
<tr>
<td><em>N. clevelandii</em></td>
<td>good</td>
<td>✓</td>
</tr>
<tr>
<td><em>N. tabacum</em></td>
<td>good</td>
<td>✓</td>
</tr>
<tr>
<td><em>N. benthamiana</em></td>
<td>good</td>
<td>✓</td>
</tr>
</tbody>
</table>

VII.3.4: Conclusion

Transient expression of genomic *RPW8.1* and *RPW8.2* in *N. benthamiana* leaves caused a weak HR, between three and five dpinf., in mesophyll cells. This *RPW8*-induced, Avr-independent HR was taxonomically restricted to *Nicotiana* sp. and did not require protein kinase C or cyclic nucleotide-dependent protein kinases. *RPW8*.*1* and *RPW8*.*2* proteins were necessary, but probably not sufficient for the *RPW8*-induced, Avr-independent HR in *N. benthamiana*. A third component, which I have called Factor X (probably encoded by *RPW8fx*), might also be necessary for the *RPW8*-induced, Avr-independent HR in *N. benthamiana*. The *RPW8fx* predicted polypeptide was not predicted to contain a TM helix.
Complete suppression of the $RPW8$-induced hypersensitive response by virus-induced gene silencing of $Sgt1$

VIII.1: Introduction

Transient expression of the $RPW8$ locus generated an Avr-independent HR in $N. benthamiana$ (VII.3). $RPW8$ is the first example of a non-solanaceous R gene that is functional in the Solanaceae (VII.3.3), implying $RPW8$ might initiate the HR via an unusual pathway for a cruciferous R gene. Hence, components of the $RPW8$-induced, Avr-independent HR signal transduction pathways are interesting because they might be key to understanding $RPW8$ function.

Virus-induced gene silencing (VIGS) technology is a genetic approach to plant gene function with all the benefits of transient expression systems (Baulcombe, 1999). Infection of a plant with viral vectors carrying a sequence homologous to the exons of any targeted plant gene will result in VIGS (Baulcombe, 1999; Marathe et al., 2000). VIGS involves post-transcriptional gene silencing (PTGS), a mechanism of systemic, sequence-specific, mRNA degradation in the cytoplasm, which probably acts as an anti-viral defence mechanism (Baulcombe & English, 1996; Baulcombe, 1999; Lindbo et al., 2001; Marathe et al., 2001; Meins, 2000; Sijen & Kooter, 2000; van Kammen, 1997; Vaucheret et al., 2001; Wang & Waterhouse, 2001). Excessive sequence-specific mRNA abundance, above a certain threshold, is thought to initiate PTGS. It is proposed that the steady-state abundance of homologous mRNAs is reduced during PTGS by RNA-dependent-RNA-polymerase synthesis of sequence-specific sense and anti-sense RNAs, complimentary to the mRNA, that with the mRNA form double stranded (ds) RNAs, which are cleaved by dsRNAses (Depicker & Van Montagu, 1997; Lindbo et al., 2001; Meins, 2000; Sijen & Kooter, 2000; Vaucheret et al., 2001; Wang & Waterhouse, 2001). Hence, VIGS systemically suppresses the expression of the targeted plant gene, producing a phenotype analogous to a loss-of-function mutation, even if there is redundancy in the genome.

VIGS has been used to investigate R gene signal transduction pathways. VIGS demonstrated that $Eds1$
(Peart et al., 2002) and Sgt1 (Peart et al., 2002b) were required for R gene mediated resistance in *N. benthamiana*. Peart et al. (2002b) also showed by silencing Sgt1 that it was essential for R gene mediated HRs in *N. benthamiana* when R genes and their corresponding Avr genes were transiently co-expressed (via *Agrobacterium*-mediated DNA transfer). VIGS of CDPK2 and closely related subfamily members significantly reduced and delayed *Avr4* and *Avr9*-dependent *Cf-4* and *Cf-9*-mediated HRs after *Agrobacterium*-mediated co-expression in *N. benthamiana* (Romeis et al., 2001). Also, VIGS did not significantly interfere with *Agrobacterium*-mediated transient expression. There was no detectable difference in GFP protein abundance, produced by *Agrobacterium*-mediated transient expression of GFP, in silenced and non-silenced *N. benthamiana* leaves (Peart et al., 2002b). Therefore VIGS can effectively silence components of resistance signal transduction pathways in *N. benthamiana* and is compatible with *Agrobacterium*-mediated transient expression.

The Tobacco rattle virus (TRV) VIGS vector was developed by Ratcliff et al. (2001). TRV initiated VIGS in all *N. benthamiana* tissues from 10 days post-inoculation (dpi) and its very mild symptoms have no effect on the systemic silencing signal. VIGS in TRV inoculated *N. benthamiana* declined after 28 dpi, a phenomenon known as recovery. New tissue that emerged after 28 dpi was not subject to VIGS. VIGS initiated by TRV and Potato virus X (PVX) vectors in *N. benthamiana* were similar (Ratcliff et al., 2001). PVX initiated VIGS suppression of *GFP* in *N. benthamiana* was greatest at 21 dpi (Ruiz et al., 1998). PVX:GFP initiated VIGS meant that the abundance of the targeted *GFP* mRNA was only detectable in the uppermost, silenced *N. benthamiana* leaves after 27 dpi (Ruiz et al., 1998). Therefore TRV initiated VIGS develops between 10 and 21 dpi; after 27 dpi TRV initiated VIGS in *N. benthamiana* plants begins to recover.

In this study the expression of *Eds1*, *N* and *Sgt1* resistance signal transduction pathway component genes was suppressed in *N. benthamiana* plants by VIGS. Silenced plants were then tested for *RPW8*-induced, Avr-independent HR competence.

**VIII.2: Methods**

*N. benthamiana* plants, at the Sainsbury laboratory, the John Innes Centre, were inoculated on three separate occasions by J. Peart, following the procedure described by Ratcliff et al. (2001) with TRV
vectors (Ratcliff et al., 2001) carrying target sequences, specific for components of R gene mediated resistance signalling pathways, as described in Peart et al. (2002). Target sequences carried by TRV vectors for VIGS were NbEds1, N, NbSgt1.1 and NbSgt1.2 (Peart et al., 2002; Peart et al., 2002b). Control plants were either not inoculated or inoculated with an empty TRV vector (TRV:00).

At 21 dpi (or 42 dpi for VIGS recovered) A. tumefasciens cultures carrying pSLJ755I5, 6.2kb-pSLJ755I5, pBIN19+ and 7.5kb-pBIN19+ binary vectors were agro-infiltrated into at least three leaves of at least three N. benthamiana plants inoculated with TRV vectors. Cultures were prepared and agro-infiltrated as described in VII.2.1.

Representative leaves for each TRV vector inoculation were harvested at 6 dpinf. and photographed.

VIII.3: Results

In three independent tests (of at least three agro-infiltrations on at least three plants), all leaves on N. benthamiana plants silenced by either TRV:NbSgt1.1 or TRV:NbSgt1.2 and transiently expressing either 7.5kb-pBIN19+ or 6.2kb-pSLJ755I5 binary vector T-DNAs failed to generate an RPW8-induced, Avr-independent HR (Figure VIII.1 e and f), unless plants had recovered from VIGS (Figure VIII.1 c). An RPW8-induced, Avr-independent HR was consistently seen in TRV:00 (Figure VIII.1 d), TRV:NbEds1 (data not shown), TRV:N (Figure VIII.1 b) and non-inoculated (Figure VIII.1 a) N. benthamiana plants transiently expressing either 7.5kb-pBIN19+ or 6.2kb-pSLJ755I5 binary vector T-DNAs. The RPW8-induced, Avr-independent HR appeared to be stronger in N. benthamiana plants inoculated with TRV:N (Figure VIII.1 b). Therefore VIGS of NbSgt1.1 or NbSgt1.2, but not NbEds1 or N, completely suppressed the RPW8-induced, Avr-independent HR in N. benthamiana.

VIII.4: Discussion

The transcription of NbEds1 and the expression of NbSgt1.1 and NbSgt1.2 was significantly reduced in N. benthamiana plants after VIGS with TRV:NbEds1, TRV:NbSgt1.1 and TRV:NbSgt1.2 vectors. NbEds1 mRNA abundance was specifically reduced in N. benthamiana inoculated with TRV:NbEds1 (Peart et al., 2002). Western blots, probed with anti-SGT1 antibodies showed that at 21 dpi SGT1
protein was substantially less abundant in plants inoculated with TRV:NbSgt1 than TRV:00 (Peart et al., 2002b). Therefore Eds1 and Sgl1 expression were silenced in N. benthamiana plants inoculated with TRV:NbEds1, TRV:NbSgt1.1 and TRV:NbSgt1.2 vectors and transiently expressing either 7.5kb-pBIN19+ or 6.2kb-pSLJ75515 binary vector T-DNAs. However, the remote possibility that a close homologue of Sgl1 in N. benthamiana was actually responsible for completely suppressing the RPW8-induced, Avr-independent HR cannot be excluded.

Complete suppression of RPW8-induced, Avr-independent HR by VIGS of Sgl1 in N. benthamiana demonstrates that SGT1 is an essential signal transduction pathway component for RPW8-induced, Avr-independent HR. SGT1 is probably a positive regulator of ubiquitin-mediated proteolytic degradation (1.3.2.2.5), which suggests that the degradation of ubiquitinated proteins is essential for RPW8-induced, Avr-independent HR. VIGS of Eds1 and N did not suppress the RPW8-induced, Avr-independent HR in N. benthamiana, demonstrating that EDS1 and N are not necessary for RPW8-induced, Avr-independent HR in N. benthamiana. VIGS of NtSgt1 by Peart et al. (2002b) completely suppressed N, Pto and Rx-mediated, Avr-dependent HRs and N, Pto and Rx-mediated resistances in N. benthamiana. Therefore R genes which required Sgl1 for HR generation in N. benthamiana also required Sgl1 for resistance. Hence, RPW8-mediated resistance is expected to require Sgl1.

VIGS of Sgl1 might permit the extraction and purification of proteins involved in generating the HR for biochemical studies, such as the co-immunoprecipitation of RPW8 interacting proteins and the structural determination of RPW8.1 and RPW8.2. Significant quantities of heterologous protein can be purified from transiently transformed N. benthamiana plants within a short time-scale. For instance, Vaquero et al. (1999) purified sufficient functional antibody from tobacco leaf extracts within a week after agro-infiltration of Agrobacterium carrying binary vector T-DNA encoding antibody sequence for SDS-PAGE, Western blotting and ELISA. Therefore extracting and purifying RPW8.1 and RPW8.2 from N. benthamiana leaf tissue is a viable alternative to heterologous expression in E. coli or S. pombe. Perhaps significant quantities of functional RPW8.1, RPW8.2 and RPW8 interacting proteins could be extracted from N. benthamiana leaf tissue transiently expressing either 7.5kb-pBIN19+ or 6.2kb-pSLJ75515 by completely suppressing the RPW8-induced, Avr-independent HR with VIGS of Sgl1.

Expression and subcellular localisation of RPW8.1 and
IX.1: Introduction

An important aspect of any R protein function is its subcellular localisation. The subcellular localisation of an R protein coincides with the site of elicitor recognition and therefore matches the pathogen’s *modus operandi* (I.3.1.2). Powdery mildews are biotrophic fungal pathogens (II.1), meaning their site of elicitor recognition and the subcellular localisation of RPW8 proteins is expected to be at the apoplastic face of the plasma membrane (I.3.1).

Green fluorescent protein (GFP) can be tagged to any protein for monitoring expression and subcellular localisation (Conn, 1999; Cubitt *et al*., 1995; Haseloff & Siemering, 1998; Gerdes & Kaether, 1996). GFP fluorescence and the biological function of the fused protein is not affected by fusion with proteins at either N or C terminus, although it is advisable to use a linker to prevent steric hindrance (Cubitt *et al*., 1995; Gerdes & Kaether, 1996; Yang *et al*., 1996). For example, EXU proteins tagged at N or C termini with GFP were both fluorescent and had identical subcellular localisations and functions (Wang & Hazelrigg, 1994). Also, subcellular localisation with GFP fusions are superior in plants to GUS, because of staining artefacts (Grebenok *et al*., 1997; Haseloff *et al*., 1997; Haseloff & Siemering, 1998) and are more sensitive than antibody-based techniques, due to non-specific background labelling (Wang & Hazelrigg, 1994). Above all, GFP’s intrinsic fluorescence means it (and fused proteins) can be monitored *in vivo, in situ*, in cells, tissues or whole plants, at high resolution and in real time (Cubitt *et al*., 1995; Haseloff *et al*., 1997; Haseloff, 1999; Haseloff & Siemering, 1999; Gerdes & Kaether, 1996; Kain *et al*., 1995; Leffel *et al*., 1997).

GFP was discovered in photocytes (bioluminescent cells) at the margins of the vellum of the cnidarian jellyfish *Aequorea victoria* (Hastings & Morin, 1998; Shimomura, 1998). GFP cDNA was sequenced by Prasher *et al*., (1992) and encodes a 238 residue polypeptide of approximately 27 kDa. Heterologous expression of GFP cDNA produces an intrinsically fluorescent protein, without exogenous co-factors or substrates (Chalfie *et al*., 1994), except oxygen (Heim *et al*., 1994; Inouye & Tsuji, 1994). The intrinsic fluorescence of GFP is due to its chromophore, formed by autocatalytic cyclisation,
dehydration and oxidation of the tripeptide Ser$_{65}$-dehydroTry$_{66}$-Gly$_{67}$ (Cody et al., 1993). The chromophore is located at the geometric centre of an 11 stranded β-barrel (24Å in diameter and 42Å long) on the coaxial α-helix; the β-barrel has α-helical loops at both ends and is known as the β-can (Ormö et al., 1996; Yang et al., 1996). The high photostability of GFP to pH, temperature and denaturing agents is attributed to chromophore protection from the chemical environment by the β-can (Ward, 1998). GFP is a stable, intrinsically fluorescent protein, even when expressed heterologously.

GFP can be efficiently detected in plant cells. Early attempts to express GFP in plants resulted in poor or undetectable fluorescence, unless a viral expression vector was used (Haseloff & Amos, 1995; Haseloff et al., 1997). However, Haseloff et al. (1997) showed that GFP expression in A. thaliana was curtailed due to aberrant mRNA processing with the removal from the coding sequence of a cryptic plant intron. Modification of the codon usage in the cryptic plant intron splice site substantially increased fluorescence (Haseloff et al., 1997).

Enhanced GFP (EGFP) and enhanced cyan fluorescent protein (ECFP) are variants of GFP. EGFP (CytoGem Product Manual) and ECFP (pECFP Vector Information) are based on the “humanised” sequence produced by Haas et al. (1996). The “humanised” GFP sequence is altered by more than 190 synonymous nucleotide substitutions to conform to the favoured codon usage of highly expressed human proteins, which substantially increased fluorescence intensity (Haas et al., 1996). “Humanised” GFP doesn’t have the cryptic plant intron splice site and is functional in plants (Chiu et al., 1996; Galbraith et al., 1999). EGFP has seven amino acid substitutions compared to GFP: F64L, S65T, S72A, N149K, M153T, I167T and H231L (CytoGem Product Manual) and excitation and emission maxima at 484 nm and 508 nm, respectively (CytoGem Product Manual). ECFP has six amino acid substitutions compared to GFP: F64L, S65T, Y66W, N146I, M153T and V163A and excitation and emission maxima at 433 nm and 475nm, respectively (pECFP Vector Information). F64L, S65T, V163A amino acid substitutions were reported to increase fluorescence intensity (Anderson et al., 1996; Cheng et al., 1996; Chiu et al., 1996; Cormack et al., 1996; Haas et al., 1996; Pang et al., 1996; Patterson et al., 1997; Reichel et al., 1996; Siemereng et al., 1996; Stauber et al., 1998; Zolotukhin et al., 1996). The substitution S72A occurs in mutants GFPmut2 and GFPmut3, which have increased fluorescence intensity (Cormack et al., 1996). A higher fluorescence intensity in the Cycle 3 mutant
(F99S, M153T and V163A) is probably due to its greater protein solubility, allowing correct folding and hence chromophore formation (Crameri et al., 1996). The Y66W amino acid substitution shifts the emission maxima to shorter wavelengths; and so, to a lesser extent, does I167T (Heim et al., 1994).

It is necessary to determine whether \textit{RPW8.1} and \textit{RPW8.2} cDNAs are transiently expressed in \textit{N. benthamiana} to confirm that both RPW8.1 and RPW8.2 proteins were required for \textit{RPW8}-induced, Avr-independent HR in \textit{N. benthamiana}. Transient assays, using non-viral DNA delivery methods, in tobacco leaf epidermal and mesophyll cells have been used to localise proteins fused to GFP \textit{in planta} (Baba et al., 2001; Batoko et al., 2000; Crawford & Zambryski, 2000; Hibberd et al., 1998; Helliwell et al., 2001; Itaya et al., 1997; Köhler et al., 1997; Kotlizky et al., 2000; Millen et al., 2001; Oikawa et al., 2000; Oparka et al., 1999; Takechi et al., 2000; Zhang et al., 2001). Transiently expressing \textit{RPW8.1} and \textit{RPW8.2} cDNAs fused to the 5’ end of \textit{GFP}, via \textit{Agrobacterium}-mediated transformation, in \textit{N. benthamiana} will prove that the full length proteins are transiently expressed.

\textit{RPW8.1} and \textit{RPW8.2} cDNAs fused by a linker to the 5’ end of \textit{EGFP} and \textit{ECFP} sequences (Figure IX. 1), respectively, were transiently expressed, via \textit{Agrobacterium}-mediated transformation, in \textit{N. benthamiana}. \textit{RPW8.1::EGFP} and \textit{RPW8.2::ECFP} fluorescence (and therefore expression) was detected in leaf abaxial epidermal and mesophyll cells with laser scanning confocal microscopy (LSCM) and subcellularly localised with reference to counter-fluors.

Fig. IX.1: Fused polypeptides a) \textit{RPW8.1::EGFP} and b) \textit{RPW8.2::ECFP} – not to scale, from N-terminal (Nt) to C-terminal (Ct). The black line that separates the \textit{RPW8} polypeptide region from the \textit{GFP} polypeptide region represents the (Gly-Ala)$_5$ linker.
IX.2: Methods

IX.2.1: Cloning RPW8 cDNAs with GFPs into binary vector T-DNAs

RPW8.1 and RPW8.2 cDNAs were each amplified in seventeen 20 µl PCRs, of 10 fold sterile Milli-Q H₂O diluted Ms-Pc cDNA (0.5 µl). The PCR mix contained 1.25U of Expand HiFi Taq (ROCHE) and 1.25U of Pfu Turbo™ (Stratagene™) in Expand buffer No.2 (BOEHRINGER MANNHEIM), 0.2 mM of dNTPs (Bioline), 0.2 µM of each primer (MWG Biotech AG) and was in thin-walled, 0.2 ml PCR tubes (LAZER). MC2 (AFP) NT F (5’ – GGG CGG ATG CAT ATG CCG ATT GGT GAG CTT GCG – 3’) and MC2 (AFP) CT R (5’ – CCC CCG ATG CAT TCA AGC TCT TAT TTT ACT ACA AGC – 3’) primers were used to amplify RPW8.1 cDNA. MC3 (ECFP) NTF2 (5’ – ACG GGA TCC ATG ATT GCT GAG GTT GCC GCA – 3’) and MC3 (AFP) CTR (5’ – ACG TCG CCA TGG TTG CGC CGG CTC CTG CGC CGG CTC CAG AAT CAT CAC TGC AGA ACG TA – 3’) primers were used to amplify RPW8.2 cDNA. The negative control contained no DNA. The thermal cycle conditions (Figure IX.2) were generated in a Hybaid Omnigene Thermal Cycler.

Fig. IX.2: Pfu Turbo PCR thermal cycle conditions to amplify a) RPW8.1 cDNA and b) RPW8.2 cDNA.

Each 20 µl PCR was mixed with 4 µl of 6X orange G loading dye and electrophoresed on a 1% agarose gel at 100V. PCR fragments of approximately 500 bp (size confirmed against 500 ng of 1 kb DNA ladder marker, GIBCOBRL™) were excised from the gel, over a glass plate, with an EtOH cleaned scalpel. The PCR fragments were extracted from the gel slices using the QIAGEN™ QIAEXII kit. For every 1 mg of gel, 3 µl of QX1 was added to the Eppendorf tube with 10 µl of re-suspended QIAEXII beads, vortexed for 30 seconds and incubated for 15 minutes at 50°C (inverting every 2 minutes). The Eppendorf tubes were centrifuged for 1 minute at 14,000 rpm and the supernatant was discarded. The pellet was then re-suspended in 0.5 ml of QX1 and vortexed for 30 seconds. The Eppendorf tubes were centrifuged for 1 minute at 14,000 rpm, the supernatant was discarded, then the pellet re-suspended in 0.5 ml of PE and vortexed for 30 seconds. The previous centrifugation and re-
suspension step was repeated once. Then the Eppendorf tubes were centrifuged for 1 minute at 14,000 rpm, the supernatant was discarded and the QIAEXII bead pellet was then dried by incubating the tubes at room temperature with the caps open until the pellet had just turned white. PCR fragment DNA was eluted from the dried QIAEXII bead pellet with 20 µl of 10 mM Tris·Cl (pH8.0), vortexed for 30 seconds, then incubated for 10 minutes at 50°C. The elute and beads were vortexed for 30 seconds and centrifuged for 14,000 rpm for 1 minute before removing the elutes of PCR fragments, which were combined (total of 40 µl). A 1 µl aliquot was mixed with 4 µl of sterile Milli-Q H₂O and 1 µl of 6X orange G loading dye and electrophoresed (against 500 ng of 1 kb DNA ladder marker) on a 1% agarose gel at 120V. The RPW8.1 and RPW8.1 cDNA concentrations were estimated as 40 ng µl⁻¹ and 5 ng µl⁻¹ with 500 ng of λHindIII marker ( GibcoBRL™).

The remaining 39 µl of the RPW8.1 cDNA and RPW8.2 cDNA elutes were diluted to 95 µl with sterile Milli-Q H₂O to approximately 15 ng µl⁻¹ and 2 ng µl⁻¹. Diluted RPW8.1 cDNA and 500 ng of pEGFP (CLONTECH) were independently digested by 20U of NcoI ( GibcoBRL™) and 20U of EcoRI (BOEHRINGER MANNHEIM) in REACT3 buffer (Stratagene™) for 15 hours at 37°C. The RPW8.1 cDNA and pEGFP digests were centrifuged at 14,000rpm for 60 seconds and inactivated in a water-bath at 65°C for 15 minutes. Diluted RPW8.2 cDNA and 1.3 µg of pECFP (CLONTECH) were independently digested by 40U of NcoI and 40U of BamHI (BOEHRINGER MANNHEIM) in REACT3 buffer (Stratagene™) for 10 hours at 37°C. The RPW8.2 cDNA and pECFP digests were centrifuged at 14,000 rpm for 60 seconds and inactivated in a water-bath at 85°C for 20 minutes. Inactivated digests were centrifuged at 14,000rpm for 60 seconds and then incubated at room temperature for 30 minutes. Digested pEGFP and pECFP were 5’ de-phosphorylated by adding 2U of Shrimp Alkaline Phosphatase (Amersham Life Science) to the digests and incubating for 3 hours at 37°C. Shrimp Alkaline Phosphatase was inactivated at 65°C for 15 minutes, centrifuged at 14,000rpm for 60 seconds and incubated at room temperature for 30 minutes. Aliquots of the pEGFP and pECFP
digests (2 µl) were diluted to 5 µl with sterile Milli-Q H₂O, mixed with 1 µl of 6X orange G loading dye and electrophoresed (against 500 ng of 1 kb DNA ladder and 500 ng of λHindIII markers) on a 1% agarose gel at 120V. The pEGFP and pECFP vectors had been completely digested to produce approximately 10 ng µl⁻¹ and 40 ng µl⁻¹ of linear 4.5 kb and 3.4 kb fragments, respectively.

Optimal ligation substrate concentrations (i & j values) are dependent on the ligation volume and the size and number of DNA molecules involved (Hadfield, 1986). For a 5 µl ligation i and j values, for an i = 2j reaction are approximately 150 ng µl⁻¹ (750 ng) for the 500 bp RPW8.1 or RPW8.2 cDNA inserts and 65 ngµl⁻¹ (325 ng) for 4.5 kb pEGFP or 70 ng µl⁻¹ (350ng) of 3.4 kb pECFP vectors.

Insert and vector were combined in the correct quantities and EtOH precipitated at -20°C for 2 hours after adding 1/10X volume of 3M NaAc and 2X volumes of EtOH. The precipitated DNA was centrifuged at 14,000 rpm for 15 minutes and the supernatant discarded. The DNA pellet was washed with 150 µl of 70% EtOH (v/v), re-centrifuged at 14,000 rpm for 2 minutes, and the supernatant discarded. Incubating the DNA pellet in the opened microfuge tube for 1 minute at room temperature dried the DNA before it was re-suspended in 3.5 µl of 10mM TrisCl (pH8.0) and incubated for 30 minutes at room temperature. Re-suspended DNA was mixed with 1 µl of 5X BRL ligase buffer (GIBCOBRL™) and a 0.45 µl (10%) aliquot was removed and diluted to 8 µl with sterile Milli-Q H₂O. T4 DNA ligase HC (2.25U, GIBCOBRL™) was then added to the buffered, re-suspended DNA and the ligation was incubated at 12°C for 12 hours. A 0.45 µl (10%) aliquot was removed from the ligation and diluted to 8 µl with sterile Milli-Q H₂O. Pre- and post-T4 DNA ligase diluted aliquots were mixed with 6X orange G loading dye and electrophoresed on a 2% agarose gel at 100V to confirm RPW8.1-pEGFP and RPW8.2-pECFP ligation.

RPW8.1-pEGFP and RPW8.2-pECFP ligated DNA was diluted 10 fold with TE. Aliquots (20 µl) of
Electrocompetent DH10B *E. coli* cells were transformed with 1 µl of diluted, ligated DNA in a Cell Porator® (Life Technologies™). Electroporated cells were then transferred to 1 ml of LB in 1.5 ml Eppendorf tubes and grown for 1 hour in a 37°C shaker at 250 rpm. The electrophoresed cells were centrifuged at 14,000 rpm for 5 minutes, the supernatants discarded and the pellets re-suspended in 0.4 ml of LB. A 0.2 ml aliquot of the re-suspended, electroporated cells was plated on to LB-Amp agar plates (200 µgml⁻¹ filter sterilised ampicillin, DUCHEFA) and grown at 37°C overnight.

*E. coli* colonies transformed with *RPW8.1*-pEGFP ligated DNA were picked with sterile P2 pipette tips, inoculated into 10 ml of LB-Amp, and cultured overnight in a 37°C shaker at 250 rpm. The cultures were put on ice for 1 hour and then 1 ml of each culture was centrifuged at 5,000 xg for 5 minutes at 4°C before discarding the supernatants. The pellets were re-suspended in 100 µl of ice-cold Solution A, and incubated on ice for 30 minutes. After incubation solution B (200 µl) was added and the suspensions were incubated for 5 minutes at room temperature. Ice-cold 5M KAc/AcH (150 µl) was added and the suspensions were incubated for 10 minutes on ice. Then the suspensions were centrifuged for 5 minutes at 14,000 rpm and the supernatants were transferred to Eppendorf tubes. The supernatants were mixed with 400 µl of 25:24:1 phenol:chloroform:isoamyl alcohol (SIGMA®), shaken, and centrifuged at 14,000 rpm for 2 minutes. The upper, aqueous phases were removed, mixed with 2X volumes of EtOH, and incubated at room temperature for 2 minutes. After centrifugation at 14,000 rpm for 5 minutes the supernatants were discarded and the DNA pellets were dried by incubation in opened Eppendorf tubes for 1 minute at room temperature before they were re-suspended in 50 µl of TE+RNAse (10 mg ml⁻¹ RNAse, QIAGEN™). A 5 µl portion of re-suspended plasmid (10%) was mixed with 1 µl of 6X orange G loading dye and separated on a 1% agarose gel at 120V against 500 ng of 1 kb DNA ladder marker.

Each re-suspended *RPW8.1*-pEGFP plasmid was tested for the presence of the *RPW8.1* cDNA insert by digesting 5 µl of re-suspended plasmid DNA in 20 µl reactions with 20U of *NcoI* and 20U of *EcoRI*
buffered by REACT3, for 10 hours at 37°C. The digests were mixed with 4 µl of 6X orange G loading dye and half (12 µl) was electrophoresed (against 500 ng of 1 kb DNA ladder and 500 ng of λ HindIII markers) on a 1% agarose gel at 120V. The presence of the RPW8.1 cDNA insert was confirmed by the excision of a fragment of approximately 500 bp.

E. coli colonies transformed with RPW8.2-pECFP ligated DNA were picked with sterile P2 pipette tips on to LB-Amp agar plates and then into thin-walled 0.2 ml PCR tubes. The LB-Agar plates were incubated overnight at 37°C. The 0.2 ml PCR tubes were heated for 1 minute at maximum temperature in a Panasonic microwave oven before adding the PCR mix. The PCR mix contained 1.25U of BIOTAQ (Bioline) in Mg2+ Tris buffer (BOEHRINGER MANNHEIM), 0.2 mM of dNTPs (Bioline) and 0.2 µM of MSC3 ESPF and AFP R primers (MWG Biotech AG). The negative control contained no DNA. The thermal cycle conditions (Figure IX.3) were generated in a Hybaid Omegene Thermal Cycler. The PCRs were mixed with 4 µl of 6X orange G loading dye and were electrophoresed (against 500 ng of 1 kb DNA ladder marker) on a 1% agarose gel at 120V. The presence of the 1.2 kb, \textit{RPW8.2::ECFP} fragment in PCRs confirmed ligation. In PCRs that contained the 1.2 kb fragment the corresponding colony was inoculated into 10 ml LB-Amp and grown overnight on a shaker at 37°C.

Fig. IX.3: The thermal cycle conditions for amplifying \textit{RPW8.2-pECFP} to test for the presence of ligated DNA.

Sequencing quality \textit{RPW8.1-pEGFP} or \textit{RPW8.2-pECFP} DNA was extracted using the the QIAGEN™ mini-prep kit, from either the remaining 9 ml of each 10 ml culture of the colonies transformed \textit{RPW8.1-pEGFP} ligated DNA that contained inserts or all of the 10 ml transformed with ligated \textit{RPW8.2-pECFP}. The 9 ml cultures were centrifuged in a 50 ml centrifuge tube (Nalge NUNC™) at 5,000 xg for 5 minutes at 4°C. The supernatants were discarded and the pellets were re-suspended in 0.4 ml of buffer P1. The re-suspended pellets were transferred to Eppendorf tubes, mixed gently by inverting with 0.4 ml of buffer P2 and incubated at room temperature for 5 minutes. After 5 minutes, chilled buffer P3 (0.3 ml) was added and the re-suspended pellets were incubated on ice for at least 10 minutes. Then the re-suspended pellets were centrifuged at 14,000 xg for 10 minutes and the supernatant was transferred
through a layer of miracloth (CALBIOCHEM®) to QIAGEN Tip20 columns, already equilibrated with 1 ml of buffer QBT. The columns were then washed four times with 1 ml of buffer QC. The DNA from each column was eluted to an Eppendorf tube with 0.8 ml of buffer QF and precipitated with 0.56 ml of isopropanol at room temperature. The precipitating elutes were immediately centrifuged for 30 minutes at 14,000 xg. The supernatants were discarded and the DNA pellets washed with 150 µl of 70% EtOH (v/v) before they were centrifuged for 2 minutes at 14,000 xg and the supernatants discarded. The DNA pellets were dried in the Eppendorf tubes by incubation at room temperature with their caps open for 1 minute. Dried DNA pellets were re-suspended in 20 µl of TE and incubated at room temperature for 30 minutes. A portion (1 µl) of each of the re-suspended DNAs was diluted to 5 µl with sterile Milli-Q H2O, mixed with 1 µl of 6X orange G loading dye and electrophoresed (against 500 ng of 1 kb DNA ladder and 500 ng of λHindIII markers) on a 1% agarose gel at 120V. The DNA concentrations of RPW8.1-pEGFP and RPW8.2-pECFP were both estimated at 100 ng µl^{-1}.

QIAGEN™ mini-prep RPW8.1-pEGFP DNAs (2 µl) were sequenced in four 10 µl reactions with 1.6pmol of each of the four primers and with 2 µl of PR mix (ABI). The four primers (MWG Biotech AG) used were MSC2 ESPF (5’ – GGG CGG ATG CAT ATG CCG ATT GGT GAG CTT GCG – 3’), MC2 (g) Seq RNt (5’ – GCT TAT CGA TTT GAC CAA AAA CG – 3’), MC2 (g) Seq FCt (CGA TAA TCT CCA AAC ACT AAA CG – 3’) and MSC2 ESPR (5’ – CCC CCG ATG CAT TCA AGC TCT TAT TTT ACT ACA AGC – 3’). QIAGEN™ mini-prep RPW8.2-pECFP DNAs (2 µl) were sequenced in five 10 µl reactions with 1.6pmol of each of the five primers and with 2 µl of PR mix (ABI). The five primers (MWG Biotech AG) used were MSC3 ESPF (5’ – GGG CCG ATG CAT ATG ATT GCT GAG GTT GCC GCA – 3’), MC3 (g) Seq RNt (5’ – CGG TGT GAT ACT ATC GAT TGT AGC – 3’), MC3 Seq F (5’ – GAA CGC TGT TTC TCT TGT TGA – 3’), MSC3 Seq F3Ct (5’ – GAT ATC AAA GAG TTC GAA GC – 3’) and MSC3 ESPR (5’ – CCC CCG ATG CAT TCA AGA ATC ATC ACT GCA GAA CGT – 3’). The thermal cycle conditions used to sequence RPW8.1-pEGFP and RPW8.2-pECFP are shown in Figure IX.4.
Sequencing reactions were EtOH precipitated by vortexing the reactions with 8 µl of sterile Milli-Q H$_2$O and 32 µl of 95% EtOH (v/v) and then incubated at room temperature for at least 2 hours. The precipitated sequencing reactions were then centrifuged at 14,000 xg for 20 minutes, the supernatant was discarded and the DNA pellet was washed in and vortexed with 250 µl of 70% EtOH (v/v). Washed DNA pellets were centrifuged at 14,000 xg for 10 minutes, the supernatants discarded and then the DNA pellets were dried by heating the tubes, with their lids open, on a block at 90°C for 1 minute. Dried sequencing DNA was then sent for processing to the Bioanalytical Laboratory (Department of Biological Sciences, UEA). Sequence chromatogram files in UNIX format were converted to text and compared to the cDNA sequences of $RPW8.1$ for $RPW8.1$-pEGFP DNAs and of $RPW8.2$ for $RPW8.2$-pECFP, using ted and xgap programs.

*Fig. IX.4: The thermal cycle conditions used to sequence $RPW8.1$-pEGFP and $RPW8.2$-pECFP with the ABI Big Dye Terminator$^\text{TM}$ system.*

$RPW8.1$::EGFP and $RPW8.2$::ECFP were amplified from 0.5 µl of an $RPW8.1$-pEGFP DNA and an $RPW8.2$-pECFP DNA, respectively, diluted 10 fold with sterile Milli-Q H$_2$O, in seventeen 20 µl PCRs.

The PCR mix contained 1.25U of Expand HiFi Taq and 1.25U of Pfu Turbo$^\text{TM}$ in Expand buffer No.2, 0.2 mM of dNTPs (Bioline), 0.2 µM of each primer and was in thin-walled, 0.2 ml PCR tubes (LAZER). MC2 (AFP) NT F (5’ – ACG GAA TTC ATG CCG ATT GGT GAG CTT GCG – 3’) and AFP R (5’ – ACG GGA TCC TTA CTT GTA CAG CTC G – 3’) primers were used to amplify $RPW8.1$::EGFP. MC3 (AFP) NTF and AFP R were used to amplify $RPW8.2$::ECFP. The negative control contained no DNA. The thermal cycle conditions (Figure IX.5) were generated in a Hybaid Omnimite Thermal Cycler.

*Fig. IX.5: PCR thermal cycle conditions used to amplify $RPW8.1$::EGFP and $RPW8.2$::ECFP from $RPW8.1$-pEGFP and $RPW8.2$-pECFP templates, respectively.*

The PCRs were electrophoresed and then PCR fragments of approximately 1.2 kb were excised from the gel and extracted as above. The remaining elute of $RPW8.1$::EGFP, $RPW8.2$::ECFP and 1280 ng
of pBluescript were independently digested by 20U of BamHI and 20U of EcoRI in REACT3 buffer (Stratagene™) for 10 hours at 37⁰C. The digests were centrifuged at 14,000 rpm for 60 seconds and inactivated in a water-bath at 85⁰C for 15 minutes. Inactivated digests were centrifuged at 14,000 rpm for 60 seconds and then incubated at room temperature for 30 minutes. Digested pBluescript was 5' de-phosphorylated by Shrimp Alkaline Phosphatase and electrophoresed as above. The pBluescript had been completely digested to produce approximately 10 ng µl⁻¹ of a linear 3 kb fragment.

RPW8.1::EGFP and RPW8.2::ECFP concentrations were both estimated at 20 ng µl⁻¹.

For a 5 µl ligation, i and j values (for an i = 2j reaction) are approximately 150 ng µl⁻¹ (650 ng) for the 1.2 kb RPW8.1::EGFP insert and 76 ng µl⁻¹ (380 ng) for the 3 kb pBluescript vector. RPW8.1::EGFP and RPW8.2::ECFP were independently ligated with pBluescript as above. RPW8.1::EGFP-pBluescript and RPW8.2::ECFP-pBluescript ligated DNAs were diluted 10 fold with TE and 1 µl aliquots were electroporated into electrocompetent DH10B E. coli as above. RPW8.1::EGFP-pBluescript transformed E. coli were tested for the presence of the RPW8.1::EGFP insert as above, except that the plasmid DNA was digested by 20U of BamHI and 20U of EcoRI. RPW8.2::ECFP-pBluescript transformed E. coli were tested for ligated RPW8.2::ECFP-pBluescript by picking colonies with sterile P2 pipette tips on to LB-Amp agar plates and then into thin-walled 0.2 ml PCR tubes. The LB-Amp-Agar plates were incubated overnight at 37⁰C. The 0.2 ml PCR tubes were heated for 1 minute at maximum temperature in a Panasonic microwave oven before adding the PCR mix. The PCR mix contained 1.25U of BIOTAQ in Mg²⁺ Tris buffer, 0.2 mM of dNTPs and 0.2 µM of MSC3 (AFP) NT F and T3 primers. The negative control contained no DNA. The thermal cycle conditions (Figure IX.6) were generated in a Hybaid Omnigene Thermal Cycler. The PCRs were mixed with 4 µl of 6X orange G loading dye and were electrophoresed (against 500 ng of 1 kb DNA ladder marker) on a 1% agarose gel at 120V. The presence of the 1.2 kb, RPW8.2::ECFP fragment in PCRs confirmed ligation. In PCRs which contained the 1.2 kb fragment the corresponding colony was inoculated into 10 ml LB-Amp and grown overnight on a shaker at 37⁰C.
Fig. IX.6: PCR thermal cycle conditions for amplifying RPW8.2::ECFP across the ligation site with pBluescript on the RPW8.2::ECFP-pBluescript template to test for correct ligation.

Sequencing quality RPW8.1::EGFP-pBluescript DNA was extracted from the remaining 9 ml of RPW8.1::EGFP-pBluescript culture as above and estimated at 40 ng µl⁻¹. RPW8.2::ECFP-pBluescript DNA was extracted from the 10 ml RPW8.2::ECFP-pBluescript culture as above and estimated at 20 ng µl⁻¹. QIAGEN™ mini-prep RPW8.1::EGFP-pBluescript plasmid DNA (2 µl) was sequenced in eight 10 µl reactions with 1.6pmol of each of the eight primers and with 2 µl of PR mix (ABI). The primers used were M13 AF (5’ – CCC AGT CAC GAC GTG GTA AAA CG – 3’), MSC2 ESPF, MSC2 Seq RNT, MC2 Seq F (5’ – GAG GCT TAT CCG AAA CTC AGA – 3’), MSC2 (g) Seq F3Ct (5’ – CGA TAG AAC TTG CAT TAA GAA GC – 3’), MSC2 ESPR, AFP R and M13 AR (5’) – AGC GGA TAA CCA TTT CAC ACA GG – 3’). QIAGEN™ mini-prep RPW8.2::ECFP-pBluescript plasmid DNA (0.5 µl) was sequenced in six 10 µl reactions with 1.6pmol of each of the six primers and with 2 µl of PR mix (ABI). The primers used were MSC3 ESPF (5’ – GGG CCG ATG CAT ATG ATT GCT GAG GTT GCC GCA – 3’), MSC3 Seq RNT, MC3 Seq F (5’ – GAA CGC TGT TTC TTC TCT TGT TGA – 3’), MC3 Seq F2Ct (5’ – GCC TCA ACC GAA GTT TG – 3’), MSC3 ESPR and AFP R. The thermal cycle conditions used to sequence RPW8.1::EGFP-pBluescript and RPW8.2::ECFP-pBluescript are shown in Figure IX.4. The sequencing reactions were EtOH precipitated and processed as above. Sequence chromatogram files in UNIX format were converted to text and compared to the cDNA sequences of RPW8.1, RPW8.2 and EGFP, using ted and xgap programs.

Sequences confirmed as RPW8.1::EGFP-pBluescript and RPW8.2::ECFP-pBluescript plasmid DNAs were electroporated into DH10B E. coli as above. Colonies were picked from the transformation plate with sterile P2 pipette tips to thin-walled 0.2 ml PCR tubes and then inoculated into 10 ml of LB-Amp and grown in a 37°C shaker overnight. The 20 µl PCR mix contained 1.25U of BIOTAQ in Mg²⁺ Tris buffer, 0.2 mM of dNTPs and with 0.2 µM of AFP R primer and MSC2 ESPF (for RPW8.1::EGFP) and MSC3 ESPF (for RPW8.2::ECFP) primers. The negative control contained no DNA and the
thermal cycle conditions (Figure IX.7) were generated in a Hybaid Omnigene Thermal Cycler. For colonies confirmed by PCR to contain the \textit{RPW8.1::EGFP} insert of \textit{RPW8.1::EGFP}-pBluescript or the \textit{RPW8.2::ECFP} insert of \textit{RPW8.2::ECFP}-pBluescript, plasmid DNA was extracted from 1 ml of their culture as above. Extracted plasmid DNAs were digested with 20U of EcoRI and 20U of BamHI in REACT3 overnight at 37°C. The digests were electrophoresed, excised from the gel and then gel extracted as above and the concentration of \textit{RPW8.1::EGFP} and \textit{RPW8.2::ECFP} were both estimated as 20 ng µl⁻¹.

Fig. IX.7: PCR thermal cycle conditions used to confirm the transformation of E. coli with \textit{RPW8.1::EGFP}-pBluescript or \textit{RPW8.2::ECFP}-pBluescript.

For a 5 µl ligation, i and j values (for an i = 2j reaction) are approximately 150 ng µl⁻¹ (650 ng) for the 1.2 kb \textit{RPW8.1::EGFP} or \textit{RPW8.2::ECFP} inserts and 60 ng µl⁻¹ (300 ng) for the 24.5 kb pKMB binary vector. \textit{RPW8.1::EGFP} and \textit{RPW8.2::ECFP} were ligated with pKMB as above. \textit{RPW8.1::EGFP}-pKMB and \textit{RPW8.2::ECFP}-pKMB ligated DNA was diluted 10 fold with TE and electroporated into electrocompetent DH10B E. coli as above, except the electroporated E. coli were plated on to LB-Tet agar plates. The insertion of \textit{RPW8.1::EGFP} and \textit{RPW8.2::ECFP} into pKMB was confirmed with PCR as above, where colonies were picked with sterile P2 pipette tips to thin-walled 0.2 ml PCR tubes and in parallel to fresh LB-Tet agar plates. The PCR primers were AFP R and 35S A, the positive controls were \textit{RPW8.1::EGFP}-pKMB and \textit{RPW8.2::ECFP}-pKMB ligated DNAs and the thermal cycle conditions are shown in Figure IX.8.

Fig. IX.8: The thermal cycle conditions to confirm \textit{RPW8.1::EGFP}-pKMB and \textit{RPW8.2::ECFP}-pKMB ligation.

The colonies confirmed by PCR to contain ligated \textit{RPW8.1::EGFP}-pKMB and \textit{RPW8.2::ECFP}-pKMB were inoculated into 10 ml of LB-Tet, and grown overnight in a 37°C shaker at 250rpm. \textit{RPW8.1::EGFP}-pKMB and \textit{RPW8.2::ECFP}-pKMB DNAs were then extracted using the QIAGEN™ plasmid mini-prep kit as above and 2 µl aliquots were digested in two 20 µl overnight reactions at
37°C: a) with 20U of *Bam*HI and 20U of *Eco*RI in REACT3 buffer, and b) with 20U of *Pst*I (BOEHRINGER MANNHEIM) in REACT2 buffer. The digests were electrophoresed on a 1% agarose gel at 120V and ligated *RPW8.1::EGFP*-pKMB and *RPW8.2::ECFP*-pKMB were identified by their insensitivity to *Pst*I in digest b) and a 1.2 kb, *RPW8.1::EGFP* or *RPW8.2::ECFP*, fragment produced by digest a).

**IX.2.2: Transient expression of *RPW8.1::EGFP* and *RPW8.2::ECFP* and detection of EGFP and ECFP fluorescence in *Nicotiana benthamiana*.

Electrocompetent *Agrobacterium tumefasciens* cells were electroporated with 1 µl of *RPW8.1::EGFP*-pKMB or *RPW8.2::ECFP*-pKMB DNAs as in VII.2.1. After 3 days at 30°C colonies were picked to fresh Agro-LB-Rif/Gent/Tet-Agar plates and simultaneously to thin-walled, 0.2 ml PCR tubes. The Agro-LB-Rif/Gent/Tet -Agar plates were incubated overnight at 37°C. The 0.2 ml PCR tubes were heated for 1 minute at maximum temperature in a Panasonic microwave oven before adding the PCR mix. The PCR mix and thermal cycle conditions were the same as those used to confirm *RPW8.1::EGFP* or *RPW8.2::ECFP* ligation with pKMB. *A. tumefasciens* colonies, confirmed by PCR to be transformed with either *RPW8.1::EGFP*-pKMB or *RPW8.2::ECFP*-pKMB, were inoculated into 10 ml Agro-LB-Rif/Gent/Tet and incubated on a shaker at 200 rpm for 3 days at 30°C. Glycerol stocks of the cultures containing the binary vectors were prepared by mixing 0.8 ml of culture with 0.2 ml of sterile glycerol in 1 ml cryovials, snap freezing them in liquid N₂ and then they were stored at -80°C. *A. tumefasciens* glycerol stocks of pKMB, *RPW8.1::EGFP*-
pKMB and RPW8.2::ECFP–pKMB were inoculated into 10 ml Agro-LB-Rif/Gent/Tet cultures and grown on a shaker for 3 days at 30°C. Portions (0.25 ml) of these cultures were inoculated into 5 ml of YEB-Gent/Tet and grown with agitation overnight at 30°C. The cultures were transferred to 15 ml centrifuge tubes (FALCON™) and centrifuged at 4,000 rpm for 10 minutes at 4°C. The supernatants were discarded and then the pellets were re-suspended in 1 ml of infiltration media before they were re-centrifuged at 4,000 rpm for 10 minutes at 4°C. The supernatants were discarded and the pellets were re-suspended in infiltration media to an OD$_{600}$ = 1.5. Re-suspended A. tumefaciens were agro-infiltrated into N. benthamiana leaves as in VII.2.1.

At 2 days post agro-infiltration, 0.5 cm$^2$ leaf sections from agro-infiltrated areas were mounted in dI.H$_2$O on microscope slides, under coverslips, with the abaxial surfaces uppermost. Some sections, as indicated, were stained with propidium iodide (SIGMA®, N-(3-triethylammoniumpropyl)-4-(6-(4-(diethylamino)phenyl)hexatrienyl) pyridinium dibromide (FM 4-64, Molecular Probes) or MitoTracker Orange (Molecular Probes). EGFP and ECFP fluorescences were detected on a Zeiss LSM570 by Dr I. Moore (Department of Plant Sciences, Oxford University). EGFP and ECFP were excited with the 458 nm line of an argon laser. FM 4-64 and MitoTracker Orange were excited with the 543 nm line of a HeNe laser. A 458/514/543 dichroic mirror was used to reflect light onto the samples. The EGFP and ECFP emission was split from FM 4-64 and MitoTracker Orange counter-fluors fluorescence with a 515 dichroic mirror and collected with a 475-525 nm band pass filter. The emission from FM 4-64 and MitoTracker Orange counter-fluors and chlorophyll fluorescence was separated from EGFP and ECFP emission using a 570 nm dichroic mirror and collected using either 650 nm long-pass or 585-615 band-pass filters. Each line in the image was scanned sequentially with 458 and 543 excitation light using the multitracking mode of the LSM570 to distinguish between the fluorescences of EGFP or ECFP and FM 4-64 or MitoTracker Orange.

IX.3: Results & Discussion
The results of this experiment are presented as micrographs. Two of the micrographs are frames taken from time-lapse micrograph sequences, which are also included as data on the attached compact disc.

IX.3.1: Subcellular localisation of RPW8.1::EGFP

IX.3.1.1: RPW8.1::EGFP was cytoplasmic and nucleoplasmic

RPW8.1::EGFP was present in the cytoplasm and nucleoplasm, with “bright-spots” (BSs), in *N. benthamiana* leaf abaxial epidermal cells, but excluded from the nucleolus which, like the cell wall was negatively stained (Figures IX.9 & IX.10). This was consistent with the predictions in III. In time-lapse micrograph movies, the extension and retraction of trans-vacuolar strands and cytoplasmic streaming of BSs (Movie 1) was evident. Cytoplasmic streaming indicated that the cells were alive (Haseloff & Siemering, 1998) when RPW8.1::EGFP fluorescence was detected. RPW8.1::EGFP fluorescence was absent from leaves transformed with the vector control (data not shown).

The cytoplasmic localisation of RPW8.1::EGFP fluorescence was confirmed by staining with FM 4-64. FM 4-64 is a plasma membrane stain that is eventually internalised by endocytosis (Vida & Emr, 1995). Aqueous FM 4-64 is non-fluorescent (Molecular Probes Handbook). Parton *et al.* (2000) stained fern rhizoids with FM 4-64; the stain was immediately detectable at the plasma membrane and after 5-15 minutes was diffusely present in the cytoplasm. Figure IX.11 shows that, after 25 minutes, FM 4-64 had internally stained the cytoplasm and trans-vacuolar strands (which co-localised with RPW8.1::EGFP fluorescence in the cytoplasm) but not the BSs. Therefore, RPW8.1 was definitely present in the cytoplasm.

The subcellular localisation of RPW8.1::EGFP was identical to GFP expressed in plant cells. The fluorescence of GFP, expressed without a fused protein or targeting sequences, accumulates diffusely throughout the cytoplasm, nucleoplasm and in trans-vacuolar strands, but is excluded from vacuoles, organelles and the nucleolus (Chiu *et al.*, 1996; Gerdes & Kaether, 1996; Grebenok *et al.*, 1997; Haseloff & Amos, 1995; Haseloff *et al.*, 1997; Haseloff, 1999; Haseloff & Siemering, 1999). Grebenok *et al.* (1997) showed that GFP enters the nucleus because it is small enough to passively diffuse through the pores; GFP fused to GUS (68 kDa) was excluded from the nucleus. Therefore
RPW8.1 (17 kDa) resided in the cytoplasm and, because it did not contain any targeting sequences and due to its size, it also entered the nucleus by passive diffusion.

IX.3.1.2: Bright-spots

The BSs were sites of locally greater fluorescence intensity. GFP fluorescence intensity – driven by a CaMV 35S promoter in tobacco – was proportional to the amount of protein (Blumenthal et al., 1999). Hence, for locally greater fluorescence intensity to form, a higher local concentration of GFP is required. Protein concentration can increase locally if the protein is sequestered in a subcellular compartment (Tsien, 1998) or because the protein aggregates. The BSs were too large to be COPI, COPII or clathrin coated vesicles; at 50-70 nm these are all below the resolution of the LSCM (I. Moore, pers. comm.). It is more likely that the BSs were organelles or aggregates.

BSs did not co-localise with mitochondria or chloroplasts. Chlorophyll fluorescence and MitoTracker Orange were used as counter-fluors to identify the chloroplasts and mitochondria. Chlorophyll fluorescence (Hibberd et al., 1998; Millen et al., 2001; Takechi et al., 2000) and MitoTracker (Duby et al., 2001; Köhler et al., 1997; Rubino et al., 2001) counter-fluors have been used to determine the subcellular localisation of GFP in planta. Chlorophyll fluorescence, detected by LSCM, corresponds to the distribution of photosystem II in higher plant-like green algae (Gunning & Schwartz, 1999) and functional chloroplasts in variegated Arabidopsis.
tissue (Takechi et al., 2000). MitoTracker counter-fluor stains are cell-permeant probes that enter the mitochondria and, once inside, are converted to a fluorescent and retained form (Molecular Probes Product Information, 2001), staining mitochondria and their movements in living plant cells (Köhler et al., 1997). In Figure IX.112 and Movie 2 the BSs did not co-localise with chloroplasts or mitochondria and were independently mobile during cytoplasmic streaming.

RPW8.1::EGFP could aggregate, by self-association, to form BSs. GFP was crystallized as a dimer by Yang et al. (1996) and does tend to dimerise at high concentrations (Cubitt et al., 1995; Palm et al., 1997; Phillips, 1998). GFP is monomeric at less than 1 mg ml\(^{-1}\), but in living cells the concentration can be greater than 5 mg ml\(^{-1}\) (Ward, 1998). EGFP formed dimers and oligomers in the secretory pathway of endocrine cells (Jain et al., 2001). However, Jain et al. (2001) demonstrated that oligomerisation was dependent on the formation of di-sulphide bridges, which are catalysed by an enzyme that resides in the ER (Alberts et al., 1994) and not the cytoplasm. Yet, BSs were not seen in plant (Itaya et al., 1997) or yeast (Chen & Ahlquist, 2000) cells containing unfused GFP, implying that GFP dimerisation alone could not cause BSs. For instance, residues 151-180 and 198-251 of the Cauliflower mosaic virus (CaMV) MP fused to GFP (Huang et al., 2001) and residues between 100-161 of the Brome mosaic virus (BMV) RNA polymerase 2a (Chen & Ahlquist, 2000) were necessary for “fluorescent foci” and “fluorescent spot” formation, respectively. RPW8.1 has putative CC motifs (Xiao et al., 2001; III.1); domains capable of homo-dimerisation. \textit{AtSERK1} contains LZ motifs – a form of CC (I.2.1.4) - and has been shown to homo-dimerise at BSs in planta by fluorescence resonance energy transfer (FRET) between \textit{AtSERK1} CFP and yellow fluorescent protein fusion proteins (Shah et al., 2001). Removal of the \textit{AtSERK1} LZ motifs abolished FRET and BS formation, proving that \textit{AtSERK1} dimerisation through the LZ motifs was responsible for BS formation. Hence, RPW8.1::EGFP probably aggregated, primarily by RPW8.1 self-association, through the CC motifs to form the BSs.

BSs have been reported in the cytoplasm of plant cells containing viral movement protein (MP) fusions to GFP. Plant cells containing MP::GFP fusions had “areas of bright fluorescence” (Kotlizky et al., 2000), “punctate, highly localised fluorescence” and “fluorescent bodies” (Padgett et al., 1996), “fluorescent dots” (Itaya et al., 1997), fluorescent “spherical punctate structures” or flecks (Heinlein et
Epel et al. (1996), Heinlein et al. (1998), Itaya et al. (1997), Kotlizky et al. (2000) and Padgett et al. (1996) assumed that, because MPs associate with plasmodesmata, the BSs were fusion protein aggregations at the plasmodesmata. The two most convincing pieces of evidence for MP::GFP BSs forming in plasmodesmata pit fields are that a) BSs were seen in the subapical cell walls of trichromes, but not in the lateral cell walls and b) the retention of BSs in washed cell walls (Padgett et al., 1996). Therefore MP::GFP BSs are probably located at plasmodesmata.

MP::GFP fusion proteins also form fluorescent filaments which protrude from the surface of protoplasts. Fluorescent “tubules” of CaMV MP fused to GFP initiate at the “fluorescent foci” and extend (protruding) from the surface of A. thaliana leaf protoplasts (Huang et al., 2000). TMV MP fused to GFP produced fluorescent “protrusions” at the surface of N. benthamiana BY2 protoplasts (Heinlein et al., 1998). Fluorescent “protrusions” were inhibited by brefeldin A (an inhibitor of secretion), but not by cytoskeleton assembly inhibitors (Heinlein et al., 1998). Hence, MP::GFP fluorescent filament development (probably by oligomerisation, initiated by aggregation at BSs) is dependent on secretion but not the cytoskeleton.

CaMV MP::GFP “fluorescent foci” are associated with the Golgi. MPI7 is an A. thaliana protein that interacts with the CaMV MP (Huang et al., 2001). MPI7::YFP and CaMV MP fused to CFP co-localise at the “fluorescent foci”, where FRET detected that the two proteins interacted (Huang et al., 2001). Huang et al. (2001) claimed that MPI7::GFP (which formed “fluorescent foci” at the cell periphery and was mobile during time-lapse microscopy) was present in reticular structures and co-localised with the distribution of GFP labelled Golgi. GFP labelled Golgi in tobacco appear as BSs that were attached to the cortical ER network and were dependent on cytoplasmic streaming for mobility (Boevink et al., 1998; Nebenführ et al., 1999). Despite CaMV MP::GFP co-localising with the Golgi, neither brefeldin A or inhibitors of cytoskeleton assembly had an effect on “fluorescent foci” formation (Huang et al., 2000). The CaMV MP was cytoplasmic and its BS formation was insensitive to inhibitors, which suggests it had a lose association with the cytoplasmic face of the Golgi.
GFP fusion protein BSs were associated with other subcellular structures which are not consistent with the plasmodesmata or Golgi examples above. Heinlein et al. (1998) found that the Tobacco mosaic virus (TMV) MP, fused to GFP, was localised to the cortical ER. Treatment with brefeldin A or cytoskeleton assembly inhibitors prevented the formation of TMV MP::GFP fluorescent “spherical punctate structures” (Heinlein et al., 1998). In yeast Chen & Ahlquist (2000) found that the “punctate spots” of BMV RNA polymerase 2a fluorescence were not associated with the Golgi or the ER. Also, BSs can be associated with two separate structures simultaneously. For example, BSs of Ob Tobamovirus MP fused to GFP were associated with the plasmodesmata as “punctate, highly localised” fluorescence and in cytoplasmically streaming “fluorescent bodies” (Padgett et al., 1996). Therefore BSs associate with several different subcellular structures, possibly simultaneously.

BS formation is not consistent with a particular subcellular structure. In plants BSs form: a) at plasmodesmata, b) are associated with the cortical ER – formation sensitive to brefeldin A and cytoskeleton assembly inhibitors – and c) are associated with the Golgi – formation insensitive to brefeldin A or cytoskeleton assembly inhibitors. Therefore BSs are phenomena which do not appear to be compartmentalised and can be formed in association with several different subcellular structures. However, the association of GFP fusion proteins to form fluorescent filaments implies that aggregation is a significant feature of BS formation. I believe that RPW8.2::EGFP BSs were formed by aggregation; primarily due to homo-dimerisation through the CCs.

IX.3.2: Subcellular localisation of RPW8.2::ECFP

IX.3.2.1: RPW8.2::ECFP was probably extracellular

RPW8.2::ECFP fluorescence in N. benthamiana leaves was found in irregular shapes (Figure IX.13). Irregular fluorescent shapes were not seen in leaves transformed with the vector control. The irregular shapes were not consistent with any known plant cell structural features. Leaf sections were stained with propidium iodide to aid identification of the irregular shapes. Propidium iodide is excluded from living cells (Yeh et al., 1981), staining cell walls and only entering dead plant cells (Haseloff & Siemering, 1998; van den Berg et al., 1995). Propidium iodide staining (Figure IX.14B) co-localises
with RPW8.2::ECFP fluorescence (Figure IX.14A) in Figure IX.14C. Therefore RPW8.2::ECFP fluorescence co-localises with either the cell walls or dead cells. It is possible that the *N. benthamiana* leaf cells in Figure IX.14 are dead, although chlorophyll fluorescence was detected in Figure IX.13 and cytoplasmic streaming was seen in Movies 1 & 2. Staining with propidium iodide implies that RPW8.2::ECFP was present in *N. benthamiana* cell walls was apoplastic and therefore must be secreted. This is consistent with a predicted SP in the RPW8.2 predicted polypeptide (III.3).

The co-localisation of RPW8.2::ECFP fluorescence and propidium iodide staining does not conclusively prove that RPW8.2 is extracellular. The CaMV 35S promoter is active in *A. tumefaciens* (Vancanneyt *et al.*, 1990) resulting in bacterial expression of RPW8.1::EGFP and RPW8.2::ECFP and hence fluorescence (data not shown). Infiltrated agrobacteria reside in the intercellular spaces of tobacco leaves and, in theory, could contribute to the fluorescent signal. However, bacterial fluorescence was not seen in the intercellular spaces of *N. benthamiana* leaves transiently expressing RPW8.1::EGFP (Figures IX.9, IX.10 & IX.11; Movie 1). Therefore, the existence of bacterial fluorescence (albeit almost certainly insignificant) and the possibility that propidium iodide has stained dead cells in Figure IX.14 prevents this technique from providing conclusive proof that RPW8.2 is extracellular.

**IX.3.2.2: Conclusively proving whether RPW8.2 is extracellular**

Three experiments could prove whether RPW8.2 is extracellular. Preventing RPW8.2::ECFP secretion could prove that RPW8.2 is extracellular. Boevink *et al.* (1999) inhibited the secretion of ER targeted GFP by cold shock and brefeldin A. However, inhibition is difficult to control and includes sample variation. More conclusive proofs would be to either a) remove the predicted 25 residue, N-terminal SP or b) to introduce a plant ER retention motif. Expression of a sequence encoding mature, genomic RPW8.2 (without the SP) would be expected to produce a cytoplasmically localised protein. Expression of RPW8.2 fused to HDEL, KDEL or RDEL – ER retention motifs (Denecke *et al.*, 1992) – at the C terminus should produce and ER localised protein. Therefore if mature RPW8.2::ECFP or RPW8.2::ECFP-HDEL were to be present in the cytoplasm and nucleoplasm, or ER, respectively, of *N. benthamiana* leaf abaxial epidermal cells it would confirm that RPW8.2 has a SP, is targeted to the ER
and subsequently the apoplast.

RPW8.2 function in the intercellular spaces of *N. benthamiana* leaves would prove that RPW8.2 is extracellular. If a solution of purified mature RPW8.2 was infiltrated into leaves already transiently expressing RPW8.1 cDNA and it induced an Avr-independent HR, then RPW8.2 would have been functional in the apoplast and RPW8.1 and RPW8.2 proteins would be necessary and sufficient.

**IX.3.3: Validating RPW8.1 and RPW8.2 subcellular localisations**

RPW8.1::EGFP and RPW8.2::ECFP proteins must be shown to confer powdery mildew resistance in *A. thaliana* accession Col-0 to prove that these GFP fusion proteins are functional. If RPW8.1::EGFP and RPW8.2::ECFP proteins are functional then it is reasonable to assume that they are correctly localised. *A. thaliana* accession Col-0 plants have been transformed with RPW8.1::EGFP-pKMB and RPW8.2::ECFP-pKMB by S. Nettleship; the T₃ generation will be scored for resistance to *E. cichoracerum* isolate UCSC1 (S. Nettleship pers. comm.).

The subcellular localisations of RPW8.1 and RPW8.2 proteins are currently being determined using an alternative technique. S. Nettleship is attempting to express and subcellularly localise HA epitope tag C-terminal fusions of RPW8.1 or RPW8.2 in *A. thaliana* accession Col-0 suspension cell cultures. If the results of this experiment are consistent with the subcellular localisations of RPW8.1::EGFP and RPW8.2::ECFP then the tentative conclusions of this chapter will be strengthened.

**General discussion**

**X.1: The functions of RPW8 proteins**

RPW8.1 and RPW8.2 were both necessary, but probably not sufficient, for the RPW8-induced, Avr-
independent HR in *N. benthamiana* (VII.3). This *RPW8*-induced HR might also require a third component, Factor X, possibly encoded by *RPW8fx*. Whether RPW8.1 and RPW8.2 require Factor X to generate an *RPW8*-induced, Avr-independent HR in *N. benthamiana* is currently being assessed.

RPW8.1 was cytoplasmic and nucleoplasmic (IX.3). RPW8.2 was probably extracellular (IX.3). The different subcellular localisations of RPW8.1 and RPW8.2 could be important for different aspects of *RPW8* function. The site of elicitor recognition and the subcellular localisation of R proteins that initiate active defence responses against biotrophic fungal pathogens is expected to be at the apoplastic face of the plasma membrane (I.3.1). Therefore RPW8.2 appeared to reside at the expected site of elicitor recognition for powdery mildew pathogens and presumably is a component of the *RPW8* recognition complex. RPW8.1 was not localised to the expected site of elicitor recognition and, because it was in the cytoplasm and nucleoplasm, probably initiates defence responses. If RPW8.1 initiates defence responses and RPW8.2 is part of the *RPW8* recognition complex then RPW8.2 probably specifically activates RPW8.1 in the presence of the elicitor, AvrRpw8. Hence, RPW8.1 is probably activated by RPW8.2 via a signal that can cross the plasma membrane; a possible function for Factor X.

Structural analysis of the RPW8.1 and RPW8.2 predicted polypeptides has identified amino acid polymorphisms T64S and D116G in RPW8.2 and S40I in RPW8.1 as candidate structural features that could be responsible for susceptibility (VI.3.2). The function of these structural features (and others) should be assessed by transient expression in *N. benthamiana* and stable transformation of *A. thaliana* ecotype Col-0; using the transient expression system as an initial screen. However, before the transient expression system can be used, it is necessary to determine whether RPW8.1 and RPW8.2 are sufficient for *RPW8*-induced, Avr-independent HR in *N. benthamiana*.

**X.2: The structures of RPW8 proteins**

The structures of RPW8.1 and RPW8.2 are needed to fully understand the functional significance of RPW8.1 and RPW8.2 site-directed-mutagenesis data. Without structures it is difficult to comprehend how particular amino acid substitutions influence the function of RPW8 proteins. Also, as structurally unique proteins, it would be interesting to see what the structures of RPW8.1 and RPW8.2 are and if
The *E. coli* and *S. pombe* heterologous expression systems were not suitable for producing sufficient RPW8.1 or RPW8.2 protein to determine their structures (IV & V). Therefore the best option for extraction and purification of RPW8.1 and RPW8.2 proteins is to produce them in planta, *i.e.* using the transient expression system, as discussed in VIII.4. An advantage of extracting and purifying RPW8.1 and RPW8.2 from plant systems is that the protein used for structural determination can also be tested functionally.

Protein structure can be determined using either nuclear magnetic resonance (NMR) or X-Ray crystallography (Alberts *et al*., 1994). NMR has certain advantages over X-Ray crystallography: it can determine the structure and dynamics of hydrated proteins in solution and under physiological conditions, without the time-consuming process of growing protein crystals (Campbell & Sheard, 1987; Kay, 1998; Jefson, 1988; Wüthrich, 1990; Wüthrich, 1991; Wüthrich, 1998). Comparisons of protein structures determined by both NMR and X-Ray crystallography agree well, although there are differences (Clore & Gronenborn, 1987; Jefson, 1988; Wüthrich, 1991). The disadvantages of NMR are that it is restricted to smaller proteins and needs highly concentrated protein solutions of a few mM (Clore & Gronenborn, 1987; Wüthrich, 1990; Wüthrich, 1991). RPW8.1 and RPW8.2 are 17 kDa and 20 kDa, respectively. This is smaller than the size limit for NMR (which is approximately 30 kDa) providing they are isotopically labelled with $^{13}$C or $^{15}$N (Wüthrich, 1998). If a 5-10 mg, highly homogeneous solution of RPW8.1 or RPW8.2 pure protein was extracted from *N. benthamiana* or *A. thaliana* then their structures could be determined in a 0.5 ml sample of a 0.5-1.0mM solution (K. Wüthrich, pers. comm.). Hence, providing RPW8.1 and RPW8.2 are sufficiently soluble, do not aggregate and can be isotopically labelled and extracted from plants, their structures could be determined by $^{13}$C or $^{15}$N-NMR.

**References**


Baulcombe, D.C. (1999) Fast forward genetics based on virus-induced gene silencing. Current...
Opinion in Plant Biology 2: 109-113


Boyes, D.C., Jaensung, N., 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. Proceedings of the National Academy of Sciences-USA **95**: 15849-15854


recessive RRS1-R gene, a member of a novel family of resistance genes. Proceedings of the National Academy of Sciences 99 (4): 2404-2409


ESPTM Yeast protein expression and purification system. Instruction manual. Stratagene
Grant, M., Brown, I., Adams, S., Knight, M., Ainslie, A., Mansfield, J. (2000) The RPM1 plant disease resistance gene facilitates a rapid and sustained increase in cytosolic calcium that is necessary for the oxidative burst and hypersensitive cell death, The Plant Journal 24 (4): 441-450
Physiological and Molecular Plant Pathology 50: 419-429
Heguy, A., Baldari, C.T., Macchia, G., Telford, J.L., Mellì, M. (1992) Amino acids conserved in interleukin-1 receptors (IL-1Rs) and the Drosophila Toll protein are essential for IL-1R signal transduction. The Journal of Biological Chemistry 267 (4) 2605-2609
class I loci reveals overdominant selection. Nature 335: 167-170
Phytopathology 36: 59-90
role of the N terminus and Leucine-Rich Repeat region of the Mi gene product in regulation of


James, C., Gschmeissner, S., Fraser, A., Evan, G.I. (1997) CED-4 induces chromatin condensation in Schizosaccharomyces pombe and is inhibited by direct physical association with CED-9. Current Biology 7: 246-252


Science 1 (4): 114-119
Enzymology 194:491-507


pECFP Vector Information, CLONTECH Laboratories Inc.


QIAexpressionist: A handbook for high level expression & purification of His6 proteins.


Ren, T., Qu, F., Morris, T.J. (2000) HRT gene function requires interaction between a NAC protein and viral capsid protein to confer resistance to turnip crinkle virus. The Plant Cell 12: 1917-1925

Renner, G. & Nguyen, P. (1982) Acute toxicities of the fungicide pentachloronitrobenzene and some of
its sulphur-containing metabolites administered to mice. Archives of Toxicology 51: 329-331
recognition. The Plant Cell 6: 511-520
Sessa, G., D’Ascenzo, M., Martin, G.B. (2000) Thr38 and Ser198 are Pto autophosphorylation sites required for the AvrPto-Pto-mediated hypersensitive response. The EMBO Journal 19 (10): 2257-2269
Sessa, G., D’Ascenzo, M., Martin, G.B. (2000b) The major site of the Pti1 kinase phosphorylated by the Pto kinase is located in the activation domain and is required for Pto-Pti1 physical interaction. European Journal of Biochemistry 267: 171-178


C. elegans CED-4, participates in cytochrome c-dependent activation of caspase-3. Cell 90: 405-413