Structure and function of tomato disease resistance proteins

van Ooijen, G.

Citation for published version (APA):
Gene-for-gene resistance in plants is based on the presence of a resistance (R) gene in the host and a matching Avirulence (Avr) gene in the pathogen. Many R genes have been cloned over the last two decades, mostly from the Solanaceae. The gene products, called R proteins, display modular domain structures. R protein function has recently been shown to require dynamic interactions between the various domains. In addition to these intramolecular interactions, R proteins interact with other proteins to form signalling complexes. These complexes are able to activate an innate immune response that arrests proliferation of the invading pathogen, thereby conferring disease resistance. In this review, we summarize current understanding of R protein structure and function, as well as the molecular mechanisms underlying the activation of defence signalling processes. Besides being a rich source for R genes, Solanaceae turn out to be a leading model system to study inter- and intramolecular interactions of R proteins.
Introduction

Solanaceae, a family of flowering dicot plants, likely derived the name from the Latin word *sol*, sun, because the flowers of the most prominent genus in this family, *Solanum*, resemble the sun and its rays. The family includes some prominent crops such as potato (*Solanum tuberosum*), tomato (*S. lycopersicum*), pepper (*Capsicum* spec.) and aubergine (*S. melongena*). Like all other plants solanaceous plants are attacked by a wide range of pathogens including oomycetes, viruses, bacteria, fungi, nematodes and insects like white fly and aphids leading to significant crop losses (Strange and Scott, 2005). In response to these attackers, passive and active defence mechanisms have evolved. Active defence responses can be subdivided into adaptive and innate immunity. Adaptive immunity in plants appears to be restricted to antiviral defense responses depending on an RNAi like mechanism (Voinnet, 2005). The innate immune system is more general and responds to a wide variety of plant pathogens. Innate immunity relies on specialized receptors that can be roughly divided into two groups: the Pathogen or Pattern Recognition Receptors (PRRs) and the Resistance (R) proteins. PRRs recognize Microbe- or Pathogen-Associated Molecular Patterns (MAMPs/PAMPs) that are often a part of highly conserved molecules shared between microorganisms of the same class (Nürnberger et al., 2004; Zipfel and Felix, 2005). PRRs allow plants to recognize distinct invaders using a limited set of receptors (Zipfel and Felix, 2005; Chisholm et al., 2006). In contrast to PRRs, R proteins respond to molecules (called avirulence proteins or elicitors) that are generally not conserved between species or even between isolates of a given pathogen. Accordingly, R proteins are encoded by large gene families, numbering several hundreds of genes per genome (Meyers et al., 2003). Because of the one-to-one relationship between a plant *R* gene and the matching *avirulence (Avr)* gene in a pathogen, this type of immunity was called ‘gene-for-gene’ resistance (Flor, 1942). Resistance mediated by R proteins is often associated with the appearance of localized cell death at the infection site, a phenomenon called the hypersensitive response (HR). This is distinct from the resistance response mediated by PRR receptors, as these generally do not induce an HR response upon pathogen recognition (Jones and Dangl, 2006). In this review, we provide an overview of the R genes that have been cloned from the Solanaceae. We describe the encoded proteins and their predicted structures. Furthermore, we discuss the current data available on the intra- and intermolecular interactions of R proteins from Solanaceae in the context of other model systems. The interaction patterns together with the identified downstream signalling components provide new insights into R protein function and downstream signalling.
Figure 1
Schematic representation of typical members of the four R protein classes. Protein domains and putative cellular localization are indicated. The Receptor-Like Protein (RLP) and the Receptor-Like Kinase (RLK) classes of R proteins span the plasma membrane (PM) and contain an extracellular Leucine Rich Repeat (LRR) domain. The CNL and TNL classes of R proteins are located intracellularly (cytoplasmic, nuclear, or membrane-bound) and contain a central NB-ARC domain (consisting of NB, ARC1 and ARC2 subdomains) coupled to an LRR domain. TNLs carry an N-terminal TIR domain, while CNLs contain either a CC or an extended CC domain.

R protein classification
Currently, over 55 R genes have been cloned from different monocot and dicot plant species ((Martin et al., 2003) and references in Table 1). Although R genes confer resistance to very different pathogens, the encoded proteins share a limited number of conserved elements. Based on these domains R proteins can be classified into five classes (Figure 1). The vast majority contains a central Nucleotide-Binding (NB) subdomain as part of a larger entity called the NB-ARC domain, which is present in the human apoptotic protease-activating factor 1 (APAF-1), R proteins, and the Caenorhabditis elegans homolog CED-4 (van der Biezen and Jones, 1998). C-terminal to the NB-ARC domain lies a leucine-rich repeat (LRR) domain, which is sometimes followed by an extension of variable length. Hence, this group is collectively referred to as NB-LRR proteins. These NB-LRR proteins are divided into two classes on the basis of their N-terminal region. If this region shows homology to a protein domain found in the Drosophila Toll and human Interleukin-1 Receptor (IL-1R), it is called the TIR domain (Whitham et al., 1994) and the proteins are referred to as TIR-NB-LRR or TNL proteins (TNL class). Since some non-TIR proteins contain predicted coiled-coil structures (CC) in their N-terminal domain, non-TIR NB-LRR proteins are collectively referred to as CC-NB-LRR or CNL proteins (CNL class). Phylogenetic analyses of the NB-ARC domains of NB-LRR proteins revealed separate clustering of TNL and CNL proteins. This suggests co-evolution of the N-terminal and NB-ARC domains and is indicative for an ancient segregation of these two classes providing an extra basis for the subdivision of NB-LRR R proteins.
All NB-LRR proteins are believed to act intracellularly. A more limited number of R proteins acts extracellularly and they contain a predicted extracellular LRR (eLRR) domain at their N-terminus. This eLRR is connected via a transmembrane domain to a variable cytoplasmic C-terminal region. When the cytoplasmic domain contains a protein kinase domain the R protein is placed in the RLK class, that of Receptor-Like Kinases. If no such domain is present it is placed in the RLP class, that of Receptor-Like Proteins. The last class represents a rest group of R proteins that cannot be placed in any of the first four classes.

Besides the genes that have been isolated and confirmed to function as R gene, numerous R gene homologues have been identified in genome sequencing and annotation programs. In Arabidopsis, TNLs form the largest group of NB-LRR proteins (Meyers et al., 1999), whereas this class is absent in monocots (Meyers et al., 1999; Pan et al., 2000). This difference could reflect differences in host/pathogen co-evolution in mono- and dicots. The majority of solanaceous NB-LRR proteins belongs to the CNL class (Table 1), whereas only three TNLs have been identified: the tomato Bs4, the potato Gro1-4 and the tobacco N gene conferring resistance to Xanthomonas campestris, Globodora rostochiöensis and Tobacco Mosaic Virus (TMV), respectively (Table1). The RLP class contains the tomato Cf and Ve proteins that confer resistance to Cladosporium fulvum and Verticillium albo-atrum, respectively (Kawchuk et al., 2001; Rivas and Thomas, 2005). The two Ve proteins differ from the Cf proteins by the presence of a putative C-terminal endocytosis signal (Kawchuk et al., 2001). No R proteins that belong to the RLK class have yet been identified within the Solanaceae. Except for the rice Xa21 gene, this group is made up of PRRs: the Arabidopsis EF-tu and flagellin FLS2 receptors (Gomez-Gomez and Boller, 2000; Chinchilla et al., 2006; Zipfel et al., 2006) and the tomato LeEIX (Ron and Avni, 2004). Tomato Asc1 has been placed in the rest class since it does not encode a ‘typical’ R protein involved in gene-for-gene resistance: resistance to Alternaria alternata is brought about by reduced sensitivity to the AAL mycotoxin rather than by specific recognition of the pathogen (Brandwagt et al., 2000).

**Avr proteins and R protein-mediated recognition**

It seems ill-considered for a pathogen to disclose its presence to a plant by secreting Avr gene products that are recognized by an R protein. Naturally, Avr genes did not evolve to serve this purpose and indeed evidence accumulates that many Avrs are virulence factors (Jones and Dangl, 2006). Plant pathogenic bacteria deliver approximately 15-30 proteins into host cells using a specialized type III secretion system (TTSS) (Buttner and Bonas, 2006). Whether these effectors suppress or trigger host defence depends on the host being attacked (Vinatzer et al., 2006).
### TABLE 1  Cloned solanaceous plant disease resistance genes

<table>
<thead>
<tr>
<th>Gene&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Plant</th>
<th>Ref&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Avr gene</th>
<th>Pathogen</th>
<th>Ref&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TNL</strong>&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bs4</td>
<td>Solanum lycopersicum</td>
<td>(Schornack et al., 2004)</td>
<td>AvrBs4, Hax4</td>
<td>Xanthomonas campestris</td>
<td>(Bonas et al., 1993; Kay et al., 2005)</td>
</tr>
<tr>
<td>Gro1-4</td>
<td>Solanum tuberosum</td>
<td>(Paal et al., 2004)</td>
<td></td>
<td>Globodera rostochiensis</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>Nicotiana tabacum</td>
<td>(Whitham et al., 1994)</td>
<td>Helicase</td>
<td>TMV</td>
<td>(Erickson et al., 1999)</td>
</tr>
<tr>
<td><strong>CNL</strong>&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bs2</td>
<td>Capsicum annuum</td>
<td>(Tai et al., 1999)</td>
<td>AvrBs2</td>
<td>Xanthomonas campestris</td>
<td>(Swords et al., 1996)</td>
</tr>
<tr>
<td>Gpa2</td>
<td>Solanum tuberosum</td>
<td>(van der Vossen et al., 2000)</td>
<td></td>
<td>Globodera pallida</td>
<td></td>
</tr>
<tr>
<td>Rxh1</td>
<td>Solanum lycopersicum</td>
<td>(Ernst et al., 2002)</td>
<td></td>
<td>Globodera rostochiensis, G. pallida</td>
<td></td>
</tr>
<tr>
<td>Hero&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Solanum lycopersicum</td>
<td>(Ori et al., 1997; Simons et al., 1998)</td>
<td></td>
<td>Fusarium oxysporum</td>
<td></td>
</tr>
<tr>
<td>I-2</td>
<td>Solanum lycopersicum</td>
<td>(Milligan et al., 1998b; Vos et al., 1998)</td>
<td></td>
<td>Meloidogyne incognita, M. arenaria, M. javanica, Bemisia tabaci</td>
<td></td>
</tr>
<tr>
<td>Mi-1.2&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Solanum lycopersicum</td>
<td>(Salmeron et al., 1996)</td>
<td>AvrPto, AvrPtoB</td>
<td>Pseudomonas syringae</td>
<td>(Ronald et al., 1992; Abramovitch et al., 2003)</td>
</tr>
<tr>
<td>Prf&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Solanum lycopersicum</td>
<td>(Ballvora et al., 2002)</td>
<td></td>
<td>Phytophthora infestans</td>
<td></td>
</tr>
<tr>
<td>R1&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Solanum tuberosum</td>
<td>(Huang et al., 2005)</td>
<td>Avr3a</td>
<td>Phytophthora infestans</td>
<td>(Armstrong et al., 2005)</td>
</tr>
<tr>
<td>R3a</td>
<td>Solanum tuberosum</td>
<td>(Song et al., 2003; van der Vossen et al., 2003)</td>
<td></td>
<td>Phytophthora infestans</td>
<td></td>
</tr>
<tr>
<td>Rpi-blb1</td>
<td>Solanum bulbocastanum</td>
<td>(van der Vossen et al., 2005)</td>
<td></td>
<td>Phytophthora infestans</td>
<td></td>
</tr>
<tr>
<td>RB</td>
<td>Solanum bulbocastanum</td>
<td>(Bendahmane et al., 1999)</td>
<td>CP</td>
<td>PVX</td>
<td>(Bendahmane et al., 1995)</td>
</tr>
<tr>
<td>Rpi-blb2&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Solanum bulbocastanum</td>
<td>(Bendahmane et al., 2000)</td>
<td>CP</td>
<td>PVX</td>
<td>(Bendahmane et al., 2000)</td>
</tr>
<tr>
<td>Sw-5&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Solanum</td>
<td>(Brommonschenk)</td>
<td></td>
<td>Tospovirus</td>
<td></td>
</tr>
</tbody>
</table>
Effectors appear to manipulate many signalling processes, but only for a small number of effectors a host target has been identified (Grant et al., 2006). Clues on how they interfere with signal transduction processes of the host can sometimes be obtained from structural homology to other proteins. For instance, members of the Xop/J/AvrRxv effector family appear to encode SUMO proteases, suggesting interference with SUMO signalling (Orth et al., 2000), while AvrPtoB is an active ubiquitin E3 ligase likely interfering with specific protein degradation of the host (Janjusevic et al., 2006). Avrs such as AvrRpt2 and AvrPphB act as proteases and cleave specific host proteins (Shao et al., 2002; Axtell et al., 2003; Shao et al., 2003; Kim et al., 2005b). For recent reviews on bacterial effectors we refer to (Mudgett, 2005; Abramovitch et al., 2006; Grant et al., 2006). Less is known on fungal effectors. Four Avr genes have been cloned from the fungus Cladosporium fulvum (Table 1), which all encode small cysteine-rich extracellular proteins.

<table>
<thead>
<tr>
<th>Effectors</th>
<th>Host</th>
<th>Effectors</th>
<th>Host</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP ToMV</td>
<td>Solanum lycopersicum</td>
<td>Lanfermeijer et al., 2005</td>
<td>(Calder and Palukaitis, 1992)</td>
<td></td>
</tr>
<tr>
<td>MP ToMV</td>
<td>Solanum lycopersicum</td>
<td>Lanfermeijer et al., 2003</td>
<td>(Calder and Palukaitis, 1992)</td>
<td></td>
</tr>
</tbody>
</table>

RLP

<table>
<thead>
<tr>
<th>Effectors</th>
<th>Host</th>
<th>Effectors</th>
<th>Host</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avr2</td>
<td>Cladosporium fulvum</td>
<td>(Loderer et al., 2002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avr4</td>
<td>Cladosporium fulvum</td>
<td>(Joosten et al., 1994)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avr4E</td>
<td>Cladosporium fulvum</td>
<td>(Westerink et al., 2004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avr9</td>
<td>Cladosporium fulvum</td>
<td>(van Kan et al., 1991)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avr9</td>
<td>Cladosporium fulvum</td>
<td>(van Kan et al., 1991)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avr9</td>
<td>Cladosporium fulvum</td>
<td>(van Kan et al., 1991)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avr9</td>
<td>Cladosporium fulvum</td>
<td>(van Kan et al., 1991)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ve1</td>
<td>Verticillium albo-atrum</td>
<td>(Kawchuk et al., 2001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ve2</td>
<td>Verticillium albo-atrum</td>
<td>(Kawchuk et al., 2001)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Other

<table>
<thead>
<tr>
<th>Effectors</th>
<th>Host</th>
<th>Effectors</th>
<th>Host</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asc-1</td>
<td>Solanum lycopersicum</td>
<td>Brandwagt et al., 2000</td>
<td>Alternaria alternata</td>
<td></td>
</tr>
</tbody>
</table>
proteins found in the tomato apoplast. Avr2 acts as a protease inhibitor (Rooney et al., 2005), while Avr4 binds to chitin present in fungal cell walls, thereby protecting it from degradation by plant chitinases (van den Burg et al., 2006). The function of the remaining C. fulvum Avrs remains unresolved but they probably exert their function extracellularly (Joosten and De Wit, 1999; Thomma et al., 2005). Other plant pathogenic fungi, like rusts and powdery mildews, form more intimate contacts with their hosts by the formation of specialized feeding structures called haustoria. These specialized feeding structures penetrate the host plant cell wall, but remain separated from the host cytoplasm by the host cell envelope. Recently, several Avr genes have been identified from haustoria-forming fungi (Catanzariti et al., 2006; Dodds et al., 2006; Ridout et al., 2006). These genes encode proteins with a predicted signal peptide for secretion into the extrahaustorial matrix, but nonetheless act inside the host cell where they are perceived by matching NB-LRR proteins. How these Avrs are taken up by the host is currently not clear (Dodds et al., 2006; Ridout et al., 2006). For a recent review on fungal Avrs and secreted proteins we refer to (Rep, 2005).

The Avr3a protein from the oomycete Phytophthora infestans also acts inside the host cell as in the absence of the R3 resistance gene it is able to suppress cell death triggered by the P. infestans elicitin INF1 (Bos et al., 2006). Avr3a and two other oomycete Avrs cloned from Hyaloperonospora parasitica share a conserved sequence motif, RxLR (Kamoun, 2006). This motif might be the signal for uptake into the host cell from the extrahaustorial space.

Based on the existence of monogenic resistance against animal pathogens such as nematodes and insects, the latter are also predicted to contain Avr genes (Vos et al., 1998; Williamson and Kumar, 2006). Unfortunately, animal Avr genes have not yet been identified (Kaloshian and Walling, 2005; Williamson and Kumar, 2006).

Because in all cases described above corresponding R and Avr proteins co-localize, one interpretation of the gene-for-gene relation assumes a direct interaction between the two proteins. Using yeast two-hybrid and in vitro pull-down approaches, a direct interaction was indeed confirmed in a few cases, i.e. between Pita / AvrPita (Jia et al., 2000), L / AvrL567 (Dodds et al., 2006), PopP2 / RRS-1 (Deslandes et al., 2003) and the TMV p50 helicase with the NB-ARC-LRR part of the tobacco N protein (Ueda et al., 2006). The latter interaction, however, was not found by others using in vivo pull-down assays (Peart et al., 2005; Mestre and Baulcombe, 2006), or rather an indirect interaction was detected (Caplan et al., 2008; also see below). A direct interaction was recently proposed for the Bs4 / AvrBs4 pair (Schornack et al., 2006) since AvrBs4 has a tetratricopeptide repeat (TPR)-like structure also found in the I-2 interacting protein phosphatase 5 (PP5) (de la Fuente van Bentem et al., 2005). Since most bacterial Avrs probably act as effector proteins that target host cellular
components, R proteins might sense the presence of an Avr protein by monitoring the state of the host target. This indirect interaction model is called the guard hypothesis (Van der Biezen and Jones, 1999; Dangl and Jones, 2001). Although a guard that waits until its guardee is killed may seem a bit odd, this theory has gained support. The best studied example of modification of a host target by Avrs, resulting in R protein activation, is the phosphorylation or cleavage of Arabidopsis Rin4 protein by AvrB, AvrRpm1 and AvrRpt2, respectively (Mackey et al., 2002; Kim et al., 2005b). These events trigger activation of the NB-LRR proteins Rpm1 and Rps2. Also, the cleavage of the Pbs1 kinase bound to Rps5 by the avrPphB protease, thereby triggering Rps5 activation, supports the guard hypothesis (Shao et al., 2003).

Three additional examples in support of the guard hypothesis have been described in tomato-pathogen model systems. Firstly, an interaction has been shown between the Pseudomonas syringae effectors AvrPto and AvrPtoB and the tomato kinase Pto, sensed by NB-LRR protein Prf. The Avr proteins interact directly with their target Pto (Tang et al., 1996; Kim et al., 2002), while Pto constitutively binds to the N-terminus of Prf (Mucyn et al., 2006). Although the presence of a tertiary complex could not be shown, interaction of all three components is essential for induction of defence signalling. Upon binding of AvrPto or AvrPtoB to Pto, a conformational change of this protein is proposed to trigger the defense signalling potential of the interacting Prf R protein (Mucyn et al., 2006).

A second example of indirect pathogen perception in tomato is the Avr2/Cf-2 system. Cf-2 resistance is strictly dependent on the tomato cysteine protease Rcr3 that is inhibited by the Avr2 protein from C. fulvum. Interestingly, the Rcr3 sequence from cultivated tomato (Solanum lycopersicum L.) differs slightly from the S. pimpinellifolium Rcr3, and SlRcr3 induces Cf-2 dependent necrosis in the absence of any Avr. This observation suggests that a small conformational difference in the SlRcr3 protein mimics the Avr2-inhibited SpRcr3 state so that it activates Cf-2 (Rooney et al., 2005).

The third example of indirect Avr binding is provided by the TNL protein N of tobacco. Using co-immunoprecipitations and bimolecular fluorescence complementation (BiFC) experiments, it was shown that N and p50, the part of the helicase protein that is required for its Avr function, associate in vivo. This association is mediated by an unknown protein binding to the TIR domain, called N interacting protein (NIP1). NIP1 is a good candidate to be the guardee of N (S. P. Dinesh-Kumar, personal communication). Besides the examples described above, indirect Avr recognition has been proposed for the tomato Mi-1.2 protein. Mi-1.2 is intriguing since it confers resistance to highly diverse animal pathogens: the root-knot nematode Meloidogyne incognita, the potato aphid Macrosiphum euphorbiae and the whitefly Bemisia tabaci (Milligan et al., 1998a; Vos et al., 1998; Nombela et al., 2003). Mi-1.2 mediated
resistance to nematodes is accompanied by an HR, while an HR is not observed in the interaction with whitefly and the potato aphid (Williamson, 1999). A mutant screen identified \textit{Rme-1}, which is unlinked to \textit{Mi-1.2}, but is required for Mi-mediated resistance against all three pathogens. As \textit{Rme-1} is not involved in other \textit{R} gene pathways and acts upstream (or at the same step) of Mi-1.2 in the signalling cascade, it is a candidate for the Mi-1.2 guardee (Martinez de Ilarduya et al., 2004). However, cloning of the \textit{Rme-1} gene is needed to confirm that \textit{Rme-1} functions as the Mi-1.2 guardee.

Based on the few examples for direct and indirect Avr perception it cannot be concluded at this stage whether there is a prevalence for one over the other. During evolution, \textit{R} protein recognition specificities are likely to be generated at random, some targeting Avr products directly, and others recognizing host factor modifications. The group targeting Avrs directly can relatively easily be overcome by mutations in the Avr protein that abolish the interaction, but not its virulence function. Resistance based on indirect recognition is (at least theoretically) more difficult to overcome, since mutations avoiding recognition will also affect virulence.

**Structural features of \textit{R} protein domains**

As described above, \textit{R} proteins can be classified based on their protein domain architecture. Figure 1 schematically presents the four structural classes of \textit{R} proteins. Unfortunately, crystal structures have not yet been obtained of any domain of plant \textit{R} proteins. However, crystal structures of LRR and NB-ARC domains have been obtained from evolutionary related proteins. These crystal structures have been used for 3D structure modelling studies of plant LRR and NB-ARC domains (van der Hoorn et al., 2005; McHale et al., 2006; Takken et al., 2006). These studies have provided insight into some structural features of these domains and they are utilized to explain the effect of certain mutations on \textit{R} protein activation and downstream signalling (McHale et al., 2006; Takken et al., 2006).

**The Leucine-Rich Repeat domain**

The LRR domain is the only domain present in all \textit{R} proteins listed in Table 1. LRRs are present in many receptors of virtually any organism, where this domain is involved in ligand recognition (Enkhbayar et al., 2004). The LRR domain consists of two to 42 repeats each comprising a \(\beta\)-sheet with the core consensus \textit{xxLxLxx} (Enkhbayar et al., 2004). In plant LRRs, individual repeats are formed by 24-28 residues and contain a core of 14 residues with consensus sequence \textit{LxxLxxLxxC/Nxx}. This core forms the \(\beta\)-sheet and the attached loop-regions. Each core is separated from the next by a spacer of variable length. Crystal structures of over 20 LRR proteins have revealed differences in overall structure, but in all proteins
a series of parallel $\beta$-strands form a right-handed super helical beta-sheet (Kobe and Kajava, 2001; Enkhbayar et al., 2004). The only plant LRR protein for which the crystal structure was solved is PGIP2, an eLRR protein that binds and inhibits polygalacturonases from fungi such as *Fusarium moniliforme* and *Aspergillus niger* (Leckie et al., 1999). The PGIP2 structure is characterized by the presence of a second $\beta$-sheet ($\beta_2$) in each repeat, with consensus NxLxGx, connecting the first ($\beta_1$) with an $\alpha$-helix in the spacer (Di Matteo et al., 2003).

In plant R proteins there are several differences between intra- and extracellular LRR domains. In the LRRs of TNLs and CNLs no clear subdomains are apparent, but a conserved motif is present in the third LRR (Meyers et al., 2003). Mutation of the D in this so-called VLDL motif in the potato Rx protein to an E produced a constitutively active protein (Bendahmane et al., 2002), whereas a mutation adjacent to this motif in the *Arabidopsis* CNL Rps5 had inhibitory epistatic effects on resistance (Warren et al., 1998). The role of this motif for NB-LRR protein function is not clear, but the sequence fits the consensus for a leucine-rich nuclear export signal (Bendahmane et al., 2002; la Cour et al., 2004). Plant eLRR proteins are characterized by a longer repeat consensus sequence than intracellular LRRs; a 24-residue motif LxxLxxLxxNxLxGxIPxxLGx instead of the general 14-residue consensus (Federici et al., 2006). The consensus of the PGIP2 $\beta_2$ is conserved in the Cf proteins and since ligand binding presumably only involves $\beta_1$, $\beta_2$ might be involved in binding of other proteins (Rivas and Thomas, 2005; Federici et al., 2006) or in homo- or hetero dimerization with other eLRRs. The Cf and Ve eLRR domains can be subdivided into three subdomains, in Cf proteins these are referred to as: C1, 2 and 3 (Rivas and Thomas, 2005). C1 forms the major part of the eLRR domain and consists of 21-28 hypervariable repeats. C2 forms a spacer domain separating C1 from C3. C3 consists of three to four relatively conserved repeats (Rivas and Thomas, 2005). R protein eLRR domains, unlike the LRR domains in NB-LRR proteins, have many putative N-linked glycosylation (NGS) sites in the exposed regions. For Cf-9, the NGS sites in the C1 have been shown to be essential for Cf-9 function, whereas introduction of NGS sites in the C3 disrupts its function (van der Hoorn et al., 2005). Another unique feature of eLRRs is the presence of N- and C-terminal ‘capping’ domains. For the small tomato eLRR protein LRP, these capping domains were shown to form disulphide bonds that are necessary to protect the eLRR from proteolysis (Kolade et al., 2006).

**The NB-ARC domain**

The recently published crystal structures of the NB-ARC domains of APAF-1 and CED-4 revealed that the NB-ARC domain consists of clearly distinguishable subdomains. Besides the NB subdomain, the NB-ARC of APAF-1 contains three
additional subdomains (ARC1-ARC3) that form a four-helix bundle, a winged-helix fold and a helical bundle, respectively (Riedl et al., 2005). Subdomain ARC3 is absent in plant R proteins and CED-4, but ARC1 and ARC2 are conserved (Figure 2)(Yan et al., 2005; Albrecht and Takken, 2006). Proteins containing an NB-ARC domain are evolutionary related to the mammalian NACHT-LRR protein family (such as NAIP, CIITA, HET-E and TP1), many of which function in innate immunity (Inohara et al., 2002; Leipe et al., 2004; Albrecht and Takken, 2006). ARC1, 2 and 3 of APAF-1 correspond to the NAD1-3 subdomains of mammalian intracellular NACHT-LRR proteins (Albrecht and Takken, 2006).

The NB-ARC and NACHT-LRR proteins belong to the STAND super family of ATPases (Signal Transduction ATPases with Numerous Domains). The nucleotide-binding domain of these proteins is proposed to function as a molecular switch, where NTP-hydrolysis induces a conformational switch required to regulate signal transduction (Leipe et al., 2004). The NB subdomains of both NB-ARC and NACHT domains form an NTP binding fold consisting of a parallel five-stranded β-sheet flanked by α-helixes (Takken et al., 2006), placing them in the large group of P-loop NTPases (Vetter and Wittinghofer, 1999). The β-sheet is five-stranded and assumes a 2-3-4-1-5 topology (Figure 2). The strands β1 and β3 encompass the most conserved motifs of the NB subdomain: the P-loop itself (also called Walker A motif), and the Walker B motif. The P-loop is defined by the consensus sequence GxxxxGKS/T that bind β- and γ-phosphates of the nucleotide through the Lys residue, and the Ser and Thr residues coordinate the Mg\(^{2+}\)-ion in the binding pocket.

The Walker B motif is defined by the hhhhhDD/E consensus site (where h represents a hydrophobic residue), in R proteins this motif is often hhhDD. The first Glu residue (D) is also important for the coordination of the Mg\(^{2+}\)-ion while the acidic residue (D/E) is thought to act as a catalytic base during nucleotide hydrolysis (Leipe et al., 2004; Hanson and Whiteheart, 2005).

Besides the P-loop and Walker B motifs, several other motifs have been identified in the NB-ARC domain (Figure 2). The hhGRExE motif is part of a linker region connecting the NB subdomain with the N-terminus. In the ARC2 subdomain, a highly conserved motif is present called the MHD-motif. Other motifs annotated in the NB-ARC domain of R proteins include Resistance NBS motifs RNBS-A, -B, -C and –D, the GxP motif and motifs VII and VIII (Figure 2). For an overview on synonymous names for these motifs, we refer to (Takken et al., 2006). Not surprisingly, many residues in the conserved motifs are essential for NB-LRR protein function and are predicted to map at positions where they are able to interact with the nucleotide (Takken et al., 2006). For instance, the conserved His residue in the APAF-1 MHD motif interacts directly with the β-phosphate of dATP (Riedl et al., 2005). Mutation of this motif in the R proteins I-2, Mi-1.2, L6 and Rx results in autoactivation (Moffett et
The NB-ARC domain of I-2, Mi-1.2 and N has been shown to specifically bind (d)ATP, similar to APAF-1 and CED4 (Tameling et al., 2002; Kim et al., 2005a; Yan et al., 2005; Ueda et al., 2006). The NB-ARC domain of these R proteins also catalyzes the hydrolysis of ATP to ADP, as shortly discussed below.

**ATP hydrolysis by NB-ARC domains**

In APAF-1, binding of cytochrome c results in hydrolysis of ATP by the NB-ARC domain, followed by replacement of the formed ADP by ATP (Kim et al., 2005a). ATP hydrolysis and nucleotide exchange are both essential to produce the activated state. Like APAF-1, CED-4 requires ATP binding for function. However, hydrolysis or nucleotide exchange has not been observed for CED-4 (Yan et al., 2005).

Specific binding and hydrolysis of ATP was recently observed for the solanaceous R proteins I-2, Mi-1.2 and N (Tameling et al., 2002; Ueda et al., 2006). For I-2,
biochemical analysis of two autoactivating mutants revealed that these mutants are affected in ATP hydrolysis but not in nucleotide binding (Tameling et al., 2006). For N, the ATP- and not the ADP-bound state of the protein was shown to interact with the elicitor (Ueda et al., 2006). Together, these observations suggest that in the active, HR signalling state, ATP is bound to the NB-ARC. Mutations in the P-loop of several other R proteins were shown to abolish function (Dinesh-Kumar et al., 2000; Tao et al., 2000; Tornero et al., 2002). These data indicate that nucleotide binding is essential for R protein signalling. To define more exactly the effect of the nucleotide binding status of R proteins on R protein function, biochemical analyses of NB-ARC domains or even entire R proteins will be a valuable development.

The N-terminal domain
As described above, the two NB-LRR classes are distinguished by the presence or absence of a TIR domain at the N-terminus. The TIR domain is conserved in metazoan Toll and IL-1R proteins (Xu et al., 2000), involved in innate immunity. Crystal structures have been published for human Toll-like receptors TLR1 and TLR2. Structures of plant TIR domains have not yet been resolved, but structures similar to animal family members are proposed, since amino acids shared between the TNL N and the metazoan Toll or IL-1R protein were shown to be essential for functioning of both (Dinesh-Kumar et al., 2000). The topology of the TIR domain consists of a parallel five-stranded β-sheet, five α-helices, and connecting loops. The loop between β2 and α2 contains several highly conserved surface-exposed residues that are important for receptor signalling by recruiting interacting proteins (Xu et al., 2000).

CNL proteins often contain a predicted coiled-coil domain (α-helix-rich domain that contains seven-residue repeat sequences). In Table 1, we have discriminated two types of CC domains: a short CC domain and an extended CC domain. This extended domain has only been found in Solanaceae and might be specific for this family. Recently, a novel protein-protein interaction domain was identified in the extended CC domain of tomato Prf. This domain has been found in solanaceous R proteins with an extended CC domain only, and was called the Solanaceous Domain (SD) (Mucyn et al., 2006).

A recent sequence analysis has revealed a short conserved motif in the middle of the short CC. The nT motif was originally described due to its high conservation in monocot CC-NB-LRRs (Bai et al., 2002). This motif is less conserved in dicots but a core EDVID motif can be identified in many but not all CC-NB-LRR proteins and is one of the few motifs that can be identified as being broadly conserved across the CC domains of NB-LRRS of different families of plants (Raiddan et al., 2008). This
EDVID motif was proven to be required for Rx function, due at least in part to its requirement for an intra-molecular interaction involving the CC domain (Rairdan et al., 2008).

**R protein domain functionality and interplay**

As detailed above, R proteins are composed of several (sub)domains that appear to function consecutively during signal perception. In this section we address the following questions: (1) What is the specific function of each (sub)domain? (2) How do the domains interact? (3) How does this result in a protein that is able to recognize Avr proteins and trigger activation of defence responses?

Many studies have shown that the LRR domain is under diversifying selection, and specifically the surface-exposed residues in the β-sheet that are putatively involved in ligand recognition (Meyers et al., 1998; Noel et al., 1999; Mondragon-Palomino et al., 2002). Diversifying selection is compatible with a receptor surface that might be involved in Avr perception (Mondragon-Palomino et al., 2002). The prediction that the LRR confers Avr specific recognition has indeed been confirmed experimentally and this specificity can be changed with a few minor changes into a new one (Ellis et al., 1999; Jia et al., 2000; Van der Hoorn et al., 2001; Wulff et al., 2001; Dodds et al., 2006; Rairdan and Moffett, 2006). Interestingly, a significant degree of diversifying selection was also observed in the N-terminal region of some NB-LRR proteins, suggesting that this region is also important for the origination of new recognition specificities (Mondragon-Palomino et al., 2002) and this has been confirmed experimentally (Ellis et al., 1999; Luck et al., 2000). In addition to a role in recognition specificity, the N-terminal domain may be involved in the recruitment of downstream signalling components. A strong indication that the N-terminus is involved in downstream signalling is the differential requirement of downstream signalling components by CNLs and TNLs, reviewed in (Martin et al., 2003).

With the LRR and the N-terminal domain conferring pathogen recognition and downstream signalling, the central NB-ARC is thought to act as the molecular switch that controls the activation state of the protein (Takken et al., 2006). The “on” and “off” state of the switch being defined by the nucleotide that is bound to the NB-ARC domain (ATP versus ADP).

Important progress has been made over the last years in understanding the function of the intra-molecular interaction in the NB-LRR R proteins. R protein signalling by NB-LRRs appears to depend on dynamic interactions between the three major domains. Positive and negative regulatory regions are scattered over the entire protein as demonstrated by many auto-activating and loss-of-function mutations (Dinesh-Kumar et al., 2000; Tao et al., 2000; Axtell et al., 2001; Bendahmane et al., 2002; Tornero et al., 2002; Takken et al., 2006). These mutations suggest that
compatibility between the different domains is required to keep NB-LRR proteins in an auto-inhibited but activatable state. Support for such harmonisation between domains is provided by domain swap experiments between homologous R proteins. For instance, LRR exchange between homologues at the maize \textit{Rp1} locus results in a necrotic phenotype (Sun et al., 2001). Likewise, domain swaps between \textit{Mi-1.2} and its parologue \textit{Mi-1.1} resulted in autoactivating incompatibilities between the LRR of \textit{Mi-1.2} and either the first 161 amino acids of the amino terminal domain (NT-1), or the next 751 residues encompassing the NB-ARC domain of \textit{Mi-1.1} (Hwang et al., 2000). Autoactivating fusions were also obtained by swapping different, non-overlapping, LRR sequences in \textit{Rx} for the equivalent region of Gpa2 (Rairdan and Moffett, 2006). These experiments suggest that the different domains interact with each other and that mis-pairing can result in spontaneous activation of the protein.

Studies on the potato protein \textit{Rx} provided the first example of direct physical and functional intra-molecular interactions between the different protein domains found in NB-LRR proteins (Moffett et al., 2002). Co-expression of either the CC and NB-ARC-LRR fragments or the CC-NB-ARC and LRR fragments reconstituted Avr-dependent HR. This transcomplementation involved physical interaction between the domains, which was disrupted by co-expression of the PVX coat protein, the activator of \textit{Rx} (Moffett et al., 2002; Rairdan and Moffett, 2006). Interestingly, the interaction between the CC domain and the NB-ARC-LRR fragment requires a functional EDVID motif in the CC domain (Rairdan et al., 2008). In vitro, the LRR domain of the tobacco TNL protein \textit{N} was found to interact with the TIR-NB-ARC fragment and this interaction was lost when the elicitor was added (Ueda et al., 2006). The latter observation however, could not be confirmed in vivo, as no intramolecular interactions were detectable upon co-expression of \textit{N} domains. Perhaps the intramolecular interactions in \textit{N} are weaker than in \textit{Rx} or are not effective for proteins expressed in trans (Mestre and Baulcombe, 2006). Immunoprecipitation studies with the \textit{R} protein \textit{Bs2} also showed an intra-molecular interaction between the CC-NB-ARC and LRR parts, but instead of disrupting this interaction, the interaction was enhanced when the elicitor (AvrBs2) was co-expressed. No interaction could be observed between the CC fragment and NB-ARC-LRR of \textit{Bs2}. Strikingly, the interaction between the CC-NB-ARC and LRR parts of \textit{Bs2} was no longer detectable upon silencing of SGT1 (Leister et al., 2005). These differences support the idea that SGT1 acts as a co-chaperone of R-protein complexes (Leister et al., 2005), as will be discussed later.

To investigate the function of the various subdomains of the NB-ARC involved in the intramolecular interactions with the N- and C-terminal domains, a series of domain swaps between \textit{Rx} and Gpa2 were made. These studies showed that the \textit{Rx} ARC1 subdomain is necessary for the interaction between the CC-NB-ARC and LRR
fragments (Rairdan and Moffett, 2006). Binding to LRR domains seems to be a general property of the ARC1 subdomain. For instance the CC-NB-ARC part of Rx was able to interact with the LRR domains of several other NB-LRRs as well (i.e. Bs2 and *Arabidopsis* HRT LRR domains). However, these interactions did not reconstitute a functional R protein capable of activating defence responses (neither autoactivation nor elicitor-dependent activation) (Rairdan and Moffett, 2006). Point mutations in the ARC1 subdomain had only a quantitative effect on the binding affinity for the LRR domain and reduced LRR binding affinity did not necessarily compromise functionality of the R protein. This observation supports an extensive and partially conserved interaction surface between the ARC1 and the LRR domain with individual contacts quantitatively contributing to the interaction. Subsequent deletion studies with the LRR of Rx revealed that the intact LRR domain is necessary for the interaction with the CC-NB-ARC part suggesting that the overall fold is needed for the interaction or that the interaction surface is scattered over the entire LRR domain (Rairdan and Moffett, 2006).

A function for the ARC2 subdomain became apparent when the Gpa2 ARC2 subdomain was swapped with the Rx ARC2 domain, as this resulted in a constitutively active protein. The LRR is believed to have a positive regulatory role in the activation process, as R proteins are generally inactive in the absence of their genuine LRR when expressed at their endogenous levels (Rathjen and Moffett, 2003; Rairdan and Moffett, 2006). Thus, the ARC2 domain of Gpa2 and LRR domain of Rx are sufficiently compatible to execute the positive regulatory role of the LRR and to induce HR, but the combination lacks the structural constraints to prevent autoactivation. Rairdan & Moffett therefore concluded that the ARC2 subdomain relays pathogen recognition mediated by the LRR domain into changes in R protein conformation, unleashing its downstream signalling potential (Rairdan and Moffett, 2006). A regulatory role for the ARC2 subdomain fits with the fact that mutations in the ARC2 subdomain often result in an autoactivating protein (Takken et al., 2006).

These observations suggest fine-tuning and co-evolution between the N-terminal end of the LRR domain and the ARC2 subdomain in NB-LRR proteins. This idea might be tested by a survey of polymorphisms in NB-LRR proteins.

A model in which Avr perception alters the interaction between the LRR and ARC2 subdomains, repositioning critical motifs and thereby allowing the molecule to progress to an active conformation, is in agreement with domain swap experiments between Mi-1.2 and Mi-1.1 (Hwang and Williamson, 2003). These swaps demonstrated autoactivating incompatibilities between the LRR and a region encompassing the ARC2 (Hwang and Williamson, 2003). Additional intramolecular interactions for Mi-1.2 were suggested based on mutational analysis. A single amino acid replacement in the first half of the Mi-1.2 LRR by the Mi-1.1 residue (R961D)
leads to autoactivation, suggesting that this residue is needed for autoinhibition of HR signalling. This autoactivation phenotype was suppressed in trans by overexpression of a subdomain of the Mi-1.1 N-terminus (NT1), suggesting that HR signalling mediated by the LRR is normally suppressed by the NT1 (Hwang et al., 2000). The necessity of compatibility between the NT1 domain and residue 961 was also shown by swapping the Mi-1.1 NT1 into a Mi-1.2 background, which results in autoactivation. However, this NT1 domain is absent in CNL proteins with a shorter N-terminus (Table 1).

The ability of NB-LRR proteins to establish intramolecular interactions depends on the conformation of the NB subdomain, regulated by the nucleotide bound. For instance, mutations in the P-loop of Rx abolished the interaction between the CC fragment and the NB-ARC-LRR fragment, whereas the LRR and CC-NB-ARC interaction was not lost (Moffett et al., 2002). Also in Bs-2 the latter interaction does not require a functional P-loop (Leister et al., 2005), whereas in N it does in vitro (Ueda et al., 2006). Direct support for a conformational change in the NB-ARC domain upon binding different nucleotides came from the tomato CNL protein I-2. The ADP-bound state of the protein displayed increased affinity for the nucleotide compared to the ATP-bound state (Tameling et al., 2006). In addition, two autoactivation mutants of I-2 were shown to have a wild-type nucleotide binding affinity but a reduced ATPase activity. These data suggested that the ADP conformation reflects a resting state whereas the ATP conformation represents the activated state of the I-2 protein (Tameling et al., 2006). In contrast, in a model proposed for the N protein the ATP-bound state reflects the resting state (Ueda et al., 2006). Upon Avr perception by N, its ATPase activity is stimulated resulting in hydrolysis of the nucleotide and transition to the activated, ADP-bound state of the protein. Unfortunately, the I-2 and N data are not directly comparable as they were obtained using truncated proteins, containing either the CC-NB-ARC (I-2) or the NB-ARC-LRR (N). The crucial next step will be to perform nucleotide binding and hydrolysis experiments using full-length R proteins.

An important additional role of the NB-ARC domain was found for non-plant NB-ARC proteins. The NB-ARC domains of the mammalian NB-ARC protein APAF-1 and its Drosophila and C. elegans analogues DARK1 and CED4, have been shown to homo-oligomerize upon activation, to form wheel-like structures of respectively 7, 8, and 4 molecules. These ‘wheels of death’ provide a platform for binding and subsequent activation of downstream procaspases (Yan et al., 2005; Yu et al., 2005; Yu et al., 2006).

Recently, also for the tobacco TNL protein N oligomerization upon recognition of the Avr was observed (Mestre and Baulcombe, 2006). Oligomerization and resistance both require an intact P-loop, hence presumably the ability to bind nucleotides.
Interestingly, mutation of the RNBS-A motif does not affect elicitor-dependent oligomerization but still abolishes resistance (Mestre and Baulcombe, 2006), indicating that resistance is not an automatic consequence of oligomerization. The N TIR domain can also oligomerize on its own (Mestre and Baulcombe, 2006). However, oligomerization of the TIR domain was independent of the Avr protein. Mutation of three of the predicted solvent exposed residues in the TIR domain abolished HR signalling coinciding with a weaker homotypic interaction of the TIR domain. However, when these mutations were introduced in the full-length protein oligomerization still occurred. This suggests that the elicitor-triggered oligomerization of the full-length R protein mainly involves NB-ARC::NB-ARC interactions, similar to the CED-4, APAF-1 and DARK oligomers. Regretfully, the N NB-ARC domain alone was not stable upon transient expression to test this hypothesis (Mestre and Baulcombe, 2006). A major remaining question is now whether oligomerization is a general feature of NB-LRR R proteins or whether it is unique for N or possibly the TNL class.

**Intermolecular interactions of R proteins**

Besides the intramolecular interactions described above, R proteins interact with other proteins to form large, dynamic, multimeric protein complexes. Yeast two-hybrid screens and recently also co-immunoprecipitation experiments have identified R protein-interacting proteins. For solanaceous NB-LRR proteins entire proteins as well as the different domains have been used as baits. For the LRR domain, thought to be the specificity determinant, this did not reveal putative “guardees” and/or Avr products. Rather, LRR interactors were found to be chaperones and chaperone-associated proteins. For example, the LRRs of N and I-2 physically interact with heat shock protein 90 (Hsp90) (Liu et al., 2004; de la Fuente van Bentem et al., 2005). Hsp90 is a chaperone mainly involved in folding receptor proteins into a signalling competent state (Pratt et al., 2004). Whereas the N-terminal part of the LRR domain of I-2 interacts with Hsp90, the C-terminal part was found to bind specifically to protein phosphatase 5 (PP5) (de la Fuente van Bentem et al., 2005). Binding is not exclusive for I-2 since in yeast two-hybrid assays PP5 was also found to interact with the solanaceous R proteins Mi-1.2 and Rx, and *Arabidopsis* Rps5 and Rpm1 (de la Fuente van Bentem et al., 2005) as well as N (S. P. Dinesh-Kumar, unpublished data). PP5 not only interacts with the LRR domain of NB-LRR proteins, but through its TPR domain also with the C-terminus of Hsp90. The biological function of the single copy gene PP5 in disease resistance remains elusive, however, since neither knock-down in tomato nor knock-out in *Arabidopsis* affects disease resistance (de la Fuente van Bentem et al., 2005). In contrast, silencing of the different Hsp90
homologs showed a requirement for Hsp90 in disease resistance and HR activation for a large number of NB-LRR and RLP R proteins (Hubert et al., 2003; Lu et al., 2003; Takahashi et al., 2003; Liu et al., 2004; de la Fuente van Bentem et al., 2005; Kaloshian and Walling, 2005; Gabriëls et al., 2006).

Another TPR-containing Hsp90-interacting co-chaperone is Sgt1. The role of Sgt1 and its interaction partner Rar1 in disease resistance has been strongly established (Austin et al., 2002; Azevedo et al., 2002; Peart et al., 2002). Several studies have demonstrated physical interaction between Hsp90, Sgt1, and Rar1 (Hubert et al., 2003; Takahashi et al., 2003; Liu et al., 2004). A role for Sgt1 as a co-chaperone in Solanaceae is illustrated by the N and Rx proteins, whose accumulation depends on Sgt1 (Azevedo et al., 2006; Mestre and Baulcombe, 2006) and the tomato Bs2 protein that requires Sgt1 for intramolecular interactions (Leister et al., 2005). Also outside the Solanaceae the combined activities of Rar1, Sgt1 and cytosolic Hsp90 are required to modulate R protein accumulation and signalling competence (Holt et al., 2005). Originally, Sgt1 was identified in yeast where it is involved in the regulation of the cell cycle and kinetochore assembly, and interacts with E3 ubiquitin ligases. Sgt1 is part of SCF E3 ubiquitin ligases via an interaction with SKP1 (Kitagawa et al., 1999). SCF E3 ligase complexes target proteins for degradation by the 26S proteasome (Willems et al., 2004). The SCF complex can associate with the COP9 signalosome (Lyapina et al., 2001; Schwechheimer et al., 2001) and interactions of COP9 and SCF subunits with either Sgt1 or Rar1 have been demonstrated in planta (Azevedo et al., 2002; Liu et al., 2002). Silencing of subunits of the SCF complex or COP9 signalosome impaired resistance mediated by N (Liu et al., 2002). Hence, besides being a chaperone, Sgt1 could be involved in R protein-mediated signalling by targeting negative regulators for degradation.

Despite extensive efforts no proteins have been identified that directly interact with the NB-ARC domain of NB-LRR proteins. This could mean that the NB-ARC is only involved in intramolecular interactions. For the N-terminal domains of NB-LRR proteins however, quite a number of interactors have been identified. This proposed “downstream signalling” domain was found to bind most of the guardees mentioned before, which supports the involvement of this domain in Avr perception. Examples of an Arabidopsis guardee that interacts with the N-terminal domain of an R protein is the Ser/Thr kinase Pbs1 that interacts with the CC domain of the Arabidopsis CNL Rps5 (Shao et al., 2003). A tertiary complex with the elicitor AvrPphB has been shown, in which cleavage of Pbs1 leads to activation of the R protein Rps5 (Ade et al., 2007; Shao et al., 2003). Another example is the interaction of Rin4 with the CC domain of the CNL proteins Rpm1 and Rps2. In addition to the interaction with Rin4, the N-terminal domains of Rpm1 and Rpp5 (CC-NB-ARC part) also interact with an ortholog of TIP49 (Holt et al., 2002). TIP49a is part of the transcriptional machinery
and interacts with the TATA binding protein (TBP) complex. Reduction of Arabidopsis TIP49a mRNA levels revealed that it does not affect Rpm1 function but it acts as a negative regulator of Rpp5 (Holt et al., 2002). Another example of indirect Avr binding is provided by the TNL protein N of tobacco, that binds the Avr protein via NRIP1 (Caplan et al., 2008). In tomato, the guardee Pto interacts not only with the SD domain of CNL Prf and with AvrPto (Mucyn et al., 2006), but also with a set of plant proteins called Pto interacting proteins or Pti’s. Pti-1 is a Ser/Thr kinase related to Pto (Zhou et al., 1995). Interestingly, Pti-4, -5 and -6 turned out to be ethylene responsive transcription factors that bind specifically to the GCC box in the promoter regions of a large number of genes encoding ‘pathogenesis-related’ (PR) proteins (Zhou et al., 1997; Gu et al., 2002).

**Figure 3**
A current model for NB-LRR R protein function and activation. (A) Direct Avr-interaction model. In the absence of a pathogen, the NB-LRR protein is in the ADP-bound resting state. Upon binding of the Avr protein, the interaction between the LRR and the ARC1 and 2 is altered (transition state), resulting in nucleotide exchange and a different conformation of the protein. This altered ATP-bound conformation represents the active state of the NB-LRR protein. Hydrolysis of ATP returns the protein into its resting state. (B) Guard model. Rather than direct binding to the Avr protein modifies a guardee bound to the N-terminus. This modification is detected by the LRR altering its interaction with the ARC1 and 2.

Although some NB-LRR proteins have been shown to associate with membranes structures, the interactions described above provide a direct link between R proteins and regulation of gene expression. An indirect link between R proteins and the nucleus is provided by Rx. Two research groups independently identified RanGAP as an interactor of the Rx N-terminus (W. I. L. Tameling and D. C. Baulcombe and M.
Sacco and P. Moffett, personal communication). In mammalian cells, RanGAP associates with RAN, a Ras-related small GTPase that has an important function in nucleo-cytoplasmic trafficking and is part of a nuclear cargo-importing complex at the nuclear envelope (Rose et al., 2004; Meier, 2006). In mammalian cells this complex is involved in the translocation of receptor-Hsp90-immunophilin complexes. The TPR-containing immunophilins (such as PP5) are required to link the heteromeric complex to a motor protein (such as dynein) for directional movement along microtubules towards the nuclear pores (Pratt et al., 2004). Interestingly, two components of the nuclear pore complex, the importin alpha homologue MOS6 (AtImpα3) and a nucleoporin 96 homologue MOS3-1, were recently identified in a suppressor screen to be required for the SNC1 phenotype (Palma et al., 2005; Zhang and Li, 2005). SNC1 is an ectopically activated NB-LRR protein that triggers constitutive activation of defence signalling (Zhang et al., 2003). A mos3-1 single mutant is compromised in both basal resistance and R protein mediated resistance (Zhang and Li, 2005).

Together, the data above imply that some R proteins may require nuclear localization for their function and could be directly involved in transcriptional regulation. Until now, nuclear localization has been reported for two TNL proteins, Rrs1 and N, which confer resistance to *Ralstonia solanacearum* in *Arabidopsis thaliana* and TMV in tobacco, respectively ((Deslandes et al., 2002) and S. P. Dinesh-Kumar, personal communication). Rrs1 is unique in that it bears a WRKY transcription factor domain at its C-terminus (Deslandes et al., 2002), providing a direct link between R protein function and transcription regulation. Activation of this R protein occurs upon binding to the Avr protein PopP2, and both proteins localize to the nucleus (Deslandes et al., 2003). On the other hand the N protein is present both in the cytoplasm and the nucleus, similar to its matching Avr protein the p50 helicase. Interestingly, nuclear localization is required for N function since equipping N and p50 with nuclear export signals reveals that disease resistance is only lost when the N protein is excluded from the nucleus and not when p50 is redirected to the cytoplasm. This observation indicates that N has a function in the nucleus, similar to RRS-1 (S. P. Dinesh-Kumar, personal communication).

An analogous type of activation was found for the prototypic animal counterpart of R proteins, the MHC class II transactivator CIITA (Steimle et al., 1993). Upon viral infection, CIITA is activated and translocated to the nucleus where it activates transcription by binding to DNA binding proteins that bind to MHC-II promoter regions (Masternak et al., 2000; Ting and Davis, 2005). Interestingly, both translocation and transactivation of CIITA require a functional NB subdomain (Harton et al., 1999). The ARC2 subdomain of CIITA adopts a winged-helix fold similar to that predicted in NB-LRR proteins (Albrecht and Takken, 2006). Such a structure is often found to be
characteristic of DNA-binding transcription factors, suggesting that both R proteins and CIITA might be directly involved in DNA binding (Albrecht and Takken, 2006). If NB-LRR proteins indeed act directly as transcriptional activators, it might explain the small number of putative downstream signalling partners that have been identified and the relatively large number of components from the transcriptional machinery among them (see below). Left aside that many putative transcription factors that bind R proteins in yeast 2-hybrid screens would have been discarded as autoactivators.

To find interactors for the membrane-spanning R proteins, the RLP class of R proteins, the cytoplasmic tail of Cf-9 has been used as bait in yeast two-hybrid screens. A thioredoxin homologue, CITRX (Cf-9 Interacting ThioRedoxin) was identified in this way as a negative regulator of Cf-9/Avr9 mediated cell death and defense responses (Rivas et al., 2004). Silencing of CITRX in tomato and N. benthamiana resulted in an accelerated Cf-9/Avr9-triggered HR accompanied by the induction of defense-related genes (Rivas et al., 2004). This interaction seems to be Cf-9 specific as no interaction was found in in vitro pull-downs with the related Cf-2 protein and Cf-2 function was not affected upon CITRX silencing (Rivas et al., 2004). Recently, the protein kinase ACIK1 (Avr9/Cf-9 Induced Kinase) was shown to interact with CITRX (Nekrasov et al., 2006). Intriguingly, silencing experiments suggested that this protein is a positive regulator of Cf-9/Avr9 function and is required for full Cf-9 disease resistance (Rowland et al., 2005).

**Partners in R protein signalling**

As detailed above, the fishing expeditions for R protein interactors thus far yielded only small numbers of components possibly involved in R protein signalling. In contrast, forward genetic screens using virus-induced gene silencing (VIGS) have been highly successful and identified a relatively large number of candidate genes. A VIGS screen of 2400 cDNAs from a normalized Nicotiana benthamiana cDNA library revealed that Pto-dependent HR was compromised in 3% of the cases (del Pozo et al., 2004). In a similar VIGS screen using 4992 cDNAs, HR was compromised in 1.6% of the cases (Lu et al., 2003). Alternative screens were performed with custom-made libraries made from cDNA-AFLP fragments that were selected based on their differential expression pattern upon Avr perception. To identify genes involved in Cf-4/Avr4-dependent HR, 192 cDNA-AFLP fragments were selected that are differentially expressed upon expression of an Avr4 transgene. These fragments were subsequently used for VIGS in N. benthamiana and twenty of them were found to correspond to genes required for Cf-4/Avr4 mediated HR (Gabriëls et al., 2006). A similar approach using Cf-9/Avr9 induced genes resulted in the identification of four genes out of the tested 43 that affect Cf-9 mediated HR (Rowland et al., 2005).
Although these VIGS screens identified relatively large numbers of genes whose silencing suppresses HR, only for a small number silencing also abolishes disease resistance. For instance, only six of 79 candidates identified in one of the Pto screens are required for Pto-mediated resistance to *P. syringae pv tabaci*, three of them derived from *Hsp90* genes (Lu et al., 2003). Likewise, of the 20 gene fragments that affect Cf-4-mediated HR, only six are essential for resistance to *Cladosporium fulvum* (Gabriëls et al., 2006).

The set of genes that is required for resistance can be roughly divided into five groups. The first consists of chaperones such as Hsp90 that have also been identified in other studies as R-protein interactors (Lu et al., 2003; Gabriëls et al., 2006). The second group has been found in each screen and consists of ribosomal proteins such as L19 (Lu et al., 2003; Gabriëls et al., 2006). The third group consist of proteins that are likely to be involved in classic signalling pathways like MAP kinase cascades (Ekengren et al., 2003) and other pathways employing protein kinases like ACIK1 (Rowland et al., 2005). The fourth group contains proteins involved in protein degradation via the 26S proteasome. Besides Sgt1, other examples are the two U-box proteins ACRE74/NICMPG1 and ACRE276 that are both ubiquitin E3 ligases (Gonzalez-Lamothe et al., 2006; Yang et al., 2006) and ACRE189, an F-box protein (Rowland et al., 2005). The fifth and last group consists of two NB-LRR proteins. The CNL NRG1 (N-requirement gene 1) appears to be specifically required for *N* function and not for other NB-LRR proteins (Peart et al., 2005). This indicates that CNL proteins may not only act as individual resistance proteins, but might also be involved in resistance mediated by TNL proteins. Similarly, the CNL NRC-1 (NB-LRR protein Required for Cf-4 function) was found to be required for Cf-4 and Cf-9 mediated resistance (Gabriëls et al., 2007). NRC-1 has also been shown to be required for HR mediated by the R proteins Prf, Rx and Mi-1.2 and for function of the RLK *Le*Eix (Gabriëls et al., 2007). The presence of more than one R protein-like factor in a single resistance pathway could indicate that resistance pathways are interwoven. Additional support for crosstalk of eLRR, TNL and CNL pathways comes from genetic studies. For instance, Cf-4 depends on EDS1, a lipase-like protein mainly required for resistance responses mediated by TNL proteins (Wiermer et al., 2005), hinting to the presence of a TNL in this pathway as well. The function of the TNL N was also found to be dependent on EDS1, while the CNL NRG1, which acts downstream of N, is not. Both are dependent on Sgt1, again showing the general importance of this protein for R signalling. Together, these data show that several resistance proteins require downstream NB-LRRs for cell death signalling (Peart et al., 2005).
**Model for R protein function**

Based on the data above a “generalised” model for the function of NB-LRR R proteins can be proposed (Figure 3). In the resting state R proteins are autoinhibited, and the NB-ARC domain interacts with both the LRR and the CC or TIR domains. This complex is molded and/or preserved in a signalling competent state by its interaction with the chaperone Hsp90 and one or more co-chaperones such as Sgt-1, Rar-1 and PP5. Perturbation of the LRR domain (either by mutations or by Avr recognition) allows transition to the active state. The Avrs can be recognized either directly, possibly by direct binding to the LRR, or indirectly by modification of a guardee that is bound to the N-terminus (CC/TIR). In the latter case, the modification of the guardee could be detected by the LRR. Such a model for indirect recognition is supported by the current data on the TNL protein N. The Avr protein (p50 helicase) is bound via an unknown protein to the TIR domain (S. P. Dinesh-Kumar, personal communication). In vitro, p50 can also interact directly with the LRR domain of N (Ueda et al., 2006). One interpretation of these data is that the LRR domain recognizes binding of p50 to the ‘guardee’ bound to the TIR domain. Conceivably, after (in)direct Avr recognition by the LRR its interaction with the ARC2 subdomain changes, resulting in a different conformation of the NB-ARC that allows nucleotide exchange (Tameling et al., 2006) or hydrolysis (Ueda et al., 2006). The exact intramolecular conformation at this stage is still unknown (e.g. dissociation of the N-terminus) and could very well differ between various R proteins (Moffett et al., 2002; Leister et al., 2005; Rairdan and Moffett, 2006; Ueda et al., 2006). The activated R protein either directly recruits downstream signalling components or could require oligomerization first, in analogy to the “wheel-of-death”. Downstream signalling components could very well be transcriptional regulators, as at least some R proteins need to be nuclearly localized for their function and one of the earliest defence signalling responses is changes in gene expression (Eulgem, 2005). Hydrolysis of ATP by the NB-ARC could return activated R proteins to the autoinhibited state. Such a mechanism would explain why hydrolysis mutants are autoactivating, and why defence signalling is only initiated when the Avr concentration in the cell is sufficiently high. Future research will put this model to the test and reveal whether it applies to all or a subset of NB-LRR proteins.

**Acknowledgements**

The authors would like to acknowledge Wladimir Tameling, David Baulcombe, Savithramma Dinesh-Kumar, Greg Rairdan, Melanie Sacco, Peter Moffett and Roger Innes for sharing their data before publication. We are grateful to Martijn Rep for providing critical review and helpful comments on this manuscript. GvO is supported by the Research Council for Earth and Life Sciences (ALW) with financial aid from
the Netherlands Organization for Scientific Research (NWO). HB is supported by Netherlands Organization for Scientific Research (NWO VENI grant 863.04.018).

References


from *Solanum bulbocastanum* confers broad spectrum resistance to potato late blight. Proc Natl Acad Sci USA 100, 9128-9133.


