Structure and function of tomato disease resistance proteins
van Ooijen, G.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
The research described in this thesis has been aimed at gaining knowledge on plant Resistance (R) protein structure and function. Resistance mediated by R proteins is dependent on the presence of a cognate Avirulence (AVR) protein in the pathogen. Because of this genetic interaction, this type of resistance was called gene-for-gene resistance (Flor, 1942). The mechanism underlying this gene-for-gene resistance is still not conclusively known. The main questions that have intrigued scientists around the world for decades (and will continue to do so) involve the process of R protein activation and the function that the different (sub)domains and conserved sequence elements play in these processes. Additionally, we know little about which other proteins are involved in orchestrating the plant R protein-mediated disease resistance mechanism. In this chapter, the results described in this thesis will be highlighted. Possible future directions and opportunities for research are presented that could help us writing the remaining blank pages in the textbook on R protein-mediated disease resistance pathways.
A functional role for the MHD motif as sensor to transduce nucleotide-dependent conformational changes

As discussed in Chapter 2, most R proteins consist of three domains (CC/TIR, NB-ARC, and LRR (Figure 1A)). The central domain is called the NB-ARC domain (van der Biezen and Jones, 1998), because of its proposed Nucleotide-Binding properties, and its presence in the human protein APAF-1 (apoptotic protease-activating factor-1), R proteins and CED-4 (Caenorhabditis elegans death-4 protein). Domains structurally related to the NB-ARC can be found in animal proteins of the NLR (nucleotide-binding domain, LRR containing) family (Ting et al., 2008). Many of these proteins act as receptors sensing intracellular perturbations.

No crystal structure has been solved for any plant NB-ARC domain, but it has been for APAF-1 (Riedl et al., 2005) and CED-4 (Yan et al., 2005). The APAF-1 NB-ARC domain consists of the nucleotide-binding (NB) subdomain and three other subdomains called ARC1, -2, and -3. In CED-4 and plant R proteins, the ARC3 subdomain is absent but the ARC1 and ARC2 subdomains are conserved (Albrecht and Takken, 2006). Specific ADP-binding by APAF-1 is achieved through interactions with various conserved residues present in the NB, ARC1, and ARC2 subdomains (Riedl et al., 2005). Because only the ADP-bound structure has been solved for APAF-1, it is not known how it senses the γ-phosphate of ATP. In CED-4, a conserved arginine in the NB subdomain sensor I region directly interacts with the γ-phosphate (Yan et al., 2005).

Numerous motifs have been identified throughout the R protein NB-ARC domain (Meyers et al., 1999; Pan et al., 2000) (Figure 1A). The functional importance of these motifs is exemplified by the many mutations of motif residues that were demonstrated to result in either loss-of-function or autoactivation of the NB-ARC-LRR protein (Chapters 3, 4, and reviewed in (Takken et al., 2006)). The MHD motif is one of the most conserved regions in R proteins (Chapter 3 and (Howles et al., 2005)). This MHD motif is located in the NB-ARC domain, at the C-terminus of the ARC2 subdomain (Figure 1A). By constructing a 3D model of the I-2 NB-ARC domain, the structural position of the MHD residues have been visualised (Chapters 3, 4, and Figure 1). As can be seen in Figure 1B, the position of the MHD histidine (H494) suggests that it directly interacts with, and positions the β-phosphate of the ADP, like it does in APAF-1 (Riedl et al., 2005; Albrecht and Takken, 2006). A direct interaction of the ARC2 subdomain with the nucleotide via this histidine is not found in related AAA+ ATPases. These proteins contain a so-called sensor II motif in the ARC1 subdomain, surrounding a conserved arginine residue that coordinates the bound nucleotide and controls intersubunit interactions (Ogura et al., 2004). APAF-1 and R proteins do not appear to share this sensor II motif, and its function might be taken
over by the MHD motif. The interaction of the MHD histidine with the ADP molecule is proposed to stabilize the ADP-bound conformation and participate in the interaction between the NB and the ARC2 subdomains. The MHD aspartate (D495) likely directly interacts with an arginine (R313) in the NB subdomain sensor I motif (Figure 1B). Mutation of this arginine in I-2 leads to a strong loss-of-function phenotype (Chapter 3).

Mutation of the 100% conserved MHD histidine or the 83% conserved aspartate would weaken interactions between the NB-ARC subdomains and between the ADP molecule and the protein. This perturbation is likely to result in destabilization of the inactive conformation, allowing transition to the active state (Chapter 3). In the cases that the MHD aspartate is not conserved, it is always replaced by another polar hydrophilic amino acid (except for the atypical rice R protein Pita), that probably does not affect the inactive conformation. Based on these observations, we propose that the MHD motif takes over both AAA+ protein sensor II functions; notably coordination of the bound nucleotide and control of intersubunit interactions.

Although the model of the plant NB-ARC domain presented in Chapter 3 has been validated by mutational analysis, it remains a model. Ultimately, to solve the R protein NB-ARC structure and dynamic changes upon activation of the defence response, crystal structures of both the ADP- and ATP-bound pockets would be essential. Although modelling leads to a good view of the inactive ADP-bound conformation, we know little about the activated, ATP-bound state. Although many people in the plant disease resistance research community seem to be convinced that these structures would be a big leap towards understanding R protein function, this goal is still not achieved. The main reason for this is probably the difficulty of obtaining sufficient amounts of properly folded R protein. Besides the opportunity to crystallize the R protein, the production of full-length R protein would also open ways to analyse the nucleotide binding and hydrolysis characteristics of (mutant) R proteins. Thus far, we have only been able to show ATPase activity for truncated I-2 and Mi-1 versions (Tameling et al., 2002) produced in *E. coli*, but we have been unable to produce sufficient amounts of properly folded full-length protein up to date (data not shown). Hopefully these experiments will be possible in the near future, and encouraging results are obtained using the yeast *Pichia pastoris* as a heterologous expression system for the production of full-length flax R protein M (Schmidt et al., 2007). Alternatively, R protein could be produced in *N. benthamiana* plants by agroinfiltration. For the subsequent purification, tandem affinity tagged R protein variants might be useful, such as the TAP-tagged Mi-1 described in Chapter 5. To boost the protein expression levels, co-infiltration with silencing suppressors could be helpful, although we observe a rapid cell death response upon expression of wild-type Mi-1 in the presence of silencing suppressors (data not shown). This might be
overcome by using a system in which HR is not induced, like a mutant background or a plant species irresponsive to the R protein. Otherwise, R protein might be produced in animal cells. Encouragingly, we have been able to detect a small amount of full-length I-2 protein produced in human cells (K.J. de Vries, unpublished data). Anyhow, at least until crystal structures of R proteins become available, the models presented in Chapters 3 and 4 will be a useful source of information to study the function of the NB-ARC domain as a molecular switch regulating R protein activity.

Figure 1
Proposed protein topology of the NB-ARC domain. (A) The position of conserved motifs in the NB, ARC1, and ARC2 subdomains is indicated in the schematic representation of a typical NB-LRR R protein. Picture adapted from Chapter 2. (B) Computationally derived 3D structure model of the NB-ARC domain of the resistance protein I-2. The model was created using the ADP bound structure of human APAF-1 (PDB code 1z6t, chain A) as structural template for I-2. The locations of R protein motifs are marked with arrows. Amino acids of the MHD motif as well as the sensor I arginine are shown in stick representation. ADP atoms are depicted as balls-and-sticks. Subdomain colouring is: NB: red, ARC1: green, ARC2: blue. Picture adapted from Chapter 3.
Intramolecular association with the LRR domain is not compromised upon activation

Domain swaps between Mi-1 (also referred to as Mi-1.2) and its parologue Mi-1.1, lead to chimaeric proteins with different characteristics. For example, incorporation of the Mi-1.2 LRR domain in a Mi-1.1 context, leads to a chimaeric protein (DS4) that strongly autoactivates (Hwang et al., 2000). These experiments initially led to a model proposing that in the repressed state, the LRR domain interacts with the N-terminus. In case of autoactivation by Mi-DS4, this interaction would be released, leading to an autoactive molecule (Hwang et al., 2000; Hwang and Williamson, 2003). The observed constitutive interaction between the CC-NB-ARC and LRR domain described in Chapter 4 shows that the initial model is oversimplified. Not only did we observe physical interactions between both autoactive (Mi-DS4) and inactive (Mi-DS2) domain combinations of Mi-1 paralogues, but also autoactivation and loss-of-function mutations in NB and ARC2 did not result in dissociation of the CC-NB-ARC and LRR domains. The ability to transcomplement and reconstitute an HR signalling-competent molecule was compromised in autoactivating MHD mutants, but not in NB autoactivation mutants or autoactive domain swaps. Likely, in the Mi-1 variants mentioned above, the functional interaction between parts of the protein is lost, even though the physical interaction between the LRR domain and its N-terminus remains. In Rx, MHD mutant D460V can still transcomplement with the LRR domain (Moffett et al., 2002). A possible reason for this difference is that Rx belongs to a different group of CC-NB-ARC-LRR proteins (Chapter 2). Further pull-down experiments with domains of different R proteins would be helpful to reveal whether subclass-specific differences are involved. Another consideration is that the effect of a given mutation might be similar, but has a different outcome in a different protein. A mutation that leads to a change in the conformation of an R protein could cause an autoactivation phenotype, whereas the analogous mutation in a different R protein could lead to inactivation (Figure 5 in Chapter 3). This could be due to a slightly more pronounced outcome of the same structural effect, leading to inactivation rather than autoactivation. Based on the modelling in Chapter 3, it is possible to point out structurally important residues, but it remains difficult to predict whether mutation would lead to loss- or to gain-of-function.

To provide more detailed information on intramolecular interactions and their dynamic changes in response to activation of Mi-1, more extensive pull-down experiments with transcomplementing protein domain combinations should be performed. Ideally, interaction between Mi-1 domains would be studied in the presence and absence of the nematode-derived stimulus that activates Mi-1 in the natural defence response. The first putative Mi-1 activating protein from nematodes was about protein MAP-1 (Semblat et al., 2001). However, co-expression of MAP-1
with Mi-1 did not induce HR, and the proteins do not interact in yeast two-hybrid analysis (S. Elzinga, unpublished data). Very recently it was reported that upon silencing of a sequence called Cg-1, virulence on Mi-1 plants was acquired in normally avirulent nematodes (Gleason et al., 2008), indicating that Cg-1 might be the factor that is recognised by Mi-1. Unfortunately, Cg-1 does not show any similarity to sequences in databases, except for a histone stem loop near the 3’ end of the transcript. There are several presumptive ORFs present in this transcript, and so far it is not clear what the gene product is. Another possibility is that the Cg-1 transcript itself is somehow responsible for the avirulence (Gleason et al., 2008). Further analyses of the Cg-1 region could contribute to reveal the basis of Mi-1 mediated nematode recognition.

RSI2 as a new member of the R protein complex

R proteins have been shown to participate in protein complexes with chaperones and chaperone-associated proteins (Chapter 5, and references herein). The LRR domain of the R proteins RPM1, N, I-2 and Rx physically interacts with Heat Shock Protein 90 (HSP90) and silencing HSP90 compromises function of a large number of NB-ARC-LRR 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; Gabriëls et al., 2007). HSP90 is a highly conserved molecular chaperone responsible for the stability and activation of a large number of signalling proteins, and is regulated by interaction with several co-chaperones (Pearl and Prodromou, 2006). HSP90-interacting co-chaperones include Protein Phosphatase 5 (PP5) (de la Fuente van Bentem et al., 2005) and SGT1, whose crystal structure was recently identified (Boter et al., 2007). Several studies have demonstrated physical interaction between HSP90, SGT1, and a protein called RAR1 (Hubert et al., 2003; Takahashi et al., 2003; Liu et al., 2004; Azevedo et al., 2006; Boter et al., 2007). Accumulation of the R proteins N and Rx depends on SGT1 and its interaction with HSP90 (Azevedo et al., 2006; Mestre and Baulcombe, 2006; Boter et al., 2007). Generally, the R protein complex members RAR1, SGT1, and HSP90 are required for R protein accumulation and signalling competence (Boter et al., 2007).

In Chapter 5 we have demonstrated physical interaction of the I-2 disease resistance gene product with RSI2, a member of the family of small heat shock proteins (sHSPs). RSI2 is also highly important in maintaining I-2 protein stability, and thereby also in initiation of HR by I-2. These results strongly establish RSI2 as a new type protein involved in R protein function. sHSPs share the three-dimensional structure fold of p23, an important regulator of HSP90 in mammals. In humans, p23 participates in the folding of a plethora of proteins (Bose et al., 1996). At its N-terminus, HSP90 has an ATP binding- and hydrolysis domain. P23 regulates HSP90
activity by direct interaction with the ATP-bound active form of HSP90 (Johnson et al., 1994), inhibiting ATPase activity. This system appears to be important for tight regulation of HSP90 functions. The central CS domain in SGT1 shows structural similarity to p23 (Boter et al., 2007). Despite this structural similarity, their mode of interacting with HSP90 is probably distinct since the CS domain of SGT1 does not share the strand of p23 that has been shown to interact with HSP90 (Boter et al., 2007). Furthermore, SGT1 has been shown not to affect HSP90 ATPase activity, unlike p23 and most other HSP90 co-chaperones.

The observation that loss of RSI2 adversely affects I-2 stability (Chapter 5), could possibly be due to reduced chaperoning activity by the HSP90/SGT1 machinery. The presumed common phylogenetic origin of p23 and HSP20 proteins (Garcia-Ranea et al., 2002), and the experimentally verified chaperoning activities of both proteins points towards a related functional behaviour as a co-chaperone of HSP90 and regulator of its activity. Since HSP90 is present both in plants and in yeast, the interaction we observe between RSI2 and I-2 could be mediated by HSP90. However, the interaction surface of HSP90 in the I-2 LRR maps to LRR1-11 and the interaction with RSI2 is mediated by LRR15-19. Possibly, RSI2 interacts with both its client I-2 and with HSP90, similar to HSP90 co-chaperone and I-2 interacting PP5 (de la Fuente van Bentem et al., 2005). Further experimentation into this specific class of small heat shock proteins could reveal whether they indeed act as ATP-dependent regulators of HSP90 chaperone activity in analogy to P23. Members of this class of proteins could be conserved factors maintaining integrity and signalling competence of the plant NB-ARC-LRR immune receptors.

**How do R proteins recognise pathogens?**

The molecular mechanism by which R proteins recognise pathogen attack is not conclusively known. The most simple interpretation of the gene-for-gene interaction would be a receptor-ligand interaction between the R and AVR proteins. This idea is supported by the fact that most R-AVR couples co-localize in the plant cell and that R proteins have an LRR domain that functions as an interaction interface in several receptors. Indeed, a number of cases are known in which the AVR protein directly interacts with the R protein. A canonical example is the direct interaction underlying flax rust resistance (Dodds et al., 2006; Ellis et al., 2007). Direct interaction has also been shown between Pita/AvrPita in rice (Jia et al., 2000) and PopP2/RRS-1 in Arabidopsis (Deslandes et al., 2003).

In contrast to the examples of direct interaction between R proteins and their cognate AVR proteins described above, direct interaction was not observed in a lot of other cases (Chapter 2). An additional mechanism has been proposed to be responsible for initiation of R protein-mediated responses to the invading pathogen. Most AVR
proteins probably act as effector proteins that target host cell components, and R proteins have been reported to sense the presence of a pathogen by monitoring the state of host targets attacked by AVR proteins. This model is referred to as the guard hypothesis (Van der Biezen and Jones, 1998). Examples in which the AVR protein does not directly interact with the R protein, include the following. (1) Phosphorylation or cleavage of Arabidopsis RIN4 protein by AvrB, AvrRpm1, and AvrRpt2 is monitored by, and activates R proteins Rpm1, Rps2, and Tao1 (Mackey et al., 2002; Kim et al., 2005; Eitas et al., 2008); (2) The Arabidopsis PBS1 kinase is inactivated by the avrPphB protease, which in turn activates RPS5 (Shao et al., 2003; Ade et al., 2007); (3) The interaction between the Pseudomonas syringae AvrPto or AvrPtoB effectors and the tomato kinase Pto is sensed and responded to by NB-ARC-LRR protein Prf (Kim et al., 2002; Mucyn et al., 2006; Xing et al., 2007); (4) Tomato cysteine protease Rcr3 is inhibited by interaction with AVR2 from Cladosporium fulvum. Tomato transmembrane R protein Cf-2 is triggered by the Rcr3-AVR2 complex (Rooney et al., 2005); (5) The helicase domain (p50) of Tobacco mosaic virus (Burch-Smith et al., 2007) interacts with the tobacco protein NRIP1. This protein is normally localized to the chloroplasts, but upon interaction with p50 it localizes to the cytoplasm and nucleus (Caplan et al., 2008), where it is recognised by the R protein N; (6) Resistance to Xanthomonas campestris in pepper, conferred by the Bs3 gene, is based on binding of AvrBs3 to the Bs3 promoter region. The AVR protein probably activates this promoter (Römer et al., 2007).

Originally, the general assumption was that the R protein waits until its guardee is attacked and inactivated by pathogen effectors. In the absence of the corresponding R protein, this inactivation of a host protein would be beneficial to the pathogens virulence. Alternatively, rather than being virulence targets, the proteins that are guarded by R proteins evolved especially to be attacked by pathogen effectors and initiate defence signalling. This would for example explain the fact that Pto is not necessary for the contribution of AvrPto towards virulence, and neither is RIN4 for AvrRpt2 and AvrRpm1 virulence functions. This idea would also solve the problem of opposing evolutionary selection on the guardee; without the cognate R protein, the guardee would evolve to avoid interaction with the effector to decrease virulence, whereas on the other hand the plant would benefit from a strong interaction in the presence of the R protein (R.A.L. van der Hoorn and S. Kamoun, personal communication).

For the R proteins I-2 and Mi-1, used as model proteins for the studies described in this thesis, we do not know much about the mechanism in which they recognise pathogen attack. In the near future, more detailed studies will be possible since very recently, candidate AVR proteins have been identified both for Mi-1 (Gleason et al., 2008) and for I-2 (P.H. Houterman and M. Rep, unpublished data).
targets for these effectors in plants (whether it is the R protein or a different plant protein) will open doors to investigate the molecular mechanism underpinning I-2 and Mi-1 mediated resistance. The emerging picture is that there is a large diversity in the way that R proteins recognise pathogen attack, since R protein recognition specificities likely evolve at random. In all cases, an important clue is to identify the plant protein(s) targeted by the cognate AVR protein, since this interaction is the frontline of the race-specific battle between the plant and the pathogen.

References


