

Structure, function and subcellular localization of the potato Resistance protein Rx1

Erik J. Smitweg

Thesis committee

Thesis supervisor

Thesis co-supervisors

Prof. dr. ir. J. Bakker, Professor of Nematology
Wageningen University

Dr. ir. A. Schots, Associate professor, Laboratory of Nematology,
Wageningen University

Dr. ir. A. Goverse, Assistant professor, Laboratory of Nematology,
Wageningen University

Other members:

Prof. dr. F.P.M. Govers, Wageningen University

Prof. dr. P. Schulze-Lefert, Max-Planck-Institut für Züchtungsforschung, Köln, Germany

Dr. ing. F.L.W. Takken, University of Amsterdam

Prof. dr. V.M. Williamson, University of California, Davis, Verenigde Staten

This research was conducted under the auspices of the Graduate School of Experimental
Plant Sciences.

Structure, function and subcellular localization of the potato Resistance protein Rx1

Erik J. Slootweg

Thesis
submitted in partial fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. dr. M.J. Kropff
in the presence of the
Thesis Committee appointed by the Doctorate Board
to be defended in public
on Friday 23 October 2009
at 1:30 PM in the Aula.

Structure, function and subcellular localization
of the potato Resistance protein Rx1

Erik J. Slootweg

Thesis Wageningen University, The Netherlands, 2009

With summaries in English and Dutch

ISBN 978-90-8585-467-8

Contents

Chapter 1	1
General Introduction	
Chapter 2	11
Fluorescent T7 display phages obtained by translational frameshift	
Chapter 3	31
T7 cDNA phage display identifies highly basic (poly)-peptides as Rx1 CC-NB-ARC interactors	
Chapter 4	53
Domain exchange between Rx1 and Gpa2 in potato reveals flexibility of CC-NB-LRR genes to switch between virus and nematode resistance	
Chapter 5	73
A docking model for the NB-ARC and LRR domains of the CC-NB-LRR Resistance protein Gpa2	
Chapter 6	107
Nuclear localization of the NB-LRR Resistance protein Rx1 is dependent on coiled-coil domain, ATP-binding, SGT1 and Rar1, and its activation is triggered in the cytoplasm	
Chapter 7	141
General discussion	
Summary	159
Samenvatting	163
Acknowledgments	169
Publications	171
Curriculum vitae	173

Chapter 1

General Introduction

Plant defence

Plants have the ability to use energy from sunlight to convert carbon dioxide and water into complex organic molecules. These energy-rich molecules are an attractive source of nutrients for organisms that cannot synthesize their own food. Probably from the very beginning, defence against other organisms was an important factor in the life of plants. All plant species that we see today were successful enough in defending themselves against attackers to survive throughout evolution. Those that could not, have gone extinct. The same, of course, holds true for the herbivores and plant pathogens. The ones we see today are the ones that successfully adapted, overcame the plant's defence system and thereby secured themselves of rich food sources. This ancient evolutionary struggle has equipped plants with an extensive arsenal of defence mechanisms and their cognate pathogens with an equally wide range of strategies to counter these (Bent and Mackey, 2007; da Cunha et al., 2007).

Obviously, plants cannot run for danger, unlike animals and humans. Therefore, their defence is mostly autonomous and very effective, even at single cell level. A first barrier against invading pathogens and herbivores is formed by physical structures like spines, trichomes, wax layers and the cell wall, and by molecules like toxins or proteases. Plants are encountering many (micro)organisms in their natural environment, both above and below ground. But the vast majority of the pathogens and pests among them are not able to invade the plant. Natural variation in these passive and generic defences between plant species forces the pathogens to specialize their life strategies towards certain host species, which appeared to be susceptible to them.

In addition to the more-or-less passive physical and chemical barriers, plants also possess responsive immune systems that allow them to activate defence mechanisms when sensing the presence of a pathogen. Despite some similarities, this immune system is significantly different from the adaptive immune system found in vertebrates. Plants lack a circulatory system and specialized mobile immune cells like leukocytes. Each plant cell harbours in principle the full repertoire of recognition specificities to detect pathogen derived molecules. Secondly, the plants repertoire of recognition specificities is not adaptive, like antibodies that diversify by somatic mutation and recombination, but is passed genetically from generation to generation.

Two evolutionary and functionally distinct recognition systems are found in plants. A relatively ancient system of receptors, which is activated by generic pathogen associated elicitors (Pathogen/Microbe associated molecular patterns: PAMPs/MAMPs) (He et al., 2007). These elicitors can be pathogen-derived molecules like flagellins, chitin or lipopolysaccharides, but also host plant-derived molecules released or modified by the pathogen, like degradation products of the cell wall. The common characteristic of these elicitors is that they are evolutionary stable; bacteria for example can not without high costs alter the structure of the lipopolysaccharides found in their outer membrane. A similar PAMP recognition system exists in vertebrates in the form of the non-adaptive innate immune system (Ausubel, 2005).

R gene mediated resistance

Although pathogens cannot easily escape this basal immune recognition, they have evolved strategies to evade the consequences. Many pathogens can deliver proteins into the host cell that target key positions in the cell's basic immune pathways and thereby suppress the defence response. These so-called effector proteins come in many varieties. Some are proteases or ubiquitin-ligases that specifically degrade host proteins. Others function as transcription factors, modify the translation mechanism or mimic plant hormones. Besides effectors targeting the basal immune response (PAMP induced immunity, PTI), pathogens also deploy effectors that modify the host cell to increase its value for the pathogen's feeding or reproduction (Janjusevic et al., 2006; da Cunha et al., 2007; Kay et al., 2007; Saijo and Schulze-Lefert, 2008).

The plants evolutionary response to the pathogens tampering with the PAMP triggered immunity and cellular machinery, are the resistance proteins (R proteins); receptor-like proteins that recognize either directly or indirectly the pathogens effectors (Jones and Dangl, 2006). In contrast to the evolutionary stable and relatively broad specificity of the PAMP recognition, R proteins respond very specific to one or a few effectors. In genetics, this phenomenon manifests itself as a gene-for-gene interaction; the presence of a single dominant resistance gene in the plant and the presence of a so-called avirulence gene in the corresponding pathogen strain determine if the plant is resistant to that pathogen (Flor, 1971). Although resistance loci were known in plant breeding for a long time, the first glimpse into the underlying mechanisms came after the cloning of several R genes about 15 years ago. With the sequencing of complete plant genomes, it became clear that plants harbour hundreds of such resistance gene homologs, for most of which no specificity is known. Many of these genes are found in large clusters or so called hot spots for resistance. Coevolution between plants and pathogens has resulted in the expansion of these R gene clusters by diversifying selection and gene conversion (Mondragon-Palomino et al., 2002; Cork and Purugganan, 2005; Mondragon-Palomino and Gaut, 2005).

Structure and function of R genes

R genes encode proteins composed of several distinct functional domains and depending on their modular architecture, at least five different classes are distinguished (Hammond-Kosack and Jones, 1997). The most prevalent type of R protein, the NB-LRRs or Nibblers, contain a central nucleotide-binding domain module, that is similar to the nucleotide-binding domain in several metazoan apoptosis regulating proteins; Apaf-1 in mammals and CED-4 in *Caenorhabditis elegans*, and is therefore called the NB-ARC domain (Apaf-1, R protein, CED-4) (van der Biezen and Jones, 1998a; Takken et al., 2006). A leucine-rich repeat (LRR) domain is usually positioned C-terminally to the NB-ARC. The LRR forms a horseshoe-like structure characterized by a parallel β -sheet on its concave surface (Kobe and Kajava, 2001). Regularly patterned leucines in this β -sheet point inwards to the hydrophobic core of the structure and are tightly packed in a stabilizing scaffold. The residues in between these leucines are often hydrophilic and are exposed to the solvent. The large non-globular conformation maximizes the potential interaction surface and makes this domain well suited for protein interactions. The LRR is thought to be involved both in intramolecular inhibitory interactions with the N-terminal R protein domains and in determining the specificity of pathogen recognition. Based on the identity of the N-terminal domain two

main classes of NB-LRR R proteins can be distinguished. The TIR-NB-LRRs carry a domain also found in the metazoan Toll-like receptors called the Toll/Interleukin-1 Receptor domain (TIR) (Burch-Smith and Dinesh-Kumar, 2007). The CC-NB-LRRs contain a loosely defined coiled coil domain, which can be formed by intertwined alpha-helices (Lupas, 1997). Smaller classes of NB-LRRs have been found with alternative N-terminal domains, like the Solanaceae specific domain (SD) or the DNA-binding BED finger domain (Kohler et al., 2008; Mucyn et al., 2009).

The structural mechanism by which these proteins function seems to centre on the conformation of the NB-ARC domain. This domain is thought to act like a molecular switch in which the activation state is determined by the identity of the nucleotide bound to it (Takken et al., 2006). Some R proteins have been shown to bind and hydrolyse ATP (Tameling et al., 2002). Based on the fact that mutants unable to hydrolyse ATP, but not affected in ATP-binding, show a constitutively active phenotype, it is hypothesized that the ATP-bound form represents the active state and the ADP-bound form represents an inactive, resting state of the R protein (Tameling et al., 2006). Such mechanism is analogous to the functioning of related NB containing proteins from the STAND (Signal Transduction ATPases with Numerous Domains) class of AAA+ ATPases (Bao et al., 2007; Faustin et al., 2007; Lukasik and Takken, 2009). How exactly the recognition of the pathogen leads to an ADP for ATP exchange and the restructuring of the protein to its signalling state is unknown. The intramolecular interactions to the NB-ARC involving the LRR and N-terminal domains have been shown to have inhibitory functions, which might be released by the recognition event, allowing the nucleotide exchange. Based on the structures of the NB-ARC containing apoptosis regulators Apaf-1 and CED-4 it is thought that the ADP bound structure is very compact with the NB and ARC closed around the ADP, and that in the ATP bound state the NB and ARC domains dramatically change in position and orientation towards an open conformation (Riedl et al., 2005; Yan et al., 2005). This open conformation exposes new interaction surfaces which could bind downstream signalling components or direct oligomerization, as seen for many related proteins (Takken and Tameling, 2009).

The recognition mechanism itself has been puzzling for several reasons. The LRR domain is the most obvious domain involved in recognition, and patterns of diversifying selection seen for the exposed residue positions support a mechanism of direct recognition (Mondragon-Palomino et al., 2002). On the other hand, direct interaction between the LRR and pathogen effectors has rarely been reported and the number of potential pathogens does not seem to be reflected in the variety of NB-LRRs the plant has to its disposition. The guard-theory describes an alternative recognition mechanism by which the NB-LRR proteins guard key protein complexes in the plant cell and detect the modifications made to these host proteins by the pathogens effectors (van der Biezen and Jones, 1998b; de Wit, 2002). Taking this theory further, one can imagine the rise of look-a-likes of the guarded proteins that function as bait for pathogen effectors, but are less limited in their evolution by the functional constraints the original guard-ee has (van der Hoorn and Kamoun, 2008). Several R protein complexes that have been described in recent years fit the guard theory (Mackey et al., 2002; Axtell and Staskawicz, 2003; Ade et al., 2007; Caplan et al., 2008), and support the view that this a common recognition mechanism.

The potato resistance genes Rx1 and Gpa2

Gpa2 and Rx1 are two highly homologous CC-NB-LRR R proteins (88% on similar in amino acid sequence) encoded in a small R gene cluster on chromosome XII of potato (van der Vossen et al., 2000). Despite their sequence similarity they confer resistance to pathogens with completely different lifestyles. Rx1 confers a relatively durable resistance against the ssRNA plant virus Potato Virus X (PVX). The resistance response mediated by Rx1 upon recognizing the PVX coat protein (CP) is highly effective in blocking virus replication and does usually not result in a visible hypersensitive cell death response (HR), therefore it is described as extreme resistance (ER). This response includes the inhibition of the translation of viral RNAs by the host's ribosomes (Bhattacharjee et al., 2009). Under specific circumstances, for example when CP accumulation is uncoupled from viral replication, Rx1 does activate a strong cell death response (Bendahmane et al., 2000). Gpa2 confers resistance against the potato cyst nematode *Globodera pallida*. Even on resistant plants the nematode initially is able to establish a feeding site. However, the feeding site does not fully develop and in later stages of development it is cut off from living tissue by necrosis of the surrounding cells.

It is surprising that the highly similar R proteins Gpa2 and Rx1 can raise these, on first sight, so dissimilar resistance responses against two distinct pathogens. One point of similarity between these two obligatory biotrophic pathogens is the sophisticated way they utilize the host cell's machineries to accommodate the cell completely to their own advantage. The potato cyst nematode uses its stylet to inject a large variety of effector proteins into a carefully selected root cell in the cortex, which expands towards the vascular system by progressive cell wall dissolution (Davis et al., 2008). These cells become multinucleate and metabolically highly active nutrient sinks, on which the sedentary endoparasitic nematode fully depends for its development and reproduction.

Compared with the variety of effectors the nematode can apply, the five proteins encoded by the PVX genome seem rather meagre. However, they compensate the lack in quantity with multifunctionality of the viral proteins (Verchot-Lubicz et al., 2007). The replicase, encoded by ORF1, is required for the replication of the viral genomic RNA, maybe with the help of host proteins. In this protein several functional domains are combined; a methyltransferase-like domain, a NTP-binding/helicase-like domain and a RNA dependent RNA polymerase domain (Draghici et al., 2009). The protein encoded by the first triple gene block gene (TGBp1) is required for cell-to-cell movement, modifies the plasmodesmata to allow larger particles pass through and was even shown to be involved in inhibiting a post-transcriptional gene-silencing pathway which is one of the plant's defence mechanisms against virus replication (Voinnet et al., 2000; Batten et al., 2003; Howard et al., 2004). With the coat protein (CP) it participates in RNA translation activation (Zayakina et al., 2008). The coat protein encapsulates the ssRNA genome into virions, but is also needed for the cell-to-cell movement of the virus and systemic infection in the host (Chapman et al., 1992; Fedorkin et al., 2001). Further insight into the tasks the viral proteins, and especially the CP, fulfil and with which host components they cooperate would help to understand how Rx1 senses the presence of the coat protein.

Rx1 has been the focus of many functional studies in the past years. The availability of its elicitor, the fact that it is functional in transient expression systems, the knowledge about many orthologous and paralogous genes make it an ideal model for R protein functioning. The structural domains of Rx1 interact with each other and even when expressed from separate genes, *in trans*, the domains are capable of forming a fully functional protein (Moffett et al., 2002). Both mutagenesis, deletion studies and sequence exchanges with Gpa2 have revealed the presence of inhibitory intramolecular interactions between the LRR and N-terminal domains and the role of the LRR in recognition (Bendahmane et al., 2002; Moffett et al., 2002; Rairdan and Moffett, 2006). Some mutations in the LRR even resulted in a broadening of the recognition specificity to include evolutionary more distantly related viruses (Farnham and Baulcombe, 2006). In two independent studies an interaction between the Rx1 and Gpa2 CC domain and the RanGTPase activating protein 2 (RanGAP2) was found. RanGAP2 overexpression makes Rx1 more sensitive for (auto)-activation and RanGAP2 silencing disrupts Rx1 mediated PVX resistance. The exact role of this interaction is not known, but it could be involved in downstream signalling or it could be a protein guarded by Rx1 and Gpa2 as a potential target for pathogen effectors. Until recently no elicitor was known for Gpa2 and its functioning could only be studied in laborious nematode assays. The discovery of RanBPM-like proteins secreted by *G. pallida* that specifically activate Gpa2 opens new possibilities for functional assays.

Outline of the thesis

The work described in this thesis aimed at gaining insight in the functioning of the resistance proteins Rx1 and Gpa2, and with the expectation that this insight would be applicable to a wider range of resistance proteins sharing similar structural and functional characteristics.

In Chapter 2 it is shown that the lytic T7 phage display system can be modified to include fluorescent protein fusions in its capsid in addition to capsid protein displaying peptides and unmodified capsid proteins. This was accomplished by means of a translation frameshift sequence, which in 10% of the translation events resulted in the shift to a second reading frame in which a yellow fluorescent protein was encoded. The idea behind creating fluorescently labelled phage applicable in phage display was to extend phage displays applications to more complex library screenings or even automated screenings.

Selecting Rx1 interactors from a tobacco cDNA library proved more difficult. In Chapter 3 cDNA phage display was explored as technique to discover proteins interacting with Rx1. The system turned out to be especially prone to picking up interactors to binding matrices like Ni-NTA or to fusion proteins like thioredoxine. This dilemma was circumvented by designing the selection procedure in such a way that it alternates between different matrices and to limit the number of selection round. Limiting the number of selections in the enrichment of the library means that selecting individual interactors from the enriched library becomes more laborious.

The proteins and peptides identified to interact with the Rx1 CC-NB-ARC had as common characteristic that they were highly basic. Several of these were shown in pull-down experiments to bind specifically to Rx1 and not to the fusion protein or to the binding matrix. The finding of a wide array of different highly basic protein fragments and peptides suggests that the CC-NB-ARC, with its netto negative charge and highly acidic regions in the CC and ARC2 domain, has the tendency to participate in interactions based on electrostatic charges. The structural modelling of the interaction between the Gpa2 NB-ARC and LRR domains, described in Chapter 5, suggested that precisely such an electrostatic interaction might play an important role in the docking of these domains. A highly acidic loop on the surface of the ARC2 domain is positioned in close proximity to a patch of basic residues on the conserved side of LRRs 5 to 7.

Recognition specificity and the concentration dependency of elicitor-dependent activation or autoactivation were studied in Chapter 4 by means of the exchange of sequence fragments between Gpa2 and Rx1. The incompatibility between the CC-NB-ARC and LRR regions of Rx1 and Gpa2 led to an autoactivation of the chimeric protein consisting of the CC-NB-ARC region of Gpa2 and the LRR of Rx1, and to loss-of-function in the construct consisting of the CC-NB-ARC region of Rx1 and the LRR region of Gpa2. Strongly lowering the expression level of constitutively active chimera abolished the autoactivity and resulted in specific PVX resistance in transgenic potato plants. The activity of the chimeric construct with the Gpa2 LRR could be restored by re-introducing the first five LRRs of Rx1 in the LRR region, restoring the compatibility between the N-terminal part of the LRR and the ARC2 domain. This rescued resistance against *Globodera pallida*. Recognition of the Rx1 CP by Rx1 was

shown to be dependent on a region of the LRR overlapping with, but not identical to the LRR region in Gpa2 required for specific recognition of the Gpa2 elicitor.

A more detailed study of the domain incompatibilities between the ARC and LRR domains was used in Chapter 5 to determine functional interactions between the ARC and LRR regions as input for a docking model of the NB-ARC and LRR 3D structures. It was assumed that a physical interaction underlies the functional interaction. A region differing in 7 amino acids between Rx1 and Gpa2 in the ARC2 domain was the main determinant for ARC2 to LRR compatibility. The LRR region that needs to match the ARC2 encompasses the most N-terminal 3 repeats. In the Gpa2 NB-ARC/LRR docking model, the interacting regions of the NB, ARC2 and LRR are shown to coevolve. Strong electrostatic interactions via conserved charged surface areas and two hydrophobic patches characterize the binding interface. Support for this docking model was found by site-directed mutagenesis aimed at altering the hydrophobic patch and clusters of charged residues.

Chapter 6 focuses on the subcellular localization of Rx1 and the role its functional domains play in the observed nucleo-cytoplasmic distribution. Recent studies have shown that several R proteins are dependent on a nuclear localization for their signalling and that some interact with nuclear factors like transcription factors. By constructing fluorescent versions of Rx1 we found that Rx1 too was localized in both the cytoplasm and the nucleus. This came somewhat unexpected because no nuclear localization signal is predicted for Rx1 (or Gpa2) and its mass (140 kDa as GFP fusion) is well above the limit for passive transport through the nuclear pore (40-60 kDa). To determine how each functional domain contributed to the dual localization of Rx1 a series of deletion constructs and single domain fusions to GFP were made. The CC and more in specific a helix-rich fragment of the CC, showed a strong nuclear localization, associated with a slow nuclear diffusion. The LRR and NB-ARC-LRR on the other hand were mostly cytoplasmic. Both silencing SGT1 and the inhibition of ATP-binding reduced the nuclear localization of Rx1, but not of a truncated version of Rx1 without the LRR. Enforcing a nuclear or cytoplasmic localization for Rx1 via nuclear localization or nuclear export sequences did not influence its activity. However, nuclear targeting of the Rx1 elicitor, the PVX CP did inhibit recognition.

The role of intermolecular interaction between NB-LRR protein subdomains, nucleotide binding and exchange and the conformational states of NB-LRR R proteins are discussed in Chapter 7. Structural and functional similarities with other proteins in the STAND class of signalling ATPases point to a conservation of the molecular switch mechanism of the NB-ARC domain. However, the regulation and signalling of this core NB-ARC is conducted by mechanistically diverse mechanisms.

References

- Ade, J., Deyoung, B.J., Golstein, C. and Innes, R.W.** (2007). Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease. *Proc Natl Acad Sci U S A* **104**, 2531-2536.
- Ausubel, F.M.** (2005). Are innate immune signaling pathways in plants and animals conserved? *Nat Immunol* **6**, 973-979.
- Axtell, M.J. and Staskawicz, B.J.** (2003). Initiation of RPS2-specified disease resistance in Arabidopsis is coupled to the AvrRpt2-directed elimination of RIN4. *Cell* **112**, 369-377.
- Bao, Q., Lu, W., Rabinowitz, J.D. and Shi, Y.** (2007). Calcium blocks formation of apoptosome by preventing nucleotide exchange in Apaf-1. *Mol Cell* **25**, 181-192.
- Batten, J.S., Yoshinari, S. and Hemenway, C.** (2003). Potato virus X: a model system for virus replication, movement and gene expression. *Molecular Plant Pathology* **4**, 125-131.
- Bendahmane, A., Farnham, G., Moffett, P. and Baulcombe, D.C.** (2002). Constitutive gain-of-function mutants in a nucleotide binding site-leucine rich repeat protein encoded at the Rx locus of potato. *Plant J.* **32**, 195-204.
- Bendahmane, A., Querci, M., Kanyuka, K. and Baulcombe, D.C.** (2000). Agrobacterium transient expression system as a tool for the isolation of disease resistance genes: application to the Rx2 locus in potato. *Plant J.* **21**, 73-81.
- Bent, A.F. and Mackey, D.** (2007). Elicitors, effectors, and R genes: the new paradigm and a lifetime supply of questions. *Annu Rev Phytopathol* **45**, 399-436.
- Bhattacharjee, S., Zamora, A., Azhar, M.T., Sacco, M.A., Lambert, L.H. and Moffett, P.** (2009). Virus resistance induced by NB-LRR proteins involves Argonaute4-dependent translational control. *Plant J* **58**, 940-951.
- Burch-Smith, T.M. and Dinesh-Kumar, S.P.** (2007). The functions of plant TIR domains. *Sci STKE* **2007**, pe46.
- Caplan, J.L., Mamillapalli, P., Burch-Smith, T.M., Czymbek, K. and Dinesh-Kumar, S.P.** (2008). Chloroplastic Protein NRIP1 Mediates Innate Immune Receptor Recognition of a Viral Effector. *Cell* **132**, 449-462.
- Chapman, S., Hills, G., Watts, J. and Baulcombe, D.** (1992). Mutational analysis of the coat protein gene of potato virus X: effects on virion morphology and viral pathogenicity. *Virology* **191**, 223-230.
- Cork, J.M. and Purugganan, M.D.** (2005). High-diversity genes in the Arabidopsis genome. *Genetics* **170**, 1897-1911.
- da Cunha, L., Sreerekha, M.V. and Mackey, D.** (2007). Defense suppression by virulence effectors of bacterial phytopathogens. *Curr Opin Plant Biol* **10**, 349-357.
- Davis, E.L., Hussey, R.S., Mitchum, M.G. and Baum, T.J.** (2008). Parasitism proteins in nematode-plant interactions. *Curr Opin Plant Biol* **11**, 360-366.
- de Wit, P.** (2002). Plant biology - On guard. *Nature* **416**, 801-803.
- Draghici, H.K., Pilot, R., Thiel, H. and Varrelmann, M.** (2009). Functional mapping of PVX RNA-dependent RNA-replicase using pentapeptide scanning mutagenesis-Identification of regions essential for replication and subgenomic RNA amplification. *Virus Res* **143**, 114-124.
- Farnham, G. and Baulcombe, D.C.** (2006). Artificial evolution extends the spectrum of viruses that are targeted by a disease-resistance gene from potato. *Proc Natl Acad Sci U S A* **103**, 18828-18833.
- Faustin, B., Lartigue, L., Bruey, J.M., Luciano, F., Sergienko, E., Bailly-Maitre, B., Volkman, N., Hanein, D., Rouiller, I. and Reed, J.C.** (2007). Reconstituted NALP1 inflammasome reveals two-step mechanism of caspase-1 activation. *Mol Cell* **25**, 713-724.
- Fedorkin, O.N., Solovyev, A.G., Yelina, N.E., Zamyatnin, A.A., Zinovkin, R.A., Makinen, K., Schiemann, J. and Morozov, S.Y.** (2001). Cell-to-cell movement of potato virus X involves distinct functions of the coat protein. *J. Gen. Virol.* **82**, 449-458.
- Flor, H.H.** (1971). Current status of the gene-for-gene concept. *Annu. Rev. Phytopathol.* **9**, 297-323.
- Hammond-Kosack, K.E. and Jones, J.D.G.** (1997). Plant disease resistance genes. In *Annu. Rev. Plant Physiol. Plant Molec. Biol.*, pp. 575-607.
- He, P., Shan, L. and Sheen, J.** (2007). Elicitation and suppression of microbe-associated molecular pattern-triggered immunity in plant-microbe interactions. *Cell Microbiol* **9**, 1385-1396.
- Howard, A.R., Heppler, M.L., Ju, H.J., Krishnamurthy, K., Payton, M.E. and Verchot-Lubicz, J.** (2004). Potato virus X TGBp1 induces plasmodesmata gating and moves between cells in several host species whereas CP moves only in *N. benthamiana* leaves. *Virology* **328**, 185-197.
- Janjusevic, R., Abramovitch, R.B., Martin, G.B. and Stebbins, C.E.** (2006). A bacterial inhibitor of host programmed cell death defenses is an E3 ubiquitin ligase. *Science* **311**, 222-226.
- Jones, J.D. and Dangl, J.L.** (2006). The plant immune system. *Nature* **444**, 323-329.
- Kay, S., Hahn, S., Marois, E., Hause, G. and Bonas, U.** (2007). A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science* **318**, 648-651.
- Kobe, B. and Kajava, A.V.** (2001). The leucine-rich repeat as a protein recognition motif. *Curr. Opin. Struct. Biol.* **11**, 725-732.
- Kohler, A., Rinaldi, C., Duplessis, S., Baucher, M., Geelen, D., Duchaussoy, F., Meyers, B.C., Boerjan, W. and Martin, F.** (2008). Genome-wide

identification of NBS resistance genes in *Populus trichocarpa*. *Plant Mol Biol* **66**, 619-636.

Lukasik, E. and Takken, F.L. (2009). STANDING strong, resistance proteins instigators of plant defence. *Curr Opin Plant Biol* **24**, 24.

Lupas, A. (1997). Predicting coiled-coil regions in proteins. *Curr. Opin. Struct. Biol.* **7**, 388-393.

Mackey, D., Holt, B.F., 3rd, Wiig, A. and Dangl, J.L. (2002). RIN4 interacts with *Pseudomonas syringae* type III effector molecules and is required for RPM1-mediated resistance in *Arabidopsis*. *Cell* **108**, 743-754.

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

Mondragon-Palmino, M. and Gaut, B.S. (2005). Gene conversion and the evolution of three leucine-rich repeat gene families in *Arabidopsis thaliana*. *Mol. Biol. Evol.* **22**, 2444-2456.

Mondragon-Palmino, M., Meyers, B.C., Michelmore, R.W. and Gaut, B.S. (2002). Patterns of positive selection in the complete NBS-LRR gene family of *Arabidopsis thaliana*. *Genome Res* **12**, 1305-1315.

Mucyn, T.S., Wu, A.J., Balmuth, A.L., Arasteh, J.M. and Rathjen, J.P. (2009). Regulation of tomato Prf by Pto-like protein kinases. *Mol Plant Microbe Interact* **22**, 391-401.

Rairdan, G.J. and Moffett, P. (2006). Distinct Domains in the ARC Region of the Potato Resistance Protein Rx Mediate LRR Binding and Inhibition of Activation. *Plant Cell* **18**, 2082-2093.

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

Saijo, Y. and Schulze-Lefert, P. (2008). Manipulation of the Eukaryotic Transcriptional Machinery by Bacterial Pathogens. *Cell Host & Microbe* **4**, 96-99.

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

Takken, F.L.W. and Tameling, W.I.L. (2009). To Nibble at Plant Resistance Proteins. *Science* **324**, 744-746.

Tameling, W.I.L., Elzinga, S.D.J., Darmin, P.S., Vossen, J.H., Takken, F.L.W., Haring, M.A. and Cornelissen, B.J.C. (2002). The tomato R gene products I-2 and Mi-1 are functional ATP binding proteins with ATPase activity. *Plant Cell* **14**, 2929-2939.

Tameling, W.I.L., Vossen, J.H., Albrecht, M., Lengauer, T., Berden, J.A., Haring, M.A., Cornelissen, B.J.C. and Takken, F.L.W. (2006). Mutations in the NB-ARC Domain of I-2 That Impair ATP Hydrolysis Cause Autoactivation. *Plant Physiol.* **140**, 1233-1245.

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

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

van der Hoorn, R.A. and Kamoun, S. (2008). From Guard to Decoy: a new model for perception of plant pathogen effectors. *Plant Cell* **20**, 2009-2017.

van der Vossen, E.A.G., van der Voort, J., Kanyuka, K., Bendahmane, A., Sandbrink, H., Baulcombe, D.C., Bakker, J., Stiekema, W.J. and Klein-Lankhorst, R.M. (2000). Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: a virus and a nematode. *Plant J.* **23**, 567-576.

Verchot-Lubicz, J., Ye, C.-M. and Bamunusinghe, D. (2007). Molecular biology of potexviruses: recent advances. *J Gen Virol* **88**, 1643-1655.

Voinnet, O., Lederer, C. and Baulcombe, D.C. (2000). A viral movement protein prevents spread of the gene silencing signal in *Nicotiana benthamiana*. *Cell* **103**, 157-167.

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

Zayakina, O., Arkhipenko, M., Kozlovsky, S., Nikitin, N., Smirnov, A., Susi, P., Rodionova, N., Karpova, O. and Atabekov, J. (2008). Mutagenic analysis of Potato Virus X movement protein (TGBp1) and the coat protein (CP): in vitro TGBp1-CP binding and viral RNA translation activation. *Mol Plant Pathol* **9**, 37-44.

Chapter 2

Fluorescent T7 display phages obtained by translational frameshift

Erik J. Slootweg¹, Hans J.H.G. Keller¹, Mark A. Hink³, Jan Willem Borst³, Jaap Bakker², Arjen Schots¹

¹*Laboratory of Molecular Recognition and Antibody Technology, Wageningen University, Wageningen, The Netherlands*

²*Department of Nematology, Wageningen University, Wageningen, The Netherlands.*

³*Microspectroscopy Centre, Wageningen University, Wageningen, The Netherlands*

Published in **Nucleic Acids Res.** 2006;**34(20):e137.**

Abstract

Lytic phages form a powerful platform for the display of large cDNA libraries and offer the possibility to screen for interactions with almost any substrate. To visualise these interactions directly by fluorescence microscopy, we constructed fluorescent T7 phages by exploiting the flexibility of phages to incorporate modified versions of its capsid protein. By applying translational frameshift sequences, helper plasmids were constructed that expressed a fixed ratio of both wild-type capsid protein (gp10) and capsid protein fused to enhanced yellow fluorescent protein (EYFP). The frameshift sequences were inserted between the 3'-end of the capsid gene and the sequence encoding EYFP. Fluorescent fusion proteins are only formed when the ribosome makes a -1 shift in reading frame during translation. Using standard fluorescence microscopy, we could sensitively monitor the enrichment of specific binders in a cDNA library displayed on fluorescent T7 phages. The perspectives of fluorescent display phages in the fast emerging field of single molecule detection and sorting technologies are discussed.

Introduction

A major challenge in the post-genome era is to unravel protein-protein interactions which are involved in transmission of information within cells, the so-called interactome. Identification of these interacting molecules would highly support the understanding of how cells function and also provide leads to the development of new drugs. To map the interactome a range of high-throughput screening technologies is employed, including yeast two-hybrid, mass spectroscopy of co-immunoprecipitated protein complexes, protein arrays, and phage display, all with their particular strengths and weaknesses (Ito et al., 2000; Li, 2000; Gavin et al., 2002; Tong et al., 2002). For phage display these strengths are the large libraries that can be created, the level of control over the binding conditions, the ease of identifying interactors by PCR (the phenotype and genotype are coupled) and the possibility to screen for interactions with almost any kind of substrate, from small chemical compounds to post-translationally modified proteins or complete cells (Sche et al., 1999; Zozulya et al., 1999; Geuijen et al., 2005). Since its introduction 20 years ago, phage display has played a crucial role in selecting interacting molecules from large libraries. The filamentous phage M13 predominated for a long time in the phage display field and is still a key player for the selection of recombinant antibody fragments (Smith, 1985; Devlin et al., 1990; Scott and Smith, 1990; Griffiths et al., 1994). However, for the assembly of the M13 phage particles all viral proteins need to be transported through the bacterial inner membrane. Sequence and folding characteristics of cytosolic proteins are often incompatible with this translocation process, which imposes serious constraints on the display of cDNA libraries on M13. To circumvent this drawback of M13, lytic phages like T4, T7 and lambda have been adapted for display applications (Sternberg and Hoess, 1995; Ren et al., 1996; Yamamoto et al., 1999). Lytic phage display does not rely on the *E. coli* secretory mechanism, because the phages are assembled in the cytoplasm and released by lysis of the bacterium.

The commercially available T7 phage display system (T7Select; Novagen) is at the moment the most widely used lytic phage display system and has successfully been employed to reveal interactions between proteins and between proteins and chemical compounds (Sche et al., 1999; Danner and Belasco, 2001; Sheu et al., 2003; Horibe et al., 2004; Krajcikova and

Hartley, 2004). The T7 phage allows the fusion of large protein fragments, up to 1000 amino acids in a low copy number, to its capsid protein. Further advantages are the availability of an efficient packaging system and the fact that due to its fast amplification, multiple selection rounds can be performed per day.

Here we report a method to construct fluorescent T7 display phages. Fluorescent phage particles enable a direct visualisation of the interaction between displayed proteins and their binding partners, which has several advantages. The enrichment of specific binders in a display library can be followed directly by standard fluorescence microscopy during the affinity selection procedure, obviating the need for more laborious and time consuming procedures like ELISA and plaque lift assays. Directly monitoring the enrichment allows making a more considered decision on when to stop the selection procedure, thereby avoiding unnecessary amplification steps. Furthermore, future developments in advanced microscopic techniques may enable the detection of single fluorescent phages in extremely small volumes, which opens the door to sensitive biolibrary sorting platforms (Bohmer and Enderlein, 2003; Visser et al., 2004).

To obtain fluorescent phage particles we incorporated the enhanced yellow fluorescent protein (EYFP) in the phage capsid by fusing it to the T7 capsid protein (gp10). The T7 phage will allow only a limited proportion of its capsid proteins to be fused to other proteins while maintaining its infectivity. For this reason, we constructed helper plasmids that express both the wild-type capsid protein and the fluorescent protein from the same gene fusion by introducing a regulated translational frameshift site between the two fusion partners. The frameshift sequence was placed in such a way that when the bacterial ribosome follows the normal reading frame it will encounter a stop codon after the capsid protein gene, but once it shifts to the -1 frame it will read through into the EYFP gene. Regulated translational frameshifts are a relatively common phenomenon in nature and have been found in viruses, prokaryotes and eukaryotes (Mardon and Varmus, 1983; Brault and Miller, 1992; Levin et al., 1993; Matsufuji et al., 1995; Zheng and McIntosh, 1995; Ivanov et al., 2000; Licznar et al., 2003). They are based on secondary structures in mRNA that combined with specific RNA sequences will cause the ribosomes to leave the original reading frame so that codons are redefined or translational bypasses occur. As far as we know this is the first report of using a translational frameshift for a biotechnological purpose.

Material and methods

Helper plasmid construction

The frameshift helper plasmids were based on the helper plasmid present in the BLT5403 *E. coli* strain (Novagen), which carries the T7 gene 10 under the T7 $\phi 10$ promoter. A fragment of gene 10 was amplified using the primers 10A-F (ACT ATA GGG AGA CCA CAA CCG) and 10A-EcoRI-mut-R (CTT CAA GAA TTC TTA CTC CAC TTT GAA AAC CAC TGC ACC AGC AGC) to restore the natural frameshift sequence by the introduction of three mutations (underlined). The PCR fragment was introduced in p5403 by exchanging the sequence between the KpnI and EcoRI restriction sites. The EYFP gene (Clontech) was amplified to introduce an EcoRI site on both ends, using the primers YFP-EcoRI-F (GTA AGA ATT CTT ATG GTG AGC AAG GGC GAG) and YFP-EcoRI-R (GT GAA TTC TTA GCT CAT GAC TGA CTT GTA GAG). The ligation of the resulting fragment into the modified p5403 resulted in p10-EYFP, wherein the *EYFP* gene is present in the -1 reading frame relative to gene 10.

To create pX1-EYFP and pX2-EYFP, first an EcoRI and a HindIII site were introduced in p5403 upstream of the stop codon of gene 10 by introducing in between the BsgI and EcoRI site a DNA fragment formed by annealing the oligonucleotides EH-F (G AAT TCC GGT TCT TAA AAG CTT TAA C) and EH-R (AA TTG TTA AAG CTT TTA AGA ACC GGA ATT CTC). The new EcoRI and HindIII sites were used to introduce the *EYFP* gene amplified with YFP-EcoRI-F (GT AAG AAT TCC ATG GTG AGC AAG GGC GAG) and YFP-HindIII-R (GC AAG CTT TTA GCT CAT GAC TGA CTT GTA GAG). This resulted in p5403-EH-EYFP. The *DnaX* frameshift cassettes were constructed from complementary oligonucleotides. The *DnaX* cassette with stem-loop (X1) with the following oligo pairs: XU1 (GCA GGG AGC AAC CAA AGC AAA AAA G)/ XL1 (ACT CTT TTT TGC TTT GGT TGC TCC CTG CTC), XU2 (AGT CAA CCG GCA GCC GCT ACC CGC)/ XL2 (CGC GCG GGT AGC GGC TGC CGG TTG) and XU3 (GCG CGG CCG GTG TGA ATT CCA TGC CCA)/ XL3 (AGC TTG GGC ATG GAA TTC ACA CCG GCC G). The underlined nucleotides represent the introduced stop codon responsible for terminating the translation when no frameshift occurs. The ligation fragment was purified from gel and ligated into the BsgI and HindIII digested p5403EH-EYFP, creating pX1-EYFP. The pX2-EYFP helper plasmid was constructed the same way, but with a cassette based on the following complementary primers: XSU1 (GCA GGG AGC AAC CAA AGC AAA AAA GAG GTG AAT TCC ATG CCC A) / XSL1 (AGC TTG GGC ATG GAA TTC ACC TCT TTT TTG CTT TGG TTG CTC CCT GCT C). All constructs were sequenced to check if sequences were correct.

Protein production

BL21(DE3) cells transformed with either pX1-EYFP, pX2-EYFP or p10-EYFP were grown at 37° C in 8 ml liquid LB with 100 µg/ml ampicillin (LB-amp). The protein production was induced with 1 mM IPTG when the cultures reached an optical density (OD₆₀₀) of 0.5. After induction the temperature was lowered to 28° C. After three hours the cells were pelleted and the produced protein was analysed on 10% SDS-PAGE. For Western blot biotinylated anti-T7 (T7-Tag antibody, Novagen) was used diluted 1:10.000 in PBS containing 0.1% BSA and 0.1% Tween-20 (PBSBT), and the antibody was detected with 1:2000 streptavidin-alkalic

phosphatase (AP) conjugate (Sigma). EYFP was detected with 1:200 anti-GFP (Clontech) and 1:5000 anti-Rabbit-AP (Jackson), both diluted in PBSBT.

T7 Select cDNA library construction and bio-panning procedures

For the construction of the cDNA display library, mRNA was extracted from total RNA of PVY^c infected *Nicotiana benthamiana* leaf material with poly-dT beads (Genoprep). cDNA was synthesised with the OrientExpress cDNA synthesis kit (Novagen) with random hexanucleotide primers. The cDNA fragments between 100 and 1500 bp length were ligated in the T7Select10-3 vector arms (Novagen) via a directional cloning strategy. After packaging the library in T7 phages the number of individual clones (5×10^5 pfu) was determined by titring and the average insert size (~500 bp) was determined by PCR amplification of 30 individual phage plaques. Three rounds of biopanning against the 5H6 monoclonal antibody were performed. The selection procedure consisted of incubating 2×10^8 phages with 3 μ g 5H6 antibody coated in microtiter plate wells for 3 hours at room temperature. Non-binding phages were removed by washing 5x with PBST. Eluted phages were amplified after each round in 2 ml BLT5403 culture. Enrichment in phages displaying binding protein fragments was monitored by plaque lifts on nitrocellulose filter (500-1500 plaques per selection round); the plaques of binding phage were detected with alkaline phosphatase-conjugated bait (5H6). The inserts from 25 binding phage clones from the third selection round were sequenced.

T7 amplification and purification

The effect of the frameshift helper plasmids on the efficiency of the phage amplification was tested by growing *E. coli* BL21(DE3) containing one of the helper plasmid constructs, in 8 ml of LB-Amp at 37° C. At an OD₆₀₀ of 0.5, the temperature was lowered to 28° C, and the cells were induced with 1 mM IPTG and simultaneously infected with 2×10^8 plaque forming units (pfu). Phage titers were determined by mixing serial dilutions of the produced phages to an overnight culture of cells containing the original helper plasmid (BLT5403). The bacteria and phages were mixed with warm top agar (48° C) and plated on LB agar plates and incubated at 37° C.

Phages needed for purification were amplified in a 300 ml culture of *E. coli* containing the required helper plasmid by infecting it with 4×10^{10} pfu of T7. After lysis of the bacteria, 50 μ l Dnase I (2 mg/ml) was added and the phages were incubated for a further 20 minutes. NaCl was added to a final concentration of 500 mM and the phages were separated from the cell debris by centrifugation for 10 min at 8000 rpm in a Sorval SLA-3000 rotor at 4° C. 10% (w/v) PEG 6000 was added to the supernatant. The lysate-PEG mixture was stored overnight at 4° C. The next day the phages were pelleted (8,000 rpm at 4° C in a Sorval SLA-3000 rotor). The pellet was dissolved in 5 ml 1M NaCl, 10 mM Tris-HCl pH 8.0, 1 mM EDTA. The resuspended phages were loaded on a CsCl gradient (20-62% w/v) and purified in a Beckman SW41 rotor centrifuged at 35,000 rpm for one hour at 20° C. The phages were collected and dialysed against PBS and stored at 4° C until further use.

Phage labelling with Alexa Fluor 488

10 μ l of 1M sodium carbonate was added to 100 μ l of the purified phages (3×10^{12} particles/ml), mixed with one vial Alexa Fluor 488 tetrafluorophenyl (TFP) ester from the

Alexa Fluor 488 protein labelling kit (Molecular Probes, Product #A10235) and incubated for one hour at room temperature. The phages were separated from the unincorporated Alexa dye by size exclusion spin column chromatography. The labelled phages were stored at 4° C in PBS buffer.

Fluorescent Correlation Spectroscopy

For fluorescence correlation spectroscopy (FCS) a Zeiss-EVOTEC ConfoCor® system was used in combination with an argon ion laser supplying a 488 nm wavelength. The laser power was reduced with a 2.0 density filter and focused in the sample by a water immersion objective (Zeiss C-Apochromat 40x, N.A 1.2). The phage samples were measured in a glass bottom 96 well plate. Data traces were analysed in FCS Data processor 1.4.1 (SSTC) and analysed using a triplet-state model (Skakun et al., 2005).

Antibodies biotinylation and coating of microspheres

Prior to the biotinylation reaction 40 µl of 0.5 M sodium carbonate of pH 9.5 was added to 400 µl of anti-T7, 5H6 (Y-5) or anti-myc (9E10) (2 mg/ml in PBS). To this mixture 4 µl N-hydroxysuccinimido-D- biotin (3 mg/ml) (Sigma) dissolved in DMSO was added and then incubated for four hours at room temperature, while gently mixing. The free biotin was removed by dialysis against PBS at room temperature. Sodium azide was added and the biotinylated proteins were stored at 4° C till further use. The streptavidin-coated microspheres (1% solids) from Bangs Laboratories with a diameter of 0.95 µm had a capacity to bind 0.38 µg biotinylated FITC per ml beads. Microspheres were vortexed and 10 µl of the suspension was mixed with 180 µl PBS containing 0.1% BSA and 0.1% Tween-20 (PBSBT). To the diluted microspheres 10 µl of the biotinylated antibody were added, and incubated for 45 minutes at room temperature. During incubation the mixture was briefly vortexed every ten minutes. The microspheres were pelleted by centrifugation (15 min, 1500 rcf). The supernatant was removed and the pellet was resuspended in 200 µl PBSBT. To remove most unbound antibody the microspheres were pelleted and resuspended in fresh PBSBT three times in total.

Fluorescent imaging of the phage interacting with antibody coated microspheres

The microspheres coated with antibodies (25 µl) were incubated with 100 µl fluorescent phages (5×10^{12} phages/ml) for three hours at room temperature prior to observing the interaction in glass-bottomed 96 well plates (Whatman). Images were made on a Biorad Radiance 2100 MP-VIS system coupled to a Nikon TE300 inverted microscope. The sample was excited by an Argon laser at 488 nm. Fluorescence was collected using a Nikon Plan Apochromat 60x water immersed objective (N.A. 1.2) and filtered by a BP505-560 emission filter (Omega).

Results

Vector construction and frameshift induction in E. coli cells

Three helper plasmids were constructed to obtain genetically fluorescent T7 phage. The function of helper plasmids is to provide sufficient wild-type capsid protein in addition to the library-displaying capsid protein that originates from the T7 genome. Different translational frameshift cassettes were placed at the 3'-end of the sequence encoding the wild type capsid protein (gp10) in the helper plasmid p5403. The expression of gp10 in these constructs is regulated by the T7 ϕ 10 promoter. The first constructed vector, named pX1, contained three recoding elements from the *E. coli dnaX* gene; the frameshift inducing sequence enclosed by a Shine Dalgarno motif at one side and a stem-loop sequence at the other. The latter two elements influence the frequency of the frameshift event, which in wild-type *dnaX* occurs in 50% of the translations. A second helper plasmid, named pX2, was constructed by including the *dnaX* frameshift sequence and the Shine Dalgarno sequence only, which was predicted to have a lower frameshift frequency (Larsen et al., 1997). A third helper plasmid, named p10, was constructed by re-introducing a short frameshift sequence found in the wild-type gene 10 of T7 at the end of the capsid protein gene in p5403. In the wild-type T7 this sequence is responsible for a -1 frameshift that causes the formation of the minor capsid protein 10B instead of the major capsid protein 10A in about 10% of the translations (Condrón et al., 1991a; Condrón et al., 1991b; Siple et al., 1991). To complete the construction of the plasmids the gene coding for the enhanced yellow fluorescent protein (EYFP, Clontech) was placed in the -1 reading frame directly downstream of the frameshift cassettes, thereby creating the vectors pX1-EYFP, pX2-EYFP, and p10-EYFP (Figure 1). In all three vectors stop codons were introduced right after the frameshift cassettes in the 0 reading frame, so that if no frameshift takes place no extensive amino acid sequence will be added to the capsid protein. However, when the ribosome shifts from the initial reading frame into the -1 frame, the translation continues and a capsid protein-EYFP fusion (gp10-EYFP) is formed.

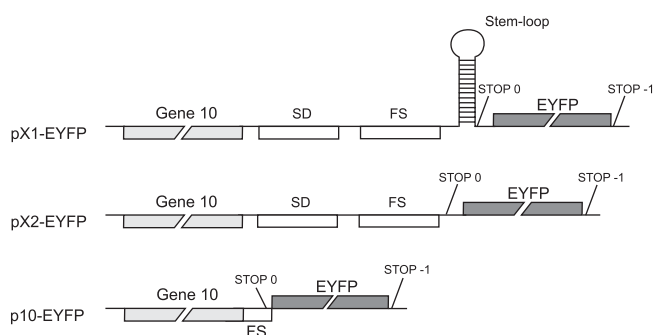


Figure 1. Schematic representation of the three translational frameshift constructs. The frameshift cassettes placed in the helper vector p5403 are located between gene 10 coding for the capsid protein (gp10) and the EYFP gene. The latter is placed in the -1 coding frame in relation to gene 10. Without recoding, the protein translation is terminated after the frameshift site (FS) indicated by STOP 0. In the -1 frame the stop codon is located beyond the EYFP gene; STOP -1. When the ribosome shifts to the -1 reading frame it produces the gp10-EYFP fusion. The constructs pX1-EYFP and pX2-EYFP contain frameshift elements originating from the *E. coli dnaX* gene; a Shine-Dalgarno (SD) sequence, the frameshift site (FS) and a stem-loop sequence. The pX2-EYFP vector is similar to pX1-EYFP, but lacks the stem-loop sequence. The frameshift site in p10-EYFP contains the short GGTTTTC frameshift inducing sequence present in T7 gene 10.

The expression of these constructs was tested in isopropyl- β -D-thiogalactopyranoside (IPTG) induced *E. coli* BL21(DE3) cells and the bacterial samples were subjected to SDS-PAGE. Coomassie Brilliant Blue staining of the proteins showed in all cases that after induction there was a high expression of wild-type coat protein without the EYFP fusion. Figure 2A shows that there are slight differences in molecular weight between the free coat proteins as they are expressed from the three vectors. These differences are caused by additional amino acids encoded in the frameshift cassettes of pX1-EYFP and pX2-EYFP (20 and 9 amino acids, respectively). The high concentrations of gp10-EYFP expressed from the vectors pX1-EYFP and pX2-EYFP are clearly visible in the Coomassie Brilliant Blue stained SDS-PAGE gel as bands with an approximate molecular weight of 70 kDa. The band intensity was used to estimate the ratio between gp10-EYFP and free capsid protein. These were approximately 1:1 for pX1-EYFP and 1:4 for pX2-EYFP. The Western blot using antibodies that either recognise T7 gp10 or EYFP showed in both cases strong signals, confirming that the bands at 70 kDa are indeed gp10-EYFP fusions, indicating that the fusion protein was abundantly expressed from pX1-EYFP and pX2-EYFP. The level of expression of gp10-EYFP produced from p10-EYFP was so low that the fusion protein could not be detected on the Coomassie Brilliant Blue stained gel. A Western blot of this construct specifically showed a thin band at 70 kDa indicating the expression of the gp10-EYFP fusion protein (Figures 2B and 2C). It was estimated that a frameshift had taken place in less than 1% of the translations from the p10-EYFP vector.

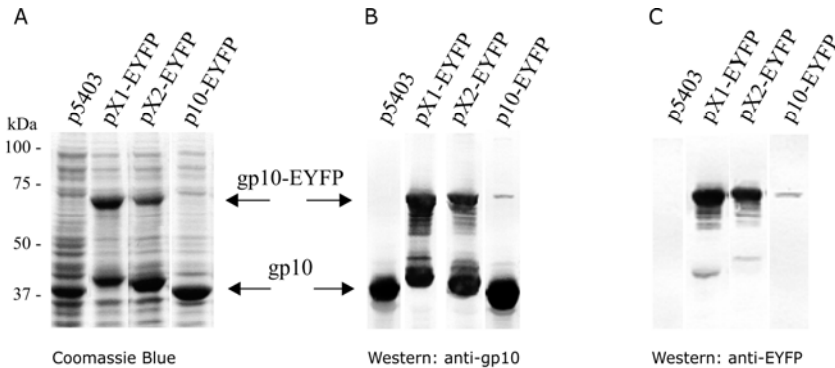


Figure 2. Different ratios of free capsid protein (gp10) and capsid protein fused to EYFP (gp10-EYFP) are produced from the three frameshift helper plasmids. Total cell samples of *E. coli* expressing the original helper plasmid (p5403), pX1-EYFP, pX2-EYFP and p10-EYFP were subjected to SDS-PAGE and analysed by Coomassie Blue staining (A) or Western blotting with anti-gp10 antibodies (B) and anti-EYFP antibodies (C). The positions on the gel of free gp10 and gp10-EYFP are indicated.

The gp10-EYFP: gp10 ratio affects the formation of infectious phage particles

Expression from each of the three frameshift constructs, pX1-EYFP, pX2-EYFP and p10-EYFP, results in the translation of both a wild-type capsid protein and a capsid protein-EYFP fusion protein in fixed ratios at three different levels. To test the functionality of the constructed helper plasmids, recombinant phages displaying a 95 amino acid fragment of the Potato Virus Y coat protein (Pcp) were amplified in bacteria containing either the helper plasmids pX1-EYFP, pX2-EYFP, or p10-EYFP. It was expected that functional helper plasmids would lead to the formation of infectious phage particles with wild-type capsid protein, capsid protein fused to EYFP, and capsid protein displaying Pcp. Three hours after infection of the cultures, the bacteria containing the vector p10-EYFP or the original helper plasmid p5403 were fully lysed, whereas the bacterial cultures containing pX1-EYFP or pX2-EYFP were still turbid. Absence of cell lysis implies the delay or inhibition of the phage amplification. To assess the efficiency of phage production for each of the helper plasmids the phage titers were determined. In the bacterial cultures containing pX1-EYFP and pX2-EYFP no infectious phage particles were detected. The titer of the phages produced from p10-EYFP was of the same order as the titer produced using the original p5403 helper plasmid (i.e. 2×10^{10} plaque forming units per ml). As from the three helper plasmids only p10-EYFP resulted in infectious phages, all further experiments were performed with this construct.

Fluorescence correlation spectroscopy measurements demonstrate the incorporation of gp10-EYFP in T7 capsids

Fluorescence correlation spectroscopy (FCS) was applied to determine if the gp10-EYFP fusion protein was indeed incorporated in the phage particles produced from the p10-EYFP helper plasmid. FCS is a spectroscopic technique that provides information about the diffusion rate and the average number of fluorescently labelled molecules in a solution. Fluorescent molecules diffusing through a confocal detection volume are detected as intensity fluctuations. By correlating these fluctuations in time the behaviour of multiple particles with different diffusion times can be resolved in a single measurement (Maiti et al., 1997; Visser and Hink, 1999; Dittrich et al., 2001). Assembled phage particles have a diameter of about 60 nm, which is more than ten times larger than free gp10-EYFP molecules, and are thus expected to have significantly longer diffusion times. Phages produced using p10-EYFP (T7-Pcp-EYFP) and p5403 (T7-Pcp) were purified and concentrated using ultracentrifugation on a CsCl gradient to eliminate most of the non-assembled gp10 molecules and bacterial debris. Similarly purified and Alexa488-labeled T7 phages served as comparison. FCS analysis of the Alexa488 labelled phage led to a diffusion time (τ) of 1.33 (1.27-1.35) milliseconds (Table 1). The autocorrelation data of free EYFP could be fitted to a diffusion time of 34 (32-35) microseconds. T7-Pcp phages do not contain any EYFP and excitation at a 514 nm wavelength resulted only in a low background fluorescence that could not be autocorrelated. Under the same conditions the T7-Pcp-EYFP phages, produced with the p10-EYFP helper plasmid, fluoresced with a maximal emission at 527 nm, typical for EYFP. Surprisingly, it was not possible to fit the autocorrelation data for T7-Pcp-EYFP for a single diffusion time. The autocorrelation curve showed that the sample contained two components. A two-component fit resulted in a relatively long diffusion time of 1.7 ms, comparable with the alexa488-labeled T7, and a short diffusion time of approximately 25 μ s that was close to the value found for free EYFP and Rhodamine Green

(Table 1). The exact diffusion time of the fast component could not be determined from these data. To exclude the possibility that this fast component was due to a photophysical phenomenon the measurements were repeated in a buffer with a higher viscosity. In the more viscous solution the diffusion times for both the fast and slow component were extended (data not shown), meaning that it must be a fluorescent particle, most likely non-incorporated gp10-EYFP. A photophysical artefact would have been independent of the viscosity (Enderlein et al., 2004).

Table 1. Results of the FCS analysis of the phage variants and fluorescent molecules.

Sample	τ_1 (μs)	D_1 (m^2s^{-1})	%F1	τ_2 (ms)	D_2 (m^2s^{-1})	%F2
Rhodamine Green	23 (21-25)	2.8×10^{-10}	100	-	-	-
YFP (Citrine)	34 (32-35)	1.9×10^{-10}	100	-	-	-
T7-Alexa488	-	-	-	1.33 (1.27-1.35)	4.8×10^{-12}	100
T7-Pcp-EYFP	25 (23-26)	2.5×10^{-10}	77	1.7 (1.7-1.8)	3.8×10^{-12}	23
T7-Pcp	-	-	-	-	-	-

Diffusion time (τ) with standard deviation derived from 10 measurements. Translational diffusion coefficient calculated as in Hink et al., Eq. 2. F1 and F2 denote the relative fractions of each component (Hink et al., 2000).

The amplification temperature influences the formation of EYFP-labelled phages

T7 phage cDNA libraries are normally amplified in bacterial cultures incubated at 37 °C. However, this is not the optimal temperature for the formation of the fluorophore in EYFP (Kimata et al., 1997; Tsien, 1998). A subset of gp10-EYFP formed at 37 °C will not become fluorescent. To investigate the influence of the incubation temperature during amplification of the library on the phage production and on the proportion of phage particles that show fluorescence, T7-Pcp-EYFP phages were amplified at 20 °C, 28 °C and 37 °C. Bacteria containing the p10-EYFP helper plasmid were first grown at 37 °C till the OD₆₀₀ reached 0.5 before the temperature was changed to the required amplification temperature, the protein production was induced, and the cells were infected with the T7-Pcp phages. After complete lysis of the cells, the infectious phage concentrations were determined (Figure 3A). At 37 °C approximately ten times more infectious phages are formed than at 20 °C. Amplification at 28 °C showed an intermediate efficiency. FCS measurements could be used to determine the amount of fluorescent phages. Thereto the phages were first precipitated, purified on a CsCl gradient by ultracentrifugation, and resuspended in 1.5 ml of PBS. The total phage concentration after purification was determined by measuring the optical density of the phage solution at 260 nm (Figure 3B). Although the amplification temperature clearly has an influence on the infectivity of the phages (Figure 3A), the total amount of phages, as determined by the optical density, did not differ significantly between the three growth temperatures (Figure 3B). Analysis of FCS measurements on the purified phages gave the average number of fluorescent phage particles present in the confocal detection volume (Figure 3C). The size of the detection volume was determined by analysis of compound with known diffusion coefficients, Rhodamine Green and EYFP. Knowing the size of the detection volume enabled us to calculate the concentration of fluorescent phages in the samples. The average concentration of fluorescent phages appeared to be the lowest at 37 °C (Figure 3C). Comparison of the concentration of fluorescent phages (Figure 3C) with the total phage concentration (Figure 3B) indicates that the relative amounts of fluorescent

phages are approximately 16% when amplified at 20 °C, 14% at 28 °C and 6% at 37 °C. Based on these differences, although not statistically significant, and the infectivity of the phages (Figure 3A) we considered 28 °C the optimal amplification temperature for fluorescent phage production and consequently used this temperature in further experiments.

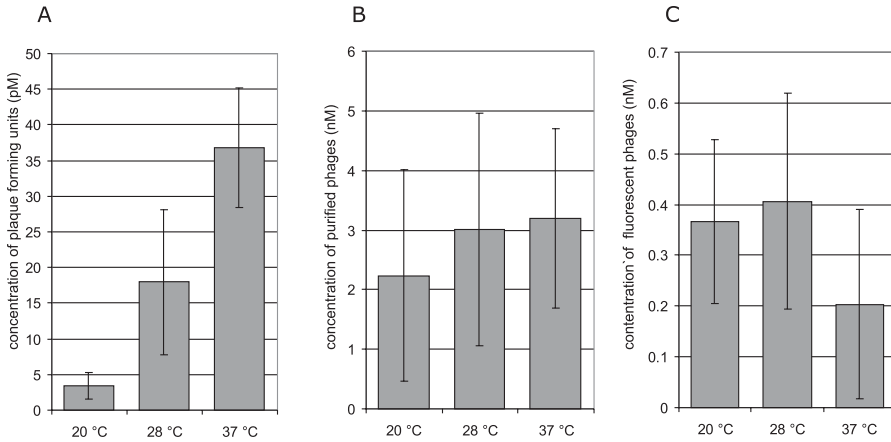


Figure 3. Influence of the temperature on the amplification efficiency and the formation of fluorescently labelled T7-Pcp-EYFP phages. Phages were amplified at 20, 28 and 37 °C in 300 ml bacteria containing the p10-EYFP vector. The concentration of infectious phages was determined by titring the phages (A, concentration in pM, based on 6 samples). After CsCl purification the total phage concentration was determined by measuring the absorbance at 260 nm (B, based on 4 samples). The concentration of fluorescent phage particles was determined by FCS analysis in four (20 °C, 37 °C) or five (28 °C) independent samples (C).

EYFP-labelled phages enable the visualisation of protein-protein interactions

To visualise protein-protein interactions using fluorescent display phages we tested the physical interaction between the PVY coat protein fragment (Pcp) and a monoclonal antibody recognising this fragment (5H6; Y-5 in ref. (Boonekamp et al., 1991)). To this purpose Pcp was displayed on fluorescent phages including gp10-EYFP (hereafter referred to as T7-Pcp-EYFP) and microspheres were coated 5H6. The 5H6-coated microspheres were expected to capture the T7-Pcp-EYFP phages, which could only be detected by fluorescence microscopy when the phage displays both the PVY coat protein fragment and EYFP molecules. Microspheres coated with an antibody to a *Myc*-tag (9E10) were used as a negative control; the phages do not contain the *myc*-tag and should not be recognised. Microspheres coated with an anti-T7 antibody that should bind all T7 phage particles were used as positive binding control. After incubation, the spheres were washed several times to remove unbound antibody. Antibody-coated microspheres that were not incubated with phages did not show a fluorescent signal and could not be detected by fluorescence microscopy under 514 nm excitation. The anti-T7 antibody-coated microspheres incubated with the T7-Pcp-EYFP phages were detected as spots on a dark background, indicating an interaction between the antibody and the fluorescent phage (Figure 4A). EYFP fluorescence could also be detected on the 5H6 antibody-coated spheres after incubation with T7-Pcp-EYFP phages in a pattern similar to what was found with the anti-T7 antibody-coated spheres (Figure 4B). No fluorescence could be detected on the anti-*myc* antibody-

coated spheres incubated with the phages, demonstrating that the fluorescence detected on the anti-T7 and 5H6 antibody-coated spheres was due to specific binding of the phages (Figure 4C). From these results we conclude that the T7-Pcp-EYFP phage particles are labelled with EYFP and display the PVY coat protein fragment. Even though only a 7% of the phages contains a fluorescent EYFP molecule, as shown by FCS analysis, this is sufficient to detect the interaction between the phage and the bait-coated microspheres.

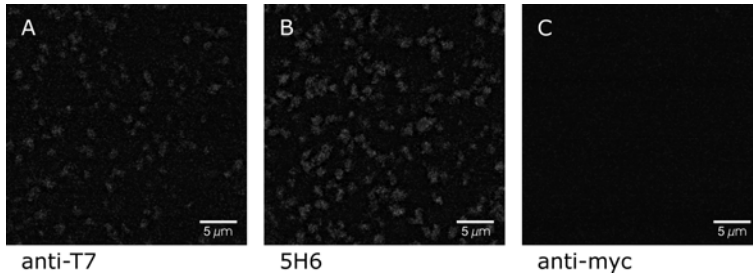


Figure 4. Visualisation of the interaction between T7-Pcp-EYFP phages and antibody-coated microspheres (0.95 µm). The microspheres coated with the anti-T7 antibodies (A) or 5H6 antibodies against the PVY coat protein (B) become fluorescent after incubation with the T7-Pcp-EYFP phage. Incubation of T7-Pcp-EYFP phages with anti-myc antibody-coated microspheres did not result in detectable EYFP fluorescence on the beads (C).

Monitoring the enrichment of T7 phage display libraries by fluorescence microscopy

The next step was to examine the display of a diverse cDNA library on fluorescent phages. As a model system it was decided to use different well characterised selection stages from a cDNA display library biopanning. The cDNA display library was constructed from PVY-infected *Nicotiana benthamiana* plants and inserted in the T7Select10-3 genome vector (Novagen). The library was selected against the 5H6 monoclonal antibody. Three selection rounds had resulted in subpopulations of this cDNA display library at different degrees of enrichment for phages displaying 5H6-binding protein fragments. The extent of enrichment was tested via plaque lift assays. Less than 0.5% of the phages in the original library displayed a 5H6-binding fragment, whereas the proportion increased to 5% after the first selection round, 31% after the second, and 59% after the third selection round. PCR Amplification and sequencing of the cDNA inserts from interacting phages from the third selection round showed that the enrichment of the library was specific. In total 10 cDNA fragments of different lengths were found, all encoding fragments of the PVY coat protein.

The original unselected library and the results from the first and third selection round were amplified in bacteria containing the p10-EYFP helper plasmid. The resulting phages were purified by ultracentrifugation on a CsCl gradient and incubated with a mixture of 5H6 and anti-myc coated microspheres (at a ratio of 1:9). The microspheres coated with 5H6 were labelled with a red fluorescent dye to distinguish them from the anti-myc coated microspheres. By fluorescence microscopy using two excitation wavelengths (514 and 633 nm) the specific interaction of the fluorescent T7 phages with the 5H6-coated microspheres was monitored. Excitation at 514 nm of the beads that were incubated with the phages displaying the original library revealed faint dots that corresponded with the location of

some of the 5H6-coated spheres as seen at 633 nm excitation (figure 5A-D). However, there was also some EYFP-fluorescence visible on the anti-myc coated spheres. This indicates that in the unselected library the proportion of phages displaying a 5H6-binding peptide (~0.5%) is still too low to elevate the signal above background level in this assay. The library that had been subjected to one round of selection shows a brighter EYFP signal on the 5H6-coated beads; the spheres that exhibit EYFP fluorescence correspond here to the ones with the red fluorescence (figure 5E-H). Incubation of phages from the third selection round, with 59% of the phages displaying a 5H6-binding peptide, resulted in the accumulation of a high amount of fluorescent phages specifically on the 5H6-coated microspheres (Figure 5I-L). The relative brightness of the beads was determined with Kodak 1D image analysis software. After one round of selection the beads were 2.5 times brighter than before selection. After three rounds the fluorescence intensity of the beads was 6.4 times higher than after incubation with the unselected library. The fluorescence screening was repeated with other ratios of 5H6- and anti-myc coated microspheres (9:1, 1:1, 1:50). A similar increase in fluorescence on the beads, even after one round of selection, could be seen (data not shown). This demonstrates that it is possible to apply the frameshift helper plasmid for the amplification and fluorescent labelling of a cDNA display library. Such a fluorescently labelled library can be used to directly monitor the enrichment in specific binders during an affinity selection procedure.

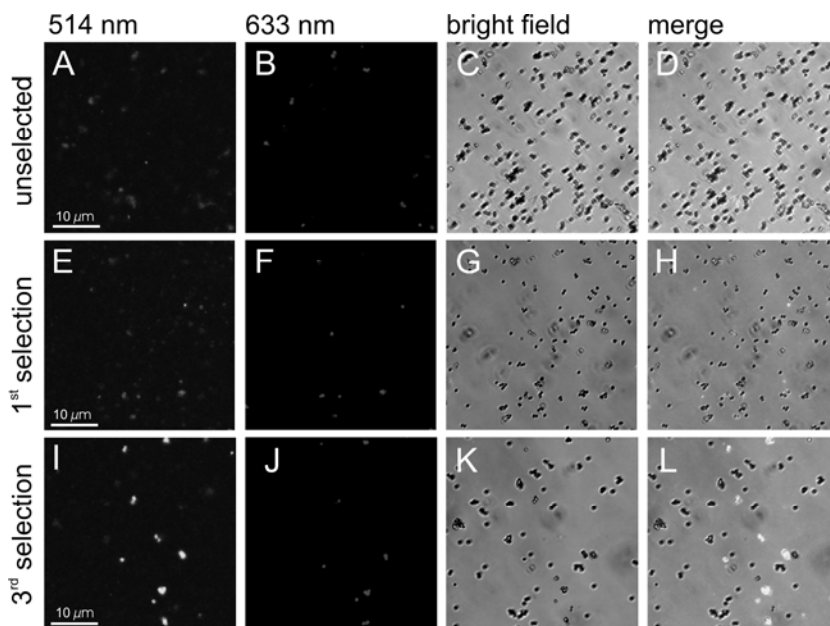


Figure 5. Microscopic images of antibody-coated microspheres incubated with EYFP-labelled T7 phages from different affinity selection rounds. Red fluorescent microspheres (0.95 μm) were coated with the 5H6 antibody against the PVY coat protein and non-fluorescent microspheres with an anti-myc antibody. The microspheres were mixed in a 1 to 9 ratio. EYFP-labelled T7 phages displaying the unselected library (A-D), the library after one round of affinity selection against 5H6 (E-H) or the library after three rounds of selection (I-L) were incubated with the microspheres. EYFP fluorescence was detected at 514 nm excitation (A, E, I), the red fluorescent beads at 633 nm (B, F, J). The bright field image (C, G, K) shows all beads. In D, H, and L the superposition of the three channels is shown.

Discussion

Translational frameshifts are a common phenomenon in nature. It is one of the mechanisms that enable cells to encode different variants of a protein in a single gene. Often these variants have specific functions in the cell. The frameshift sequence from the *E. coli DnaX* gene that is used in this study illustrates this quite well. Both the non-frameshifted product (τ) and the shorter, frameshifted variant (γ) are formed in a 1:1 ratio and serve as subunits of DNA polymerase III where they have distinct functions (Larsen et al., 1997; McHenry, 2003). The result of a translational frameshift, or ribosome recoding, is the production of two proteins in a fixed ratio and with a particular nested sequence relationship. A feature that has interesting biotechnological implications as we show in this study. We employed ribosome recoding to produce fluorescence labelled T7 phages that are fully compatible with the commercially available T7 display system (T7Select, Novagen).

Using a recoding sequence to express two gene products has a number of advantages. In this case the expression of both the wild-type capsid protein and the capsid protein fused to EYFP were needed. If no translational recoding sequence is used, for each product a complete open reading frame should be present, either on one plasmid or on two different plasmids. In both options the presence of two homologous DNA sequences in the cell will adversely influence the plasmids stability (Baneyx, 1999). Furthermore, to get the desired ratio of the two products, the promoters, regulatory elements and copy numbers of the plasmids will have to be adjusted. With our system a certain ratio can be achieved by employing the right frameshift sequence, and many have been described already, with a wide range of frameshifting frequencies (Larsen et al., 1997; Baranov et al., 2002).

The fact that T7 phages can be used as a display system is based on their ability to incorporate capsid proteins with fusions into their capsids. However, fusions larger than about 50 amino acids cannot be displayed on every copy of the capsid protein. The T7 system used in this study (T7Select10-3) displays peptides from the cDNA library on 10 to 15 out of the 415 capsid proteins (Rosenberg et al., 1996). It was not possible to predict how many of the 27 kDa EYFPs can be displayed in combination with the library. To find the right ratio we constructed three frameshift constructs with a range of shift efficiencies. The highest ratios of EYFP fused capsid protein, approximately 50% and 25% (which would correspond to 200 and 100 fusions per phage head), did not sustain the correct formation of infectious phages. The frameshift sequence in the p10-EYFP helper plasmid directed the ribosome into the -1 frame to form the EYFP fusion in less than 1% of the translations. With this helper plasmid the capsid assembly was not hampered and amplification of a cDNA phage display library in *E. coli* was indistinguishable from the original helper plasmid. The presence of both the library derived proteins and the EYFP on the capsid surface was shown with the specific binding of these fluorescent phages to microspheres coated with specific antibodies, against which antigens were present in the library. It is likely that the capsid allows the inclusion of higher numbers of fluorescent fusions than tested in this study. Different frameshift sequences with recoding frequencies between 25% and 1% should be tested to find a ratio where most of the particles are fluorescent without interfering with the phage formation (Baranov et al., 2003; Gurvich et al., 2003). Still, although the frameshift frequency was not high enough to label every single phage particle, the fluorescent signal was sufficient for the detection of the phages and their interaction with the bait. We did not

test our frameshift helper plasmid in combination with the other available T7Select vectors. The T7Select415 was designed for display of peptides smaller than 50 amino acids on every capsid protein and does not need a helper plasmid for amplification. If it would be combined with our helper plasmid a considerable fraction (50%) of unfused capsid protein derived from the helper plasmid would also be incorporated, thus lowering the copy number of the peptide fusions from the library. T7Select1-1 and 1-2 are designed for the display of large (max.1000aa) and difficult to express proteins in a low copy number, which is comparable with the EYFP-fusion copy number resulting from the use of the frameshift helper plasmid. The fact that the library fragment would be incorporated in only 10% of the phages and the fluorescent fusion in approximately 14% means that an even smaller fraction of the phages will display both the library fragment and the EYFP-fusion. Thus, a larger number of phages would be necessary for screening.

The fluorescent phage particles could be characterised by means of microspectroscopic measurements. Fluorescence correlation spectroscopy (FCS) was applied to confirm the incorporation of gp10-EYFP fusion proteins into the phage capsids and to determine what percentage of the phages carried EYFP. FCS is a very sensitive technique and is based on the measurement of fluorescence fluctuations that are caused by fluorescent particles moving in and out of the confocal detection volume. Autocorrelation of these fluctuations gives information about the translational diffusion coefficient of the fluorescent particles and their relative numbers (Visser and Hink, 1999). The diffusion of particles with fluorescence characteristics of EYFP and a diffusion coefficient corresponding to the 60 nm phages was measured. Comparing the total concentration of phages produced at 28 °C with the number of fluorescent phages measured in the FCS experiments revealed that 14% of the phages is fluorescent under these conditions.

The large difference in size and thus diffusion coefficient between phages and most molecules that will normally be used as bait can be used to detect interactions between the two. Measuring the interaction between phage displayed proteins and their interactors by FCS has been attempted by Lagerkvist et al. (Lagerkvist et al., 2001). The interaction between an M13 phage displayed Fab fragment and its cognate antigen could be detected, but in their experiments the phages were labelled by fluorescent antibodies, causing aggregation and a significant loss of sensitivity in the detection of interactions.

Our data demonstrate that the enrichment of a library during affinity selection can be sensitively monitored with fluorescence microspectroscopy. When our library was enriched to 60% 5H6 binders by three rounds of bio-panning, the bright fluorescence of phages binding the 5H6-coated microspheres could easily be detected among the spheres coated with the anti-myc control antibody. Even after one round of panning, with approximately 5% of the phages displaying a 5H6-binding protein fragment, the difference was detectable between the fluorescence on the 5H6-coated beads and the beads with the control antibody. The specificity of the binding is evident by the lack of fluorescence on the beads coated with the control antibody. The clear signal of the fluorescent phage on the beads might be utilised in fluorescence based sorting systems like fluorescence activated bead sorting (FABS) or in microfluidic sorting systems. One could envisage even an setup wherein a fluorescent phage

display library is screened against a library of bait molecules coated on beads or spotted on a chip surface (Fu et al., 1999; Visser et al., 2004).

In principle there are alternative means to create fluorescent phages. Phages can for example be labelled with fluorescent chemical dyes (Gitis et al., 2002a; Gitis et al., 2002b; Oda et al., 2004). With a wide range of dyes available offering high quantum yields, very bright phage particles can be produced. In case of M13 phage the chemical dye labelling had no negative effect on the infectivity (Jaye et al., 2004). With T7 probably being more vulnerable we experienced a ten fold decrease of the infectivity after the chemical labelling (data not shown). Chemical labelling involves also extra handling steps and is needed each time the phages have been propagated in bacteria.

T7 is not the first phage that has been labelled with a fluorescent protein. Labelling of the lytic phage T4 with GFP has been reported. However, in these studies the phages were not used in a fluorescent phage display setup, but as model to study the T4 phage head structure or as detection system for *E. coli* (Mullaney and Black, 1998; Mullaney et al., 2000; Tanji et al., 2004). To display proteins on T4 phages fusions with either HOC (Highly antigenic Outer Capsid protein) or SOC (Small Outer Capsid protein) have been made (Ren et al., 1996; Jiang et al., 1997; Ren and Black, 1998; Malys et al., 2002). Both these proteins are not essential for particle formation and bind to the capsid with high copy number and in principle one could be used for labelling and one for display. However, T4 has the disadvantage of lacking an efficient in vitro packaging method, which makes it more difficult to create large display libraries. The proteins used for display in T4 are not incorporated in the plasmid, but bind to it with high affinity. This means the link between phenotype (the displayed protein) and genotype is not as strong as with T7 where the proteins used as display platform form an inextricable part of the capsid.

Combined with the standard bio-panning selection fluorescent display phages can give information about the enrichment of binders in a display library, which enables fine tuning of the selection procedure and obviates the need for more laborious ELISA or plaque lift assays. Furthermore, with the fast pace of developments in single molecule detection technologies and sorting systems, these fluorescent phages open the way to high throughput platforms for the direct selection of binding molecules (Rigler, 1995; Auer et al., 1998; Ma et al., 2000).

References

- Auer, M., Moore, K.J., Meyer-Almes, F.J., Guenther, R., Pope, A.J. and Stoekli, K.A.** (1998). Fluorescence correlation spectroscopy: lead discovery by miniaturized HTS. *Drug Discov. Today* **3**, 457-465.
- Baneyx, F.** (1999). Recombinant protein expression in *Escherichia coli*. *Curr. Opin. Biotechnol.* **10**, 411-421.
- Baranov, P.V., Gesteland, R.F. and Atkins, J.F.** (2002). Recoding: translational bifurcations in gene expression. *Gene* **286**, 187-201.
- Baranov, P.V., Gurvich, O.L., Hammer, A.W., Gesteland, R.F. and Atkins, J.F.** (2003). Recode 2003. *Nucleic Acids Research* **31**, 87-89.
- Bohmer, M. and Enderlein, J.** (2003). Fluorescence spectroscopy of single molecules under ambient conditions: Methodology and technology. *Chemphyschem* **4**, 793-808.
- Boonekamp, P.M., Pomp, H., Gussenhoven, G.C. and Schots, A.** (1991). The Use of Immunochemical Techniques and Monoclonal-Antibodies to Study the Viral Coat Protein-Structure of Potato Virus-a, Potato Virus-Y and Beet Necrotic Yellow Vein Virus. *Acta Botanica Neerlandica* **40**, 41-52.
- Brault, V. and Miller, W.A.** (1992). Translational Frameshifting Mediated by a Viral Sequence in Plant-Cells. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 2262-2266.
- Condrón, B.G., Atkins, J.F. and Gesteland, R.F.** (1991a). Frameshifting in Gene-10 of Bacteriophage-T7. *J. Bacteriol.* **173**, 6998-7003.
- Condrón, B.G., Gesteland, R.F. and Atkins, J.F.** (1991b). An Analysis of Sequences Stimulating Frameshifting in the Decoding of Gene-10 of Bacteriophage-T7. *Nucleic Acids Research* **19**, 5607-5612.
- Danner, S. and Belasco, J.G.** (2001). T7 phage display: A novel genetic selection system for cloning RNA-binding proteins from cDNA libraries. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 12954-12959.
- Devlin, J.J., Panganiban, L.C. and Devlin, P.E.** (1990). Random Peptide Libraries - a Source of Specific Protein-Binding Molecules. *Science* **249**, 404-406.
- Dittrich, P., Malvezzi-Cameggi, F., Jahnz, M. and Schwillle, P.** (2001). Accessing molecular dynamics in cells by fluorescence correlation spectroscopy. *Biol. Chem.* **382**, 491-494.
- Enderlein, J., Gregor, I., Patra, D. and Fitter, J.** (2004). Art and artefacts of fluorescence correlation spectroscopy. *Current Pharmaceutical Biotechnology* **5**, 155-161.
- Fu, A.Y., Spence, C., Scherer, A., Arnold, F.H. and Quake, S.R.** (1999). A microfabricated fluorescence-activated cell sorter. *Nat. Biotechnol.* **17**, 1109-1111.
- Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., Remor, M., Hofert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M.A., Copley, R.R.,**
- Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G. and Superti-Furga, G.** (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141-147.
- Geuijen, C.A., Bijl, N., Smit, R.C., Cox, F., Throsby, M., Visser, T.J., Jongeneelen, M.A., Bakker, A.B., Kruisbeek, A.M., Goudsmit, J. and de Kruijff, J.** (2005). A proteomic approach to tumour target identification using phage display, affinity purification and mass spectrometry. *Eur J Cancer* **41**, 178-187.
- Gitis, V., Adin, A., Nasser, A., Gun, J. and Lev, O.** (2002a). Fluorescent dye labeled bacteriophages--a new tracer for the investigation of viral transport in porous media: 1. Introduction and characterization. *Water Res* **36**, 4227-4234.
- Gitis, V., Adin, A., Nasser, A., Gun, J. and Lev, O.** (2002b). Fluorescent dye labeled bacteriophages--a new tracer for the investigation of viral transport in porous media: 2. Studies of deep-bed filtration. *Water Res* **36**, 4235-4242.
- Griffiths, A.D., Williams, S.C., Hartley, O., Tomlinson, I.M., Waterhouse, P., Crosby, W.L., Kontermann, R.E., Jones, P.T., Low, N.M., Allison, T.J., Prospero, T.D., Hoogenboom, H.R., Nissim, A., Cox, J.P.L., Harrison, J.L., Zaccolo, M., Gherardi, E. and Winter, G.** (1994). Isolation of High-Affinity Human-Antibodies Directly from Large Synthetic Repertoires. *Embo J.* **13**, 3245-3260.
- Gurvich, O.L., Baranov, P.V., Zhou, J., Hammer, A.W., Gesteland, R.F. and Atkins, J.F.** (2003). Sequences that direct significant levels of frameshifting are frequent in coding regions of *Escherichia coli*. *Embo J.* **22**, 5941-5950.
- Hink, M.A., Griep, R.A., Borst, J.W., van Hoek, A., Eppink, M.H., Schots, A. and Visser, A.J.** (2000). Structural dynamics of green fluorescent protein alone and fused with a single chain Fv protein. *J Biol Chem* **275**, 17556-17560.
- Horibe, T., Gomi, M., Iguchi, D., Ito, H., Kitamura, Y., Masuoka, T., Tsujimoto, I., Kimura, T. and Kikuchi, M.** (2004). Different contributions of the three CXXC motifs of human protein-disulfide isomerase-related protein to isomerase activity and oxidative refolding. *J. Biol. Chem.* **279**, 4604-4611.
- Ito, T., Tashiro, K., Muta, S., Ozawa, R., Chiba, T., Nishizawa, M., Yamamoto, K., Kuhara, S. and Sakaki, Y.** (2000). Toward a protein-protein interaction map of the budding yeast: A comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 1143-1147.
- Ivanov, I.P., Matsufuji, S., Murakami, Y., Gesteland, R.F. and Atkins, J.F.** (2000). Conservation of polyamine regulation by translational frameshifting from yeast to mammals. *Embo J.* **19**, 1907-1917.
- Jaye, D.L., Geigerman, C.M., Fuller, R.E., Akyildiz, A. and Parkos, C.A.** (2004). Direct fluorochrome labeling of phage display library clones for studying binding specificities: applications in flow cytometry

and fluorescence microscopy. *J. Immunol. Methods* **295**, 119-127.

Jiang, J., AbuShilbayeh, L. and Rao, V.B. (1997). Display of a PorA peptide from *Neisseria meningitidis* on the bacteriophage T4 capsid surface. *Infection and Immunity* **65**, 4770-4777.

Kimata, Y., Iwaki, M., Lim, C.R. and Kohno, K. (1997). A novel mutation which enhances the fluorescence of green fluorescent protein at high temperatures. *Biochem. Biophys. Res. Commun.* **232**, 69-73.

Krajcikova, D. and Hartley, R.W. (2004). A new member of the bacterial ribonuclease inhibitor family from *Saccharopolyspora erythraea*. *FEBS Lett.* **557**, 164-168.

Lagerkvist, A.C., Foldes-Papp, Z., Persson, M.A.A. and Rigler, R. (2001). Fluorescence correlation spectroscopy as a method for assessment of interactions between phage displaying antibodies and soluble antigen. *Protein Sci.* **10**, 1522-1528.

Larsen, B., Gesteland, R.F. and Atkins, J.F. (1997). Structural probing and mutagenic analysis of the stem-loop required for *Escherichia coli* dnaX ribosomal frameshifting: programmed efficiency of 50%. *J Mol Biol* **271**, 47-60.

Levin, M.E., Hendrix, R.W. and Casjens, S.R. (1993). A Programmed Translational Frameshift Is Required for the Synthesis of a Bacteriophage-Lambda Tail Assembly Protein. *J. Mol. Biol.* **234**, 124-139.

Li, M. (2000). Applications of display technology in protein analysis. *Nat. Biotechnol.* **18**, 1251-1256.

Licznar, P., Mejlhede, N., Prere, M.F., Wills, N., Gesteland, R.F., Atkins, J.F. and Fayet, O. (2003). Programmed translational-1 frameshifting on hexanucleotide motifs and the wobble properties of tRNAs. *Embo J.* **22**, 4770-4778.

Ma, Y., Shortreed, M.R. and Yeung, E.S. (2000). High-throughput single-molecule spectroscopy in free solution. *Anal Chem* **72**, 4640-4645.

Maiti, S., Haupts, U. and Webb, W.W. (1997). Fluorescence correlation spectroscopy: Diagnostics for sparse molecules. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 11753-11757.

Malys, N., Chang, D.Y., Baumann, R.G., Xie, D.M. and Black, L.W. (2002). A bipartite bacteriophage T4 SOC and HOC randomized peptide display library: Detection and analysis of phage T4 terminase (gp17) and late sigma factor (gp55) interaction. *J. Mol. Biol.* **319**, 289-304.

Mardon, G. and Varmus, H.E. (1983). Frameshift and Intragenic Suppressor Mutations in a Rous-Sarcoma Provirus Suggest Src Encodes 2 Proteins. *Cell* **32**, 871-879.

Matsufuji, S., Matsufuji, T., Miyazaki, Y., Murakami, Y., Atkins, J.F., Gesteland, R.F. and Hayashi, S. (1995). Autoregulatory Frameshifting in Decoding Mammalian Ornithine Decarboxylase Antizyme. *Cell* **80**, 51-60.

McHenry, C.S. (2003). Chromosomal replicases as asymmetric dimers: studies of subunit arrangement and functional consequences. *Mol. Microbiol.* **49**, 1157-1165.

Mullaney, J.M. and Black, L.W. (1998). Activity of foreign proteins targeted within the bacteriophage T4 head and prohead: Implications for packaged DNA structure. *J. Mol. Biol.* **283**, 913-929.

Mullaney, J.M., Thompson, R.B., Gryczynski, Z. and Black, L.W. (2000). Green fluorescent protein as a probe of rotational mobility within bacteriophage T4. *J Virol Methods* **88**, 35-40.

Oda, M., Morita, M., Unno, H. and Tanji, Y. (2004). Rapid detection of *Escherichia coli* O157:H7 by using green fluorescent protein-labeled PP01 bacteriophage. *Appl Environ Microbiol* **70**, 527-534.

Ren, Z.J. and Black, L.W. (1998). Phage T4 SOC and HOC display of biologically active, full-length proteins on the viral capsid. *Gene* **215**, 439-444.

Ren, Z.J., Lewis, G.K., Wingfield, P.T., Locke, E.G., Steven, A.C. and Black, L.W. (1996). Phage display of intact domains at high copy number: a system based on SOC, the small outer capsid protein of bacteriophage T4. *Protein Sci* **5**, 1833-1843.

Rigler, R. (1995). Fluorescence correlations, single molecule detection and large number screening. Applications in biotechnology. *J Biotechnol* **41**, 177-186.

Rosenberg, A., Griffin, G., Studier, F.W., McCormick, M., Berg, J. and Mierendorf, R. (1996). T7Select Phage Display System: A powerful new protein display system based on bacteriophage T7. *inNovations* **6**, 1-6.

Sche, P.P., McKenzie, K.M., White, J.D. and Austin, D.J. (1999). Display cloning: functional identification of natural product receptors using cDNA-phage display. *Chem. Biol.* **6**, 707-716.

Scott, J.K. and Smith, G.P. (1990). Searching for Peptide Ligands with an Epitope Library. *Science* **249**, 386-390.

Sheu, T.J., Schwarz, E.M., Martinez, D.A., O'Keefe, R.J., Rosier, R.N., Zuscik, M.J. and Puzas, J.E. (2003). A phage display technique identifies a novel regulator of cell differentiation. *J Biol Chem* **278**, 438-443.

Siple, J., Stassi, D., Dunn, J. and Goldman, E. (1991). Analysis of bacteriophage T7 gene 10A and frameshifted 10B proteins. *Gene Expr* **1**, 127-136.

Skakun, V.V., Hink, M.A., Digris, A.V., Engel, R., Novikov, E.G., Apanasovich, V.V. and Visser, A. (2005). Global analysis of fluorescence fluctuation data (vol 34, pg 323, 2005). *Eur. Biophys. J. Biophys. Lett.* **34**, 972-972.

Smith, G.P. (1985). Filamentous Fusion Phage - Novel Expression Vectors That Display Cloned Antigens on the Virion Surface. *Science* **228**, 1315-1317.

Sternberg, N. and Hoess, R.H. (1995). Display of Peptides and Proteins on the Surface of Bacteriophage-Lambda. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 1609-1613.

Tanji, Y., Furukawa, C., Na, S.H., Hijikata, T., Miyanaga, K. and Unno, H. (2004). *Escherichia coli* detection by GFP-labeled lysozyme-inactivated T4 bacteriophage. *Journal Of Biotechnology* **114**, 11-20.

Tong, A.H.Y., Drees, B., Nardelli, G., Bader, G.D., Brannetti, B., Castagnoli, L., Evangelista, M., Ferracuti, S., Nelson, B., Paoluzi, S., Quondam, M., Zucconi, A., Hogue, C.W.V., Fields, S., Boone, C. and Cesareni, G. (2002). A combined experimental and computational strategy to define protein interaction networks for peptide recognition modules. *Science* **295**, 321-324.

Tsien, R.Y. (1998). The green fluorescent protein. *Annu. Rev. Biochem.* **67**, 509-544.

Visser, A. and Hink, M.A. (1999). New perspectives of fluorescence correlation spectroscopy. *Journal Of Fluorescence* **9**, 81-87.

Visser, A., Kunst, B.H., Keller, H. and Schots, A. (2004). Towards sorting of biolibraries using single-

molecule fluorescence detection techniques. *Current Pharmaceutical Biotechnology* **5**, 173-179.

Yamamoto, M., Kominato, Y. and Yamamoto, F. (1999). Phage display cDNA cloning of protein with carbohydrate affinity. *Biochem. Biophys. Res. Commun.* **255**, 194-199.

Zheng, J.H. and McIntosh, M.A. (1995). Characterization of Is1221 from *Mycoplasma-Hyorhinis* - Expression of Its Putative Transposase in *Escherichia-Coli* Incorporates a Ribosomal Frameshift Mechanism. *Mol. Microbiol.* **16**, 669-685.

Zozulya, S., Lioubin, M., Hill, R.J., Abram, C. and Gishizky, M.L. (1999). Mapping signal transduction pathways by phage display. *Nat. Biotechnol.* **17**, 1193-1198.

Chapter 3

T7 cDNA phage display identifies highly basic (poly)-peptides as Rx1 CC-NB-ARC interactors

Erik Slootweg, Matthijs Jore, Lidwien van Sprundel, Hans Keller, Aska Goverse, Jaap Bakker and Arjen Schots.

To be submitted

Abstract

In this study, we have demonstrated in a pilot experiment the value of T7 phage display to identify specific interactors by using an antibody raised against the PVY coat protein. Screening of a PVY-infected *N. benthamiana* cDNA phage display library resulted in the selection of peptides harbouring the known PRIKAI epitope. Next, phage display was explored as technique to discover proteins interacting with the potato R protein Rx1. The system turned out to be prone to pick up interactors binding to matrices like Ni-NTA or to fusion proteins like thioredoxine. A possible way to circumvent this weakness was to design the selection procedure in such a way that it alternates between different matrices and to limit the number of selection round. This adapted approach resulted in the identification of a series of highly basic protein fragments and random peptides, for which a specific interaction could be shown. Two cDNA sequences encoded the ribosomal proteins L19 and L36a, which showed a stunted growth phenotype upon gene silencing in *N. benthamiana* using VIGS and a slightly reduced Rx-mediated HR.

Introduction

The largest class of plant resistance genes (R gene) encodes NB-LRR proteins (nibblers), which are able to recognize a wide range of pathogens upon invasion of the host. In plant genomes, hundreds of genes encoding such proteins containing the conserved nucleotide-binding (NB) core and C-terminal Leucine-rich repeat (LRR) region have been found (The Arabidopsis Genome Initiative, 2000; Meyers et al., 2003; Zhou et al., 2004; Tuskan et al., 2006; Kohler et al., 2008). For most of them no pathogen recognition specificity or function is known. Some might have functions other than pathogen recognition, for example downstream in the signalling pathways of other R proteins (Peart et al., 2005; Gabriels et al., 2007), or in plant development (Hewezi et al., 2006). Of the R proteins with known specificity only a few could be shown to directly interact with pathogen derived elicitors (Jia et al., 2000; Deslandes et al., 2003; Dodds et al., 2006). Indirect recognition mechanisms, via the detection of pathogen-induced modifications to host proteins, have been suggested to explain the apparent lack of direct interactions (van der Biezen and Jones, 1998; van der Hoorn and Kamoun, 2008). Interesting examples of such indirect recognition mechanisms have come to light in recent years. The *Arabidopsis* R proteins Rps2 and Rpm1 interact with RIN4 and recognise different modifications of this protein by *Pseudomonas* effector proteins (Mackey et al., 2002; Mackey et al., 2003). Another *Arabidopsis* R protein, Rps5, forms a complex with the kinase PBS1 and is activated once this kinase is digested by a pathogen-derived protease (Ade et al., 2007).

Downstream signalling is thought to be initiated by a conformational change of the R protein's central NB-ARC domain, coinciding with the exchange of the bound nucleotide (McHale et al., 2006; van Ooijen et al., 2007; Takken and Tameling, 2009). In analogy with more distantly related, but structurally similar, NB containing proteins, the R protein activation could lead to oligomerization and the activation of interacting downstream components. Some R proteins have been shown to interact with transcription factors (Shen et al., 2007; Caplan et al., 2008), but in most cases little is known about how R protein activation translates in the eventual induction of the resistance and cell death responses. The R protein Rx1 from potato mediates an highly efficient resistance response against Potato

Virus X (Bendahmane et al., 1999). Interactions between its subdomains, the N-terminal coiled-coil (CC), the central NB-ARC and the C-terminal LRR are required for its functioning and have been studied in detail (Moffett et al., 2002; Rairdan and Moffett, 2006; Rairdan et al., 2008). Through the CC domain Rx1 interacts with RanGAP2 (RanGTPase Activating Protein) and this interaction stimulates Rx1's activation (Sacco et al., 2007; Tameling and Baulcombe, 2007). The exact role of this interaction in Rx1 signalling is not yet known.

R proteins are thought to be present in the cell as part of a larger protein complex. The modular architecture of R proteins suggests that they form a scaffold for various interacting proteins, involved in pathogen recognition, downstream signalling or protein stabilization. For example, SGT1 and HSP90 (both essential for Rx1 signalling) have been shown to interact with the LRR domain of several R proteins (Bieri et al., 2004; de La Fuente van Bentem et al., 2005; Leister et al., 2005). However, few common interactors have been found for the CC-NB-ARC domains despite extensive screenings for downstream interactors (Lukasik and Takken, 2009), with the exception of the Arabidopsis CRT1 ATPase protein that interacts with several NB-ARC domains (Kang et al., 2008).

In this study, cDNA phage display was applied as an alternative method to identify additional downstream Rx1 interactors, which could further resolve the Rx1 signalling pathway. Ever since their introduction in the mid-eighties, phage display technologies have been used successfully to select and characterize antibodies (Smith, 1985; Smith and Petrenko, 1997). This technique enables the construction of large libraries of phage clones displaying peptides as a covalent fusion to bacteriophage capsid proteins. Because the DNA encoding the displayed peptides is enclosed in the phage capsid, the phenotype and genotype are linked. This enables the powerful selection procedure called bio-panning: iterative cycles of affinity selection and phage amplification during which the proportion of phages displaying the binding peptides is highly enriched in the total phage population (Fig. 1).

The filamentous phage M13 has for long been the most extensively applied phage for display purposes, but this system has its weaknesses. Each capsid protein-fusion has to be transported through the bacterial membranes to be assembled into the phage particle. This creates a bias in the display libraries against peptides that cannot efficiently be transported to the membrane (Krumpe et al., 2006). This problem become particularly apparent when

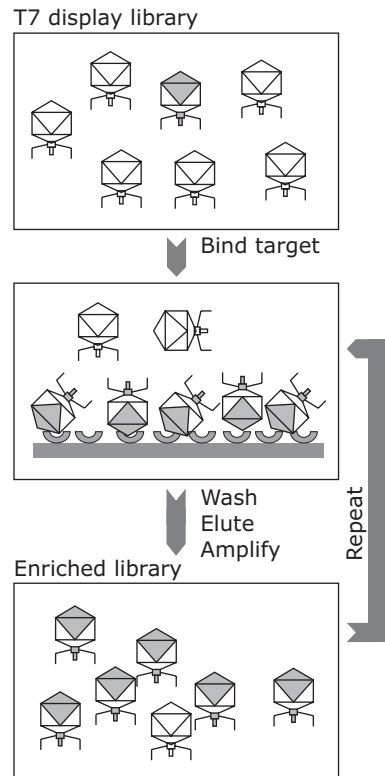


Figure 1. Schematic overview of the bio-panning procedure. A display library of phages exhibiting various peptides on their capsids is incubated with matrix-bound bait protein. Some phages bind the bait protein and are eluted and amplified. Repeating this cycle will enrich the library in phages displaying bait-interacting peptides.

cDNA encoded proteins are displayed. Lytic phage systems circumvent these disadvantages (Castagnoli et al., 2001). Their capsids are assembled in the bacterial cytoplasm and eventually released by lysis of the bacteria, without the need of transport through the cell membrane. Lytic phage cDNA display has been used successfully in the identification of RNA-binding proteins (Danner and Belasco, 2001), protein-protein interactions (Bukanov et al., 2000; Houshmand and Bergqvist, 2003), and proteins binding chemical compounds (Yamamoto et al., 1999; Savinov and Austin, 2001; Jin et al., 2002; Aoki et al., 2005).

Here, we used the commercially available T7 Select system (Novagen) based on lytic phages to display a *Nicotiana tabacum* cDNA library and to select protein fragments interacting with the N-terminal signalling domains of the CC-NB-LRR protein Rx1. In a pilot experiment to optimise the system, a PVY infected *N. benthamiana* cDNA library was screened for interactors to the monoclonal anti-PVY antibody (Boonekamp et al., 1991; Keller et al., 2005). This resulted in the successful identification of the PVY CP region encompassing the known epitope PRIKAI (Keller et al. 2005). The screening of the *N. tabacum* cDNA library for Rx1 CC-NB-ARC interactors resulted in a series of highly basic protein fragments and random peptides. Several of these could be shown to bind to Rx1 and not to GST or TRX in pull-down experiments. The biological relevance of Rx1's tendency to bind to positively charged peptides and the limitations of cDNA phage display for the selection of relatively low affinity interactions are discussed.

Results

Construction of a T7 cDNA display library from PVY-challenged N. benthamiana

When creating a T7 cDNA display library for interaction screenings certain conditions have to be taken into consideration. First of all there is a trade-off between the size of the displayed peptide and the frequency of the fusion of this peptide in the phage capsid. High frequencies of large peptides have an adverse effect on the capsid formation and would thereby hinder the library replication (Rosenberg et al., 1996). However, larger peptides could encompass full proteins or functional domains, whereas smaller peptides only represent small fragments of proteins and are less likely to retain the biologically relevant interaction surfaces or functions. Another point for consideration is the size of the library. Even though a directional cloning strategy is applied, only about a third of the cDNA fragments will be fused in the correct reading frame and direction. To ascertain that most potential interactors from the transcriptome are represented, the library must have a coverage of several times the number of transcripts, especially if the library consists of small cDNA fragments.

To test the suitability of a cDNA fragment library displayed at 3-5 copies on the T7 capsid for the selection of plant gene encoded interactors, a small test library was created. The interaction between the monoclonal antibody Y-5 and the coat protein (CP) of Potato Virus Y (PVY^C) was used to assess the efficiency of library enrichment and interactor selection. This antibody has been developed as a diagnostic tool for the detection of PVY infections in potatoes (Boonekamp et al., 1991). It recognises a conserved six amino acid epitope PRIKAI in the PVY coat protein, as has been determined earlier by screening against a peptide phage

library (Keller et al., 2005). The PVY coat protein is, part of a polyprotein encoded by a large viral mRNA and the mature protein is derived by proteolytic cleavage (Robaglia et al., 1989). *Nicotiana benthamiana* plants were infected with the PVY^C strain. After one week, the spreading of PVY throughout the infected plants was shown by a direct antibody sandwich-ELISA with the monoclonal antibodies 5C12 and Y-5 on protein extract from leaf material (data not shown). Next, a T7 library was constructed from PVY-infected plant material and as control, PCR was performed with PVY specific primers showing the presence of the coat protein encoding cDNA in the library. Determining the titer of the library after packaging revealed that it consisted of 4×10^5 individual clones.

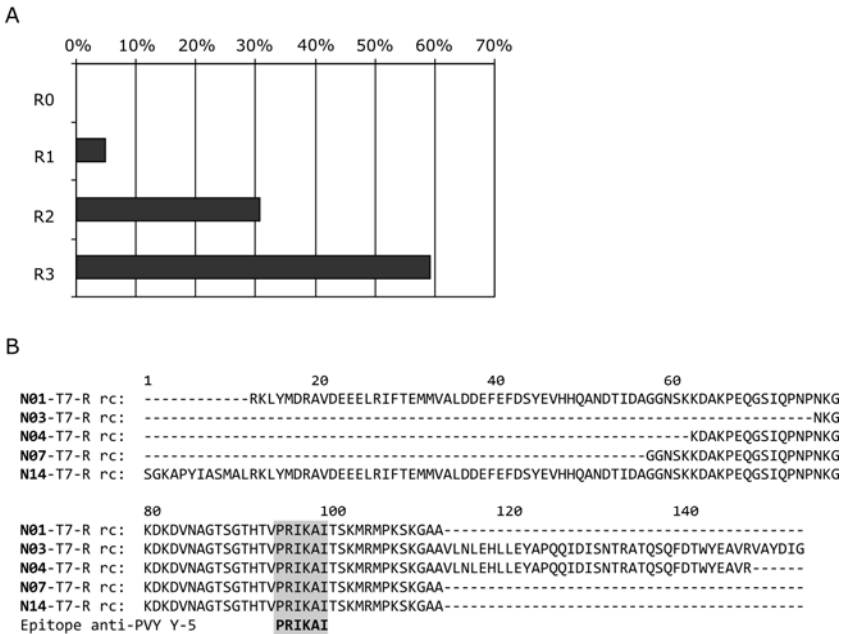


Figure 2. A. Enrichment of the T7 cDNA display library in rounds of selection against the anti-PVY antibody Y-5. The percentage of binders was determined by a plaque lift assay. Interacting phages were visualized by immunoblotting with AP-conjugated Y-5 antibody. **B.** Alignment of five displayed peptides interacting with Y-5 anti-PVY, resolving the PRIKAI epitope (highlighted in grey) as determined by Keller et al., (2005).

Selection of protein fragments from the cDNA library binding the Y-5 antibody

T7 phages displaying protein fragments with a high binding affinity for the Y-5 antibody were selected by bio-panning; successive rounds of affinity selection and amplification (Fig. 1). In the first selection round $3 \mu\text{g}$ of Y-5 antibody was coated to a microtiter well. 1.6×10^8 phages, 400 times the number of unique clones in the library, were incubated with the coated antibody. A total of three panning rounds were performed in this way. The stringency of the selection was increased in the last two rounds by adding 0.1% Tween-20 during the co-incubation of the phage and the antibody, and by including additional washing steps. Finally, an ELISA test showed that the affinity of the selected phage

population for the bait had increased (data not shown). Although extra panning rounds could result in a population that would have a higher proportion of binding phages, there would be a risk of reducing variation and introducing a bias towards smaller, more efficiently expressed fragments.

To discriminate between the individual phage clones that bound to the Y-5 antibody and the rest of the phage population, a plaque lift assay and subsequent bait protein overlay were performed. A similar number of phages from every selection round was plated and lifted onto a nitrocellulose filter. An alkaline phosphatase conjugated Y-5 antibody was applied to identify plaques of binding phages. Already in the phage population resulting from the first selection round, about 5% of the plaques were interacting with the Y-5 antibody. After three rounds, this number had increased to almost 60 % of the phages interacting with the Y-5 antibody (Fig. 2A).

The cDNA inserts of forty individual positive clones from the third selection round were amplified by PCR for further analysis with primers flanking the cDNA cloning site in the T7 genome. After agarose gel electrophoresis of the PCR products, it became apparent that these forty clones contained only 9 DNA fragments differing in size. The sizes of the cDNA inserts ranged from 200 to 500 base pairs and the most common fragment (12x coverage) was 270 base pairs long. All inserts encoded parts of the PVY coat protein, and no plant proteins were selected in the screening. By aligning the translated sequence of the five smallest fragments a minimal overlapping area could be determined of 37 amino acids, including the known epitope PRIKAI (Fig. 2B alignment). From a relatively small cDNA library, specific interacting fragments could be isolated after only three selection rounds, demonstrating the value of cDNA phage display for the identification of protein-protein interactions.

Screening a T7 cDNA display library from N. tabacum against the bait R13-V5H6
Rx1 is able to confer resistance against avirulent PVX strains as transgene in *N. tabacum* and *N. benthamiana* (Bendahmane et al., 2000). The full functionality of Rx1 in these *Nicotiana* species means that the protein-protein interactions involved in Rx1 signalling do occur in the *Nicotiana* background. Because *Nicotiana* is a more suitable model plant for further analysis of possible interacting proteins, this system was chosen to search for interactors. The transgenic *N. tabacum* T14.4 line (kindly provided by the Sainsbury Laboratory, UK) contains *Rx1* under control of its endogenous regulatory sequences and was used as source for mRNA in creating a T7 cDNA display library. Determining the initial titer after packaging revealed a library size of approximately $1 \cdot 10^7$ individual clones, about 25 times larger than the library used for the pilot selection study. The average length of the inserts was 600 bp and the distribution of fragment lengths as determined by PCR for 52 random clones is shown in Table 1.

Table 1. Insert size (bp) distribution of the T7 cDNA display library from *N. tabacum*

Insert size (bp)	Frequency
1500-850	13%
850-500	44%
500-250	29%
250-1	14%

For the selection of specific binders from a phage display library, the target protein (bait) needs to be purified and attached to a matrix to allow incubation with the phages and subsequent washing steps. Full-length *Rx1* and constructs containing either the N-terminal CC-NB-ARC domains or the C-terminal LRR domain were expressed in *E. coli* Top10 via the arabinose inducible pBAD vector system. The *Rx1* constructs contained a C-terminal fusion to a V5-epitope tag (GKPIPNLLGLDST) and a 6xHis-tag. The expression of full-length *Rx1* or the LRR domain construct proved to be challenging, and we were not able to acquire sufficient amounts of protein needed for bio-panning. The *Rx1* CC-NB-ARC domain on the other hand could be expressed as soluble and stable protein in *E. coli* and it was decided to continue with only this *Rx1* fragment. The *Rx1* CC-NB-ARC-V5-6H (R13-V5H6) was purified from bacterial extract by affinity purification on a Ni-NTA column (Fig. 3A) and used as target protein in subsequent screening of the *N. tabacum* T7 cDNA display library.

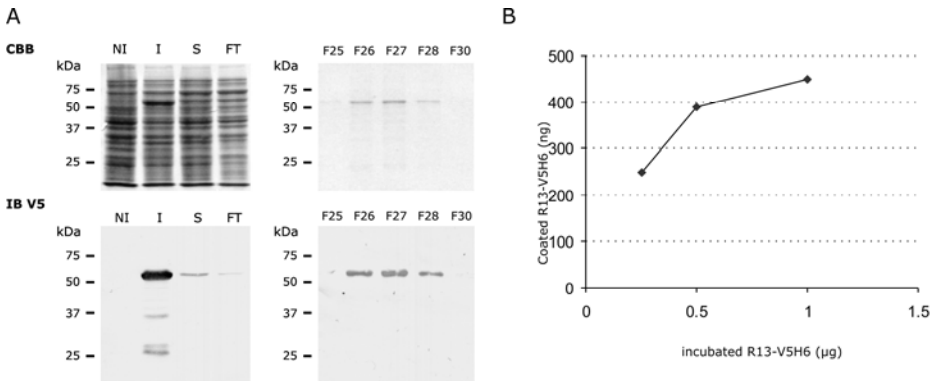


Figure 3. A. Purification of R13-V5H6 by Ni-NTA. Coomassie brilliant blue (CBB) stained 10% SDS-PAGE gel and anti-V5 immunoblot (IB V5) showing total extract of non-induced bacteria (NI), total extract of induced bacteria (I), soluble fraction of the protein extract after induction (S), and the flow-through of the Ni-NTA column (second CBB gel) showing the fractions in which R13-V5H6 was eluted from the column (F25-F30). R13-V5H6 is detected on the immunoblot as a band of approximately 60 kDa. **B.** 0.25 μg, 0.5 μg and 1.0 μg of purified R13-V5H6 was incubated in 100 μL coating buffer in a microtiter plate well overnight at 4°C. After washing 3 times the amount of coated protein was determined with a BCA assay calibrated against a BSA concentration range.

Initial affinity selection experiments showed that the use of an identical matrix to attach the target protein resulted in a strong selection for aspecific binders to this matrix from the cDNA library, which competed with the potential target binding proteins (data not shown). For example, when Ni-NTA beads were used to attach the target protein, high frequencies of cDNAs encoding histidine-rich peptides were found. When only microtiter wells were used for coating the target protein, high frequencies of hydrophobic peptides rich in tryptophans were found (data not shown). To avoid this, two changes in the selection procedure were implemented. In the first place, the type of support matrix for the target protein was alternated between the selection rounds. Secondly, an additional screening of individual phage clones was introduced after the selection was completed to distinguish target interacting phages from matrix interacting phages. Furthermore, the efficiency of the target protein coating was determined to minimise the available uncoated surface for matrix binding. The surface coating was saturated at about 450 ng of R13-V5H6 protein when 1 µg of protein was incubated per well for 12 hours (Fig. 3B).

Identification of interacting proteins of R13-V5H6 using T7 phage display

In total, 8 rounds of panning were performed with R13-V5H6 (Table 2). For the first selection round the target (bait) protein was coated in a 4 ml NUNC MaxiSorb tube to maximise the binding surface and allow the incubation with 3×10^{10} phages (1 mL, 3000x coverage of the library diversity). For round 2, 3, 6, and 7 the bait protein was coated in wells of a 96 well microtiter plate. For selection round 4, 5, 8 the bait protein was attached to paramagnetic Ni-NTA agarose beads (Table 2).

After each round the change in total binding capacity for the enriched phage population was monitored by determining the titer of phages bound to a fixed amount of coated R13-V5H6 protein. The ratio between input phage and binding phage increased in the first round to $1 \times 10^5:1$ and decreased steadily to $1 \times 10^2:1$ in round 8 (Table 3). The phages from the last selection round were further analyzed.

Table 2. Selection of R13-V5H6 binders from a *N. tabacum* T7 cDNA library in 8 rounds of bio-panning against R13-V5H6. The matrix used to attach the bait protein was in the first round a Nunc Maxisorb tube, and in the other rounds either microtiter plate wells (P) or Ni-NTA agarose (N).

Round	Matrix	Phage input	Eluted phage	Input/Eluted
1	Nunc	3.2×10^{10}	3.6×10^5	8.9×10^4
2	P	1.0×10^{10}	7.2×10^6	1.4×10^3
3	P	1.6×10^9	1.4×10^5	1.1×10^4
4	N	2.5×10^9	2.5×10^5	1.0×10^4
5	N	5.0×10^8	1.5×10^5	3.3×10^3
6	P	1.6×10^9	3.6×10^6	4.4×10^2
7	P	1.0×10^9	2.4×10^6	4.2×10^2
8	N	1.0×10^8	5.6×10^5	1.8×10^2

Table 3. Binding assay to monitor enrichment. Fixed amounts of phage from each selection round (1-8) and from the unselected T7 library (0) were incubated with coated R13-V5H6. The amount bound phage was determined by titering after several washing steps. The table gives the ratio of input:eluted phage.

Selection round	Input phage/ Eluted phage
0	6.0×10^4
1	1.0×10^5
2	1.6×10^5
3	1.3×10^4
4	6.4×10^3
5	3.7×10^2
6	5.2×10^2
7	3.0×10^2
8	1.0×10^2

The resulting phage populations are a mixture of non-binding, specific target-binding and matrix-binding peptide displaying phages, enriched in R13-V5H6 binders. To identify individual bait protein binding phage clones a screening system was set up. Phages were plated on *E. coli* in Petri dishes. Individual plaques, representing monoclonal phages, were picked and amplified in 96 well plates. These monoclonal phages were analysed in a binding assay for their interaction with the bait protein and the binding matrix (after blocking with BSA). The levels of binding T7 phage were detected by a monoclonal anti-T7 antibody and a peroxidase conjugated secondary antibody. Peroxidase activity was visualised via the oxidation of the ABTS (2,2'-azino-bis(3-ethylbenzthiazoline-6-sulphonic acid) substrate. After screening 1,000 phage clones from the R13-V5H6 selection, unique inserts were identified that did bind R13-V5H6 specifically, and not Ni-NTA or the microtiter well. Sequencing showed that five of these were in-frame protein fragments (Fig. 4A), and nine consisted of peptides formed by out-of-frame translation of cDNA inserts (Fig. 4B).

Three out of the five in-frame protein fragments could be identified as ribosomal proteins. R0-R1 shared the highest homology to the C-terminal 180 residues of the ribosomal protein L2. R2-A11 was homologous with the central 87 residues of ribosomal protein L19, and R2-B11 showed homology with the N-terminal 100 residues of ribosomal protein L36a (Fig. 4A.). The remaining two in-frame protein fragments shared most similarity with a putative myosin-heavy chain-like protein (R0-R11) and with an unknown DUF1022 domain-containing protein called ELM1 (R1-B6).

A striking characteristic of all the identified peptides was the large proportion of positively charged residues in their amino acid sequences. The average isoelectric point (pI) of the non-ORF peptides was 11.5 (± 0.5) and of the protein fragments 10.1 (± 2.1). An analysis of the average amino acid composition of the protein fragments (Fig. 4C) and the non-ORF peptide

fragments (Fig. 4D) shows that they are similar in composition. Both contain high fractions of arginine (R), lysine (K), leucine (L), serine (S) and glycine (G).

Proteins with isoelectric points in this range are usually only found in the nucleus, for example DNA binding proteins like histones and the DNA-binding domains of transcription factors. A search for proteins in the Swiss-Prot protein database with amino acid compositions similar to the ones found in the non-ORF peptides identified mainly ribosomal proteins (data not shown).

A Protein fragments

R0R1, 135 AA
 AIHNIEITLGGKQLARAAG AVAKLIAEKG KSATLKLPSG EVLRLSKNCS ATVQGNQVW NQKSLGRAGS
 KRWLGRPV RGVNPNVPHV PHGGEGRAP IGRKKPTTPW GYPALGRSK KRNKYSNLI FXPRS
 Theoretical pI: 11.43
 Negatively charged residues (D + E): 6 / 4%
 Positively charged residues (R + K): 25 / 19%

BLAST-P Match
 GENE ID: 800420 rp12 | 50S ribosomal protein L2 [Nicotiana tabacum]
 Score = 254 bits (650), Expect = 1e-66, Method: Compositional matrix adjust.
 Identities = 131/136 (96%), Positives = 131/136 (96%), Gaps = 1/136 (0%)

Query 1 AIHNIEITLGGKQLARAAGAVAKLIAEKGK SATLKLPSG EVLRLSKNCSATVQQ-GNWG 59
 Sbjct 138 AIHNIEITLGGKQLARAAGAVAKLIAEKGK SATLKLPSG EVLRLSKNCSATVQQ GNWG 197

Query 60 VVQKSLGRAGSKRWLGRKRPVVRGVVMPVDPHGGEGRAPIGRKKPTTPWGYPALGRRS 119
 Sbjct 198 VVQKSLGRAGSKRWLGRKRPVVRGVVMPVDPHGGEGRAPIGRKKPTTPWGYPALGRRS 257

Query 120 XRNKYSNLI FXPRS 135
 Sbjct 258 XRNKYSNLI RRS 273

R0R11, 76 AA
 EESRLVAKK EIEAAKASEK LALEAINALQ ESELARSSND EDSPSGVTL S LKEYFDLSKW PMRQRKRLIR
 GWLLLS
 Theoretical pI: 6.61
 Negatively charged residues (D + E): 13 / 17%
 Positively charged residues (R + K): 13 / 17%

BLAST-P Match
 >gb|EEF52409.1| Paramyosin, putative [Ricinus communis] Length=879
 Score = 79.3 bits (194), Expect = 9e-14, Method: Compositional matrix adjust.
 Identities = 43/58 (74%), Positives = 52/58 (89%), Gaps = 0/58 (0%)

Query 2 ESRLVAKK EIEAAKASEK LALEAINALQ ESELARSSND EDSPSGVTL S LKEYFDLSK 59
 Sbjct 656 ESRLVAKK EIEAAKASEK LALEAINALQ ESELARSSND EDSPSGVTL S LKEYFDLSK 713

R1B6, 127 AA
 DVVSNQIKVT LSGVFKDST TNVPKGLGAS TLMDSRETA NNRLNPRRS EKISTGKMRN IMSDFKETF
 TSSMKNL REGLTRPFT GLEDMSESYS YPPLNDATA ASRVNEALAE RQWKIKA
 Theoretical pI: 9.47
 Negatively charged residues (D + E): 15 / 12%
 Positively charged residues (R + K): 20 / 16%

BLAST Match
 - AA 1-81 were identified as an out-of-frame translated sequence of
 N. tabacum Salicylic Acid-binding catabase (SGN-U44741)
 - AA 82-127:
 GENE ID: 832295 ELM1 | ELM1 (ELONGATED MITOCHONDRIA 1) [Arabidopsis thaliana]
 Score = 72.0 bits (175), Expect = 1e-11, Method: Compositional matrix adjust.
 Identities = 31/46 (67%), Positives = 37/46 (80%), Gaps = 0/46 (0%)

Query 1 ERGLTRPFTGLEDMSESYSYPLNDATAASRVNEALAE RQWKIKA 46
 Sbjct 382 ERGVVRSFTGFEDMSESYSYPLNDATAEAATIRRELAARGLRS 427

R2A11, 88 AA
 RKGTREARLP TKVLMRRMR VLRLLLKRYR ESKIDRHHY HDMYMKVGNV VFKNKRVLME NIHKTKAEKA
 REKLSDFE ARAANK
 Theoretical pI: 11.35
 Negatively charged residues (D + E): 9 / 10%
 Positively charged residues (R + K): 30 / 34%

BLAST-P Match
 >gb|AAR83877.1| 60S ribosomal protein L19 [Capsicum annuum] Length=232
 Score = 176 bits (445), Expect = 8e-43, Method: Compositional matrix adjust.
 Identities = 86/87 (98%), Positives = 87/87 (100%), Gaps = 0/87 (0%)

Query 1 RKGTREARLPTKVLWRRMRVLRLLRKYRESKIDRHHYHDMYMKVGNVFNKRVLME 60
 Sbjct 81 RKGTREARLPTKVLWRRMRVLRLLRKYRESKIDRHHYHDMYMKVGNVFNKRVLME 140

Query 61 NIHKTKAEKAREKLSDFEARRAKN 87
 Sbjct 141 NIHKTKAEKAREKLSDFEARRAKN 167

R2B11, 94 AA
 LAVGLNKGHV VTKELAPRP SDRKGTSKRV VLFVRSILRE VAGFAPYEKR ITELKVGKD KRALKVAKRK
 LGTHRAKK REEMSSVLRK MRTK
 Theoretical pI: 11.46
 Negatively charged residues (D + E): 8 / 9%
 Positively charged residues (R + K): 31 / 33%

BLAST-P Match
 >gb|EEF49092.1| 60S ribosomal protein L36e, putative [Ricinus communis] Length=107
 Score = 174 bits (440), Expect = 3e-42, Method: Compositional matrix adjust.
 Identities = 90/93 (96%), Positives = 91/93 (97%), Gaps = 0/93 (0%)

Query 1 LAVGLNKGHVTKELAPRPSDRKGTSKRVVLFVRSILREVAGFAPYEKRITELKVGKD 60
 Sbjct 7 LSVGLNKGHVTKELAPRPSDRKGTSKRVVLFVRSILREVAGFAPYEKRITELKVGKD 66

Query 61 KRALKVAKRKLGTHRAKKREEMSSVLRKMR 93
 Sbjct 67 KRALKVAKRKLGTHRAKKREEMSSVLRKMR 99

B non-ORF peptides

R4B5, 5x, 33AA
 ARQPMKNGY RDHLRRHWK RLSSYLGRIG HMG
 Theoretical pI: 11.73
 Negatively charged residues (D + E): 1 / 3%
 Positively charged residues (R + K): 10 / 30%

R8F7, 4x, 68 AA
 REEKIGKSDY GGRYGRSVO ERGREGRR PKMLQWVRA RLPEETVMS
 VKKMLWQPL LAHQEGL
 Theoretical pI: 10.92
 Negatively charged residues (D + E): 8 / 12%
 Positively charged residues (R + K): 17 / 25%

R7A5, 2x, 82 AA
 CQTFVTSAR SRSTNIIYKI KHTHTVTGP YKTFSTRMS RMTTFSRLK
 RQNKQENPIY ISQVRRTRA SGNPTGILV CK
 Theoretical pI: 11.53
 Negatively charged residues (D + E): 1 / 1%
 Positively charged residues (R + K): 16 / 20%

R6G3, 2x, 100 AA
 HRRFQIQIKR SPRTRGRGK SQTQMLRNQ IGRMHQRL KMRKLSRPR
 GWMTLKRSM ILRQSLKIG LMRKMNGLR QKLITRSVPR ALDVGSGEDQ
 Theoretical pI: 12.38
 Negatively charged residues (D + E): 4 / 4%
 Positively charged residues (R + K): 26 / 26%

R1A2, 1x, 66 AA
 REEKIGKSDY GGRYGRSVO ERGREGRR PKMLQWVRA RLPEETVMS
 VKKMLWQPL LAHQEGL
 Theoretical pI: 10.89
 Negatively charged residues (D + E): 8 / 12%
 Positively charged residues (R + K): 16 / 24%

R5B9, 1x, 38 AA
 GRRKXKLSN ISNKLPLKSD NRVCSNMKL SLLSLACRQ
 Theoretical pI: 11.14
 Negatively charged residues (D + E): 1 / 3%
 Positively charged residues (R + K): 11 / 29%

R5B12, 1x, 35 AA
 GNLKEEQRL SLLRHMNM NIFYWGVKVK SPTLR
 Theoretical pI: 11.48
 Negatively charged residues (D + E): 2 / 6%
 Positively charged residues (R + K): 9 / 26%

R6G5, 1x, 40 AA
 KKYRGRQRP RKKQVLMKEN LLSILQVMS RQQRNQQR
 Theoretical pI: 12.02
 Negatively charged residues (D + E): 1 / 3%
 Positively charged residues (R + K): 15 / 38%

R3M4, 1x, 32 AA
 RKMKGGRR PDGERKRS RMIEYLRHM LL
 Theoretical pI: 11.80
 Negatively charged residues (D + E): 3 / 9%
 Positively charged residues (R + K): 13 / 41%

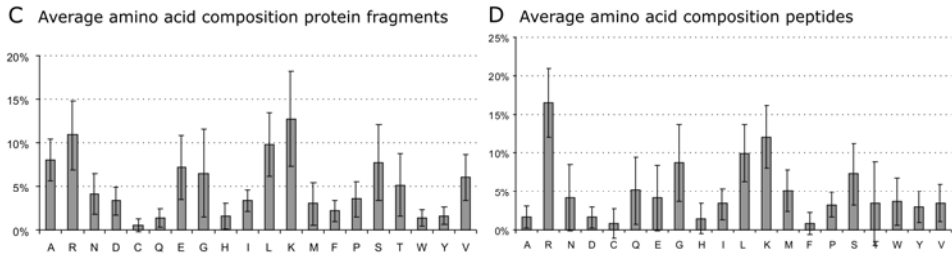


Figure 4. Overview of the *N. tabacum* cDNA sequences encoding protein fragments (A) or non-ORF peptides (B) that bind to R13-V5H6 in a T7 phage display screening. The average amino acid composition of the protein fragments (C) and non-ORF peptides is shown in a diagram (error bars depict standard error).

Confirmation of the interactions with R13-V5H6 in a pull down assay

The overall basic nature of these proteins could either mean that in the selection procedure we inadvertently selected for positively charged peptides, or that the interaction with basic proteins was a characteristic of the Rx1 CC-NB-ARC domain. However, in the selection of the monoclonal phages displaying these peptides the aspecific interaction with the polystyrene walls of the ELISA plate wells and the Ni-NTA beads were excluded. To test if the selected peptides bound specifically to the Rx1 CC NB-ARC when they were no longer displayed on phages, their interaction was tested in a pull-down assay. All 5 in-frame protein fragments (R0-R1, R0-R11, R1-B6, R2-A11, R2-B11) and one of the non-ORF peptides (R1-A2) were produced in *E. coli* as fusions to thioredoxine (TRX). These TRX fused peptides were used in a pull-down assay with GST-R13 or free GST as negative control to determine if they specifically bound the Rx1 CC-NB-ARC. Bacterial protein extracts containing the bait (GST-R13 or GST) were mixed with protein extracts containing either free TRX or the TRX fused peptide derived from the phages. Glutathione beads were applied to pull-down the GST fused constructs and the complexes were specifically eluted with glutathione. Western blotting with anti-TRX or anti-GST antibodies was used to detect if the potential interacting peptides were pulled down with the GST-R13 or GST proteins (Fig. 5). One construct, R0-R1 (ribosomal protein L2), could be pulled down with both GST and GST-R13, and was therefore considered an aspecific binder. All other fragments could be pulled-down specifically with GST-R13, but not with GST alone. The non-ORF peptide R1-A2 showed the strongest binding. The confirmed interaction with the Rx1 CC-NB-ARC and the absence of homology between the selected peptides, other than their amino acid composition and high pI, led us to believe that binding to basic peptides is a characteristic of the Rx1 CC-NB-ARC.

Functional analysis of the interactors in *N. benthamiana* using VIGS

To study if the detected interactions with the protein fragments were relevant for the functioning of Rx1 in the host cell, the genes encoding the four protein fragments were silenced in *Nicotiana benthamiana* via TRV Induced Gene Silencing (VIGS) (Fig. 6A). Twenty days after inoculation with the TRV silencing vectors the new leaves and stem of the *N. benthamiana* plants infected with the control Phytoene desaturase (PDS) TRV showed an almost complete bleaching phenotype. Plants silenced for R2-A11 (ribosomal protein L19)

and R2-B11 (ribosomal protein L36a) showed severely stunted growth and reduced levels of chlorophyll. Plants infected with the R1-B6 TRV vector (unknown protein with DUF1022 domain) or R0-R11 (putative myosin-heavy chain-like protein) showed no visible phenotype and could not be distinguished from non-infected plants. Three weeks after infection with the TRV vector the plants silenced for R0-R11, R1-B6, R2-A11 and R2-B11, and plants infected with empty silencing vector were co-infiltrated with Rx1 and the avirulent CP106 or virulent CP105 in an agroinfiltration assay (Fig. 6B). In the empty vector, R0-R11 and R1-B6 TRV plants the Rx1 mediated hypersensitive response developed without difference in strength or timing. In the R2-A11 and R2-B11 plants the HR did not develop fully. The phenotypes of these plants suggest that they are severely hampered in protein synthesis, which would be expected for plants in which components of the ribosome are silenced. In our opinion, it is not possible to separate a possible role in Rx1 signalling of these proteins from their role in protein synthesis.

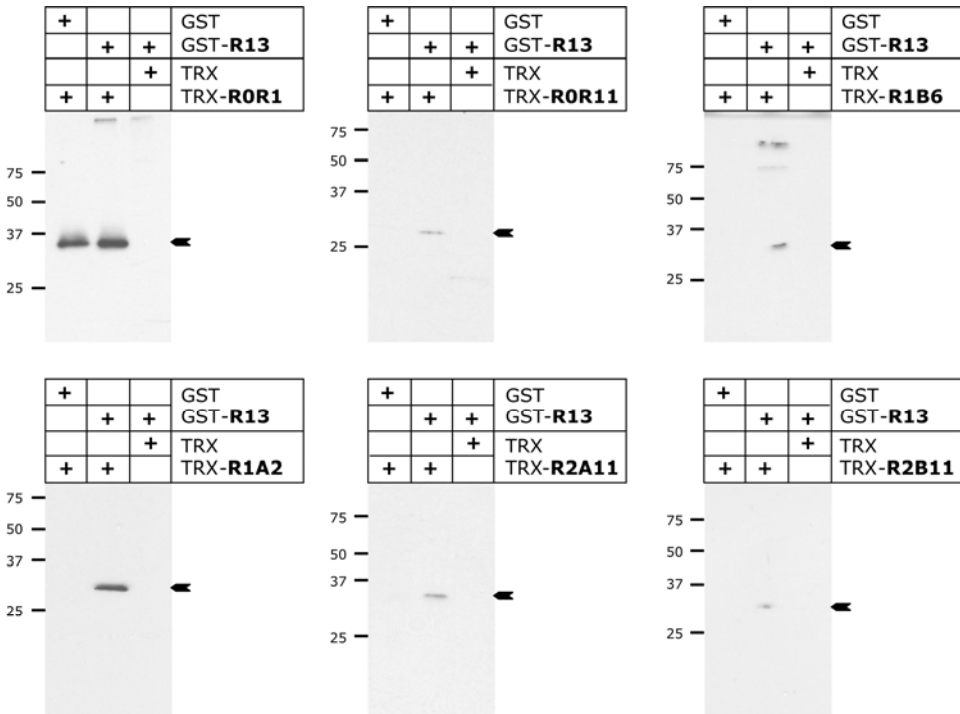
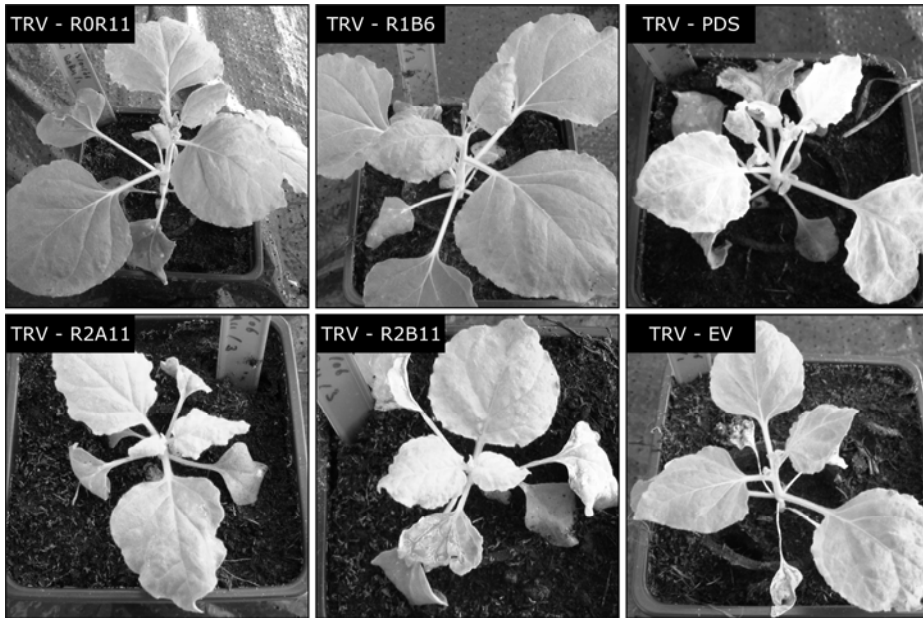


Figure 5. GST pull-down experiment. Purified Glutathione-S-Transferase (GST) fusion product of the Rx1 CC-NB-ARC (R13) was incubated with a protein extract containing thioredoxine (TRX) fusion products of the potential interactors. The GST protein was bound to glutathione sepharose 4B beads and eluted with a glutathione solution. Western blotting with anti-TRX antibodies was used to detect if the potential interacting peptides fused to TRX were pulled down with GST-R13 (target) or GST (control).

A



B

	CP106	GFP
EV	+	-
R0R11	+	-
R1B6	+	-
R2A11	+/-	-
R2B11	+/-	-

Figure 6. A. Phenotypes of TRV induced gene-silencing in *N. benthamiana* at 18 days post inoculation. R0-R11 (myosin-like protein) and R1-B6 (ELM1 homolog) did not show a silencing phenotype. R2-A11 (RPL19), and R2-B11 (RPL36A) VIGS plants were all severely stunted and developed bleaching and spontaneous necrosis phenotypes. TRV:PDS (phytoene desaturase) and the empty vector TRV:EV were used as controls. **B.** Rx1-mediated HR phenotypes in the presence of the elicitor (CP106) or GFP (control) upon TRV silencing of the interactors R0-R11, R1-B6, R2-A11, R2-B11 in *N. benthamiana*.

Discussion

In this chapter, we have demonstrated the potential of using a tobacco T7 cDNA phage library to identify specific epitopes in a pilot screening against a monoclonal antibody that recognizes the coat protein of PVY. This resulted remarkably fast in the selection of several positive and correct clones encoding the PRIKAI motif (Keller et al., 2005). If no epitope was known for this antibody, an approximate epitope of 37 amino acids would have been found by overlaying the identified binders. This strategy could be a valuable method for the identification of epitopes from cDNA libraries in for example the field of antibody-based therapies or in case of the auto-antibodies involved in autoimmune diseases. Proteomic approaches based on the affinity purification of antigens and subsequent identification by mass-spectroscopy might be more elaborate than the rapid screening method used in this study.

Unfortunately, T7 cDNA phage display did not prove to be a suitable method for the selection of protein-protein interactions involved in signalling. None of the known interactors of the Rx1 CC-NB-ARC (RanGAP2, CTR1 (Kang et al., 2008), and the Rx1 LRR) were detected in the screening of 1×10^7 cDNA fragments. Interactions with the binding matrices (or fusion proteins) were difficult to circumvent and made it necessary to include time-consuming screenings of monoclonal phages. Furthermore, the Rx1 CC-NB-ARC interacting protein fragments identified by the T7 phage display screening seem unrelated except for their strong basic nature, and several of these are involved in a housekeeping function of protein synthesis. We observed no effect on Rx1 cell death signalling other than the effect expected when disturbing the protein translation machinery. The high content of positively charged amino acid residues in the sequence of the protein fragments is shared by the peptides formed by out-of-frame translations of cDNA fragments. Together these observations make us assume that the detected interactors bind Rx1 not because they are components of the Rx1 signalling pathway, but because the Rx1 CC-NB-ARC has a strong tendency to interact with positively charged peptides.

The Rx1 CC-NB-ARC domain is acidic in nature as it contains 75 acidic residues (Asp, Glu) versus 46 basic residues (Lys, Arg), and it has a hypothetical isoelectric point of 4.84. Especially in the ARC2 domain a high number of acidic residues is concentrated in a small area in between the RNBS-D and MHD motifs.

In contrast, an area on the surface of the N-terminal half of the LRR contains a cluster of basic residues that is most likely involved in the intramolecular interaction with the NB-ARC domain (Chapter 5, this thesis). This might explain the selection of basic proteins and peptides when using the acidic CC-NB-ARC as bait in such a T7 cDNA phage library screening. The involvement of the CC-NB-ARC domain of Rx1 in electrostatic interactions can have interesting biological implications, for they could mimic an existing relevant inter- or intramolecular interaction directed by similar surface charges.

However, the one known intermolecular interaction of the Rx1 CC with RanGAP2 through its N-terminal WPP domain is probably not based on such electrostatic forces (Sacco et al., 2007; Tameling and Baulcombe, 2007). The N-terminal 154 amino acid region of *Solanum tuberosum* RanGAP2 (CAL69642) has an isoelectric point of 5.36 and contain 23 strongly

basic residues (R, K) versus 26 strongly acidic residues (D, E) and does not share the amino acid composition seen for the highly basic R13 interactors (Fig. 4C & D).

Most of the proteins in plants that do have isoelectric points in the range found for the described interactors are DNA or RNA-binding proteins, like histones or ribosomal proteins. RPL19 is partly buried in the large subunit of the ribosome and part of it lies in the interface with the small subunit. It is one of around 4 ribosomal proteins interacting with the Sec61 complex in yeast, connecting the ribosome-nascent chain complexes and the ER protein-conducting channel (Beckmann et al., 2001). L19 is linked to Myc expression in *Drosophila* and overexpressed in many cancer cell lines. Both RPL19 and (Davies and Fried, 1995; Gallant, 2005) RPL36a were found to be involved in IRES (Internal ribosome entry site)-dependent translation (Cherry et al., 2005). This alternative translation mechanism is often used by RNA viruses hijacking the host cell's translational machinery (Kieft, 2008) and by the host cell itself for translation during apoptosis (Holcik and Sonenberg, 2005).

Remarkably, ribosomal proteins have been identified several times in R gene mediated cell death. Of the 79 cDNAs 22 encoded ribosomal proteins (Lu et al., 2003), including RPL36, which resulted in a suppression of Pto-mediated HR when silenced. Two encoded histones and one a sec61 gamma subunit. It is interesting to see that the silencing of several ribosomal proteins does not cause lethality, but does affect an R protein mediated cell death. RPL19 was also identified two times in a study of the CF-4/Avr4 mediated HR (Gabriels et al., 2006). In this study, 192 transcript-derived fragments whose expression responded to the Cf-4 mediated HR induction, were silenced via TRV VIGS. For only five a severe HR suppression was observed, and two out of these five encoded RPL19 (the other three were identified as LeHSP90-1, a nuclear GTPase and the CC-NB-LRR NRC1). Both HSP90 and NRC1 have been shown to be necessary for full functionality of Rx1 (Lu et al., 2003; Gabriels et al., 2007). Although an effect on Rx mediated HR was observed, we attribute this to impaired protein synthesis upon silencing of RPL19. Therefore, the role of ribosomal proteins in Rx disease signalling against PVX remains elusive.

Material and methods

Plant material, RNA isolation and T7 display libraries

N. benthamiana plants used for the PVY-containing library were infected with the PVY^C strain. After one week, the spreading of PVY throughout the infected plants was shown by a direct antibody sandwich-ELISA with the monoclonal antibodies 5C12 and Y-5 on protein extract from leaf material. Total RNA was extracted from 500 mg leaf material by Trizol extraction (Invitrogen). mRNA was extracted from total RNA with Genoprep oligo-(dT) paramagnetic beads (Qiagen). The library was constructed from cDNA fragments generated by RT PCR with random primers from the purified mRNA (OrientExpress cDNA kit, Novagen). The cDNA fragments were ligated into the T7 genome vector arms in an orientation specific way as described in the T7 system manual (Novagen). The chosen T7 genome vector (T7Select 10-3, Novagen) results in about 3 to 5 displayed fusion peptides per capsid (Rosenberg et al., 1996). The library from *N. tabacum* T14.4 (kindly provided by D. Baulcombe, Sainsbury Laboratory, UK) was constructed in a similar way, but total RNA was extracted from 6 g leaf material. After packaging the libraries in phage, they were titered to determine the number of unique clones and amplified. The average insert length was determined by PCR with primers flanking the insert site in the T7 genome, on individual phage plaques (T7Down: 5'-AAC CCC TCA AGA CCC GTT TA, T7Up: 5'-GGA GCT GTC GTA TTC CAG TC). Phages were amplified in *E. coli* strain BLT5403 as described in Chapter 2.

Plasmid construction

pBAD:R13-V5H6: The DNA encoding the CC-NB-ARC domains of Rx1 (R13) was cloned in the Arabinose inducible expression vector pBAD (invitrogen) under control of the AraBAD promoter via NcoI and EcoRI (The NcoI overlaps the Rx1 startcodon, the EcoRI site follows directly after the codons coding for M473 (GAA GCT CGA AAC ATG AAT TC C GCC). Both were introduced via PCR. The construct pET42b-R13 was constructed by inserting the R13 fragment from the pBAD vector via NcoI-EcoRI in the multiple cloning site of the Novagen pET42b vector. Expression from this vector is controlled by the T7 promoter and lac operator and results in a GST fusion to the N-terminus of R13. The TRX-fusions of the selected binders from the T7 libraries were constructed by BamHI-XhoI ligation from the T7 genome background into the Novagen pET32b multiple cloning site. Cloning into pET32b results in an N-terminal fusion to the thioredoxine tag.

Protein production and purification

Protein expression from pBAD: *E. coli* Top10 cells with the pBAD:R13-V5H6 plasmid were grown at 37°C in 50 mL LB M9 medium (standard LB medium + 0.4% glucose, 100 mM MgSO₄, 1x M9 salts) to an OD₆₀₀ of 0.5. The bacteria are then pelleted and resuspended in 50 mL of LB medium. These cells were diluted 1:100 in LB with the appropriate antibiotics and incubated at 37°C at 250 rpm in large flasks until the culture reached an OD₆₀₀ of 0.5. Expression is induced by adding arabinose to a final concentration of 0.02% and cells are incubated overnight at 20°C, shaking 250 rpm. For further processing the bacteria are pelleted and resuspended in lysis buffer (400 mM NaCl, 100 mM KCl, 10 mM imidazol, 10 mM Tris-HCl pH8), concentrating the cell density 100x. Protein expression from pET vectors: *E. coli* BL21(DE3) containing the pET plasmids were grown in LB medium as

described above. After induction of expression with 0.4 mM IPTG the bacteria were incubated for 5 hours at 20°C, pelleted and resuspended in PBS or lysis buffer. *E. coli* suspension with added protease inhibitors (Pefabloc, Roche) were lysed using a French pressure cell at 100 MPa. All constructs were purified using the Amersham HiTrap HP Ni-NTA column with the Amersham Akta Prime. Protein was eluted in fractions by applying a gradient towards higher imidazol concentration with pure elution buffer (400 mM NaCl, 100 mM KCl, 10 mM Tris-HCl- pH 8, 500 mM Imidazol) as highest imidazol concentration. Fractions containing the highest protein concentration were dialyzed against the lysis buffer without imidazol.

T7 library screening

Bio-panning was performed as described in the results section. Phages were incubated with the bait protein, either coated in microtiter plates (Nunc Maxisorp) or bound to magnetic Ni-NTA agarose beads (Qiagen). Microtiter wells and beads were blocked with BSA before incubation with the phages. After 2-4 hours incubation the non-binding phages were removed by three washes with PBS-T (PBS + 0.1% Tween-20). Binding phages were eluted with 50 µl elution buffer (1% SDS, 0.1% Tween-20) and added to *E. coli* BLT5403 for amplification (as described in the T& manual, Novagen). After every elution a small fraction of the eluate was used to determine the number of bound phage by titering. Plaque lift assays were performed by mixing a dilution of phages with melted Topagar (LB + 0.7% agar) and BLT5403. The Topagar-phage-bacteria mixture was poured on LB agar plates and incubated until clear plaques in the bacterial lawn became visible. The agar was overlaid with a nitrocellulose membrane. Phage bound adsorbed to this membrane were visualized by antibodies.

GST pull-down assay

50 µl slurry of Glutathione sepharose 4B beads (GE Healthcare) was used per sample in the GST pull-down. Before incubation with the protein extracts the beads were washed three times in PBS. Purified GST or GST-R13 were mixed with raw protein extracts from *E. coli* expressing TRX-fusions of the protein fragments identified earlier as positive in the library screening. These mixtures were incubated with the beads at 4°C overnight and washed three times with PBS-T. GST and GST-R13 were eluted by incubation with 200 µl 10 mM glutathione at room temperature for 5 minutes. Eluates were further analyzed by SDS-PAGE and immunoblotting using monoclonal anti-TRX (Invitrogen) or anti-GST antibodies.

Virus induced gene-silencing

Three week old *N. benthamiana* plants were co-infiltrated with *A. tumefaciens* GV3101 containing pTRV1 and pTRV2 (Liu et al., 2002). The cDNA fragments from the T7 library were cloned via BamHI-XhoI into the multiple cloning site of pTRV2. The empty TRV vector was used as negative control and TRV2:PDS (phytoene desaturase, (Liu et al., 2002)) was used to visualize the silencing progression. 21 days after inoculation with the silencing vectors the plants were tested for their ability to initiate an HR response via the agroinfiltration of pBINPLUS:Rx1 combined with pBINPIUS:CP106 or CP105.

References

- Ade, J., Deyoung, B.J., Golstein, C. and Innes, R.W.** (2007). Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease. *Proc Natl Acad Sci U S A* **104**, 2531-2536.
- Aoki, S., Ohta, K., Yamazaki, T., Sugawara, F. and Sakaguchi, K.** (2005). Mammalian mitotic centromere-associated kinesin (MCAK) - A new molecular target of sulfoquinovosylacylglycerols novel antitumor and immunosuppressive agents. *Febs Journal* **272**, 2132-2140.
- Beckmann, R., Spahn, C.M., Eswar, N., Helmers, J., Penczek, P.A., Sali, A., Frank, J. and Blobel, G.** (2001). Architecture of the protein-conducting channel associated with the translating 80S ribosome. *Cell* **107**, 361-372.
- Bendahmane, A., Kanyuka, K. and Baulcombe, D.C.** (1999). The Rx gene from potato controls separate virus resistance and cell death responses. *Plant Cell* **11**, 781-791.
- Bendahmane, A., Querci, M., Kanyuka, K. and Baulcombe, D.C.** (2000). Agrobacterium transient expression system as a tool for the isolation of disease resistance genes: application to the Rx2 locus in potato. *Plant J.* **21**, 73-81.
- Bieri, S., Mauch, S., Shen, Q.H., Peart, J., Devoto, A., Casais, C., Ceron, F., Schulze, S., Steinbiss, H.H., Shirasu, K. and Schulze-Lefert, P.** (2004). RAR1 positional controls steady state levels of barley MLA resistance proteins and enables sufficient MLA6 accumulation for effective resistance. *Plant Cell* **16**, 3480-3495.
- Boonekamp, P.M., Pomp, H., Gussenhoven, G.C. and Schots, A.** (1991). The Use of Immunochemical Techniques and Monoclonal-Antibodies to Study the Viral Coat Protein-Structure of Potato Virus-a, Potato Virus-Y and Beet Necrotic Yellow Vein Virus. *Acta Botanica Neerlandica* **40**, 41-52.
- Bukanov, N.O., Meek, A.L., Klinger, K.W., Landes, G.M. and Ibraghimov-Beskrovnya, O.** (2000). A modified two-step phage display selection for isolation of polycystin-1 ligands. *Functional Integrative Genomics* **1**, 193-199.
- Caplan, J.L., Mamillapalli, P., Burch-Smith, T.M., Czymbek, K. and Dinesh-Kumar, S.P.** (2008). Chloroplastic Protein NR1P1 Mediates Innate Immune Receptor Recognition of a Viral Effector. *Cell* **132**, 449-462.
- Castagnoli, L., Zucchini, A., Quondam, M., Rossi, M., Vaccaro, P., Panni, S., Paoluzi, S., Santonico, E., Dente, L. and Cesareni, G.** (2001). Alternative bacteriophage display systems. *Comb. Chem. High Throughput Screen* **4**, 121-133.
- Cherry, S., Doukas, T., Armknecht, S., Whelan, S., Wang, H., Sarnow, P. and Perrimon, N.** (2005). Genome-wide RNAi screen reveals a specific sensitivity of IRES-containing RNA viruses to host translation inhibition. *Genes Dev.* **19**, 445-452.
- Danner, S. and Belasco, J.G.** (2001). T7 phage display: A novel genetic selection system for cloning RNA-binding proteins from cDNA libraries. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 12954-12959.
- Davies, B. and Fried, M.** (1995). The L19 ribosomal protein gene (RPL19): gene organization, chromosomal mapping, and novel promoter region. *Genomics* **25**, 372-380.
- de La Fuente van Bentem, S., Vossen, J.H., de Vries, K.J., van Wees, S., Tameling, W.I., Dekker, H.L., de Koster, C.G., Haring, M.A., Takken, F.L. and Cornelissen, B.J.** (2005). Heat shock protein 90 and its co-chaperone protein phosphatase 5 interact with distinct regions of the tomato I-2 disease resistance protein. *Plant J* **43**, 284-298.
- Deslandes, L., Olivier, J., Peeters, N., Feng, D.X., Khounlotham, M., Boucher, C., Somssich, I., Genin, S. and Marco, Y.** (2003). Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. *Proc Natl Acad Sci U S A* **100**, 8024-8029.
- Dodds, P.N., Lawrence, G.J., Catanzariti, A.M., Teh, T., Wang, C.I., Ayliffe, M.A., Kobe, B. and Ellis, J.G.** (2006). Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proc Natl Acad Sci U S A* **103**, 8888-8893.
- Gabriels, S.H., Takken, F.L., Vossen, J.H., de Jong, C.F., Liu, Q., Turk, S.C., Wachowski, L.K., Peters, J., Witsenboer, H.M., de Wit, P.J. and Joosten, M.H.** (2006). CDNA-AFLP combined with functional analysis reveals novel genes involved in the hypersensitive response. *Mol Plant Microbe Interact* **19**, 567-576.
- Gabriels, S.H., Vossen, J.H., Ekengren, S.K., Ooijen, G., Abd-El-Halim, A.M., Berg, G.C., Rainey, D.Y., Martin, G.B., Takken, F.L., Wit, P.J. and Joosten, M.H.** (2007). An NB-LRR protein required for HR signalling mediated by both extra- and intracellular resistance proteins. *Plant J* **50**, 14-28.
- Gallant, P.** (2005). Myc, Cell Competition, and Compensatory Proliferation. *Cancer Res* **65**, 6485-6487.
- Hewezi, T., Mouzeyar, S., Thion, L., Rickauer, M., Alibert, G., Nicolas, P. and Kallerhoff, J.** (2006). Antisense expression of a NBS-LRR sequence in sunflower (*Helianthus annuus* L.) and tobacco (*Nicotiana tabacum* L.): evidence for a dual role in plant development and fungal resistance. *Transgenic Res* **15**, 165-180.
- Holcik, M. and Sonenberg, N.** (2005). Translational control in stress and apoptosis. *Nat Rev Mol Cell Biol* **6**, 318-327.
- Houshmand, H. and Bergqvist, A.** (2003). Interaction of hepatitis C virus NS5A with La protein revealed by T7 phage display. *Biochem Biophys Res Commun* **309**, 695-701.
- Jia, Y., McAdams, S.A., Bryan, G.T., Hershey, H.P. and Valent, B.** (2000). Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *Embo J.* **19**, 4004-4014.
- Jin, Y., Yu, J. and Yu, Y.G.** (2002). Identification of hNopp140 as a binding partner for doxorubicin with a phage display cloning method. *Chem. Biol.* **9**, 157-162.

- Kang, H.-G., Kuhl, J.C., Kachroo, P. and Klessig, D.F.** (2008). CRT1, an Arabidopsis ATPase that Interacts with Diverse Resistance Proteins and Modulates Disease Resistance to Turnip Crinkle Virus. *Cell Host & Microbe* **3**, 48-57.
- Keller, H., Pomp, R., Bakker, J. and Schots, A.** (2005). Epitope identification and in silico prediction of the specificity of antibodies binding to the coat proteins of Potato Virus Y strains. *European Journal of Plant Pathology* **111**, 391-397.
- Kieft, J.S.** (2008). Viral IRES RNA structures and ribosome interactions. *Trends Biochem Sci* **33**, 274-283.
- Kohler, A., Rinaldi, C., Duplessis, S., Baucher, M., Geelen, D., Duchaussoy, F., Meyers, B.C., Boerjan, W. and Martin, F.** (2008). Genome-wide identification of NBS resistance genes in *Populus trichocarpa*. *Plant Mol Biol* **66**, 619-636.
- Krumpe, L.R., Atkinson, A.J., Smythers, G.W., Kandel, A., Schumacher, K.M., McMahon, J.B., Makowski, L. and Mori, T.** (2006). T7 lytic phage-displayed peptide libraries exhibit less sequence bias than M13 filamentous phage-displayed peptide libraries. *Proteomics* **6**, 4210-4222.
- Leister, R.T., Dahlbeck, D., Day, B., Li, Y., Chesnokova, O. and Staskawicz, B.J.** (2005). Molecular genetic evidence for the role of SGT1 in the intramolecular complementation of Bs2 protein activity in *Nicotiana benthamiana*. *Plant Cell* **17**, 1268-1278.
- Liu, Y., Schiff, M. and Dinesh-Kumar, S.P.** (2002). Virus-induced gene silencing in tomato. *Plant J* **31**, 777-786.
- Lu, R., Malcuit, I., Moffett, P., Ruiz, M.T., Peart, J., Wu, A.J., Rathjen, J.P., Bendahmane, A., Day, L. and Baulcombe, D.C.** (2003). High throughput virus-induced gene silencing implicates heat shock protein 90 in plant disease resistance. *Embo J* **22**, 5690-5699.
- Lukasik, E. and Takken, F.L.** (2009). STANDING strong, resistance proteins instigators of plant defence. *Curr Opin Plant Biol* **24**, 24.
- Mackey, D., Belkhadir, Y., Alonso, J.M., Ecker, J.R. and Dangl, J.L.** (2003). Arabidopsis RIN4 is a target of the type III virulence effector AvrRpt2 and modulates RPS2-mediated resistance. *Cell* **112**, 379-389.
- Mackey, D., Holt, B.F., 3rd, Wiig, A. and Dangl, J.L.** (2002). RIN4 interacts with *Pseudomonas syringae* type III effector molecules and is required for RPM1-mediated resistance in Arabidopsis. *Cell* **108**, 743-754.
- McHale, L., Tan, X., Koehl, P. and Michelmore, R.W.** (2006). Plant NBS-LRR proteins: adaptable guards. *Genome Biol* **7**, 212.
- Meyers, B.C., Kozik, A., Griego, A., Kuang, H. and Michelmore, R.W.** (2003). Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. *Plant Cell* **15**, 809-834.
- Moffett, P., Farnham, G., Peart, J. and Baulcombe, D.C.** (2002). Interaction between domains of a plant NBS-LRR protein in disease resistance-related cell death. *Embo J* **21**, 4511-4519.
- Peart, J.R., Mestre, P., Lu, R., Malcuit, I. and Baulcombe, D.C.** (2005). NRG1, a CC-NB-LRR protein, together with N, a TIR-NB-LRR protein, mediates resistance against tobacco mosaic virus. *Curr. Biol.* **15**, 968-973.
- Rairdan, G.J., Collier, S.M., Sacco, M.A., Baldwin, T.T., Boettrich, T. and Moffett, P.** (2008). The Coiled-Coil and Nucleotide Binding Domains of the Potato Rx Disease Resistance Protein Function in Pathogen Recognition and Signaling. *Plant Cell* **20**, 739-751.
- Rairdan, G.J. and Moffett, P.** (2006). Distinct Domains in the ARC Region of the Potato Resistance Protein Rx Mediate LRR Binding and Inhibition of Activation. *Plant Cell* **18**, 2082-2093.
- Robaglia, C., Durand-Tardif, M., Tronchet, M., Boudazin, G., Astier-Manificier, S. and Casse-Delbart, F.** (1989). Nucleotide sequence of potato virus Y (N Strain) genomic RNA. *J Gen Virol* **70** (Pt 4), 935-947.
- Rosenberg, A., Griffin, G., Studier, F.W., McCormick, M., Berg, J. and Mierendorf, R.** (1996). T7Select Phage Display System: A powerful new protein display system based on bacteriophage T7. *in* *Innovations* **6**, 1-6.
- Sacco, M.A., Mansoor, S. and Moffett, P.** (2007). A RanGAP protein physically interacts with the NB-LRR protein Rx, and is required for Rx-mediated viral resistance. *Plant J* **52**, 82-93.
- Savinov, S.N. and Austin, D.J.** (2001). The cloning of human genes using cDNA phage display and small-molecule chemical probes. *Comb. Chem. High Throughput Screen* **4**, 593-597.
- Shen, Q.H., Saijo, Y., Mauch, S., Biskup, C., Bieri, S., Keller, B., Seki, H., Ulker, B., Somssich, I.E. and Schulze-Lefert, P.** (2007). Nuclear activity of MLA immune receptors links isolate-specific and basal disease-resistance responses. *Science* **315**, 1098-1103.
- Smith, G.P.** (1985). Filamentous Fusion Phage - Novel Expression Vectors That Display Cloned Antigens on the Virion Surface. *Science* **228**, 1315-1317.
- Smith, G.P. and Petrenko, V.A.** (1997). Phage Display. *Chem Rev* **97**, 391-410.
- Takken, F.L.W. and Tameling, W.I.L.** (2009). To Nibble at Plant Resistance Proteins. *Science* **324**, 744-746.
- Tameling, W.I. and Baulcombe, D.C.** (2007). Physical association of the NB-LRR resistance protein Rx with a Ran GTPase-activating protein is required for extreme resistance to Potato virus X. *Plant Cell* **19**, 1682-1694.
- The Arabidopsis Genome Initiative** (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796-815.
- Tuskan, G.A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhallerao, R.R., Bhallerao, R.P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G.L., Cooper, D., Coutinho, P.M., Couturier, J., Covert, S., Cronk, Q.,**

- Cunningham, R., Davis, J., Degroeve, S., Dejardin, A., Depamphilis, C., Detter, J., Dirks, B., Dubchak, I., Duplessis, S., Ehlting, J., Ellis, B., Gendler, K., Goodstein, D., Gribskov, M., Grimwood, J., Groover, A., Gunter, L., Hamberger, B., Heinze, B., Helariutta, Y., Henrissat, B., Holligan, D., Holt, R., Huang, W., Islam-Faridi, N., Jones, S., Jones-Rhoades, M., Jorgensen, R., Joshi, C., Kangasjarvi, J., Karlsson, J., Kelleher, C., Kirkpatrick, R., Kirst, M., Kohler, A., Kalluri, U., Larimer, F., Leebens-Mack, J., Leple, J.C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D.R., Nelson, C., Nieminen, K., Nilsson, O., Pereda, V., Peter, G., Philippe, R., Pilate, G., Poliakov, A., Razumovskaya, J., Richardson, P., Rinaldi, C., Ritland, K., Rouze, P., Ryaboy, D., Schmutz, J., Schrader, J., Segerman, B., Shin, H., Siddiqui, A., Sterky, F., Terry, A., Tsai, C.J., Uberbacher, E., Unneberg, P., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**, 1596-1604.
- van der Biezen, E.A. and Jones, J.D.G. (1998). Plant disease-resistance proteins and the gene-for-gene concept. *Trends Biochem.Sci.* **23**, 454-456.
- van der Hoorn, R.A. and Kamoun, S. (2008). From Guard to Decoy: a new model for perception of plant pathogen effectors. *Plant Cell* **20**, 2009-2017.
- van Ooijen, G., van den Burg, H.A., Cornelissen, B.J. and Takken, F.L. (2007). Structure and function of resistance proteins in solanaceous plants. *Annu Rev Phytopathol* **45**, 43-72.
- Yamamoto, M., Kominato, Y. and Yamamoto, F. (1999). Phage display cDNA cloning of protein with carbohydrate affinity. *Biochem. Biophys. Res. Commun.* **255**, 194-199.
- Zhou, T., Wang, Y., Chen, J.Q., Araki, H., Jing, Z., Jiang, K., Shen, J. and Tian, D. (2004). Genome-wide identification of NBS genes in japonica rice reveals significant expansion of divergent non-TIR NBS-LRR genes. *Mol Genet Genomics* **271**, 402-415.

Chapter 4

Domain exchange between Rx1 and Gpa2 in potato reveals flexibility of CC-NB-LRR genes to switch between virus and nematode resistance

Erik Slootweg¹, Kamila Koropacka², Jan Roosien², Robert Dees¹, Casper van Schaik², Rikus Pomp², Liesbeth Bouwman², Arjen Schots¹, Geert Smant², Jaap Bakker², Aska Goverse²

¹ Lab of Molecular Recognition and Antigen Technology and ² Lab of Nematology, Dep. of Plant Sciences, Wageningen University, Binnenhaven 5, 6709 PD, Wageningen, the Netherlands.

To be submitted

Abstract

Plants are challenged by a myriad of pathogens and to defend themselves they have evolved numerous disease-resistance (*R*) genes, of which most encode a nucleotide binding domain (NB) and a specificity determining leucine-rich repeat region (LRR). NB-LRR proteins are often rapidly evolving molecules and it has been shown that only few changes in the LRR domain are required to alter the resistance specificity towards novel variants of a pathogen. However, little is known about the ability of NB-LRR genes to generate resistance to phylogenetically unrelated pathogens. Here we exchanged the LRR domains of the paralogs *Gpa2* and *Rx1*, which mediate resistance to the cyst nematode *Globodera pallida* and *Potato virus X* (PVX), respectively, in potato (*Solanum tuberosum*). Both *R* genes have a predicted coiled-coil domain (CC) at their amino terminus. The genetic fusion of the CC-NB of *Gpa2* with the LRR of *Rx1* (*Gpa2_{CC}/Rx1_L*) showed autoactivation, but lowering the expression levels resulted in extreme resistance to PVX as observed in wild type potato plants. In contrast, transgenic potato expressing the reciprocal construct (*Rx1_{CC}/Gpa2_L*) showed a loss-of-function phenotype. Reintroduction of the first 5 LRRs of *Rx1* resulted in a gain of resistance, and a mild inhibition of nematode development was obtained similar to wild type resistance to *G. pallida*. Our results show that the CC, NB, and ARC domains are non-pathogen specific modules and support the hypothesis that changing the recognition specificities of LRR domains is sufficient to switch the resistance specificities of NB-LRR genes towards taxonomically unrelated pathogens, irrespective the route of invasion or mode of parasitism.

Introduction

Plants are constantly exposed to a diverse array of pathogens and parasites that attempt to invade leaves, stems or roots by various mechanisms. To sense foreign invaders, plants have evolved a sophisticated immune system consisting of receptor-like resistance (*R*) proteins and a more generic microbe-associated molecular pattern (MAMP) recognition system (Holt et al., 2003; Jones and Takemoto, 2004). The dominant *R* genes operate in a gene-for-gene system in which the *R* proteins limit the growth of viruses, bacteria, fungi and invertebrate pests by triggering a host defence response upon recognition of pathogen-derived elicitors. This recognition may involve a direct interaction between the *R* protein and its cognate elicitor, or an indirect interaction by sensing elicitor-dependent modifications of host proteins. The subsequent host defence response may include the production of anti-pathogenic compounds, the induction of a reactive oxygen burst and a local programmed cell death, or a so-called hypersensitive response (HR) (Lam et al., 2001). Most known *R* genes encode a nucleotide binding site (NB) and a leucine-rich repeat domain responsible for the direct or indirect recognition of the pathogen. Within the NB-LRR class of disease resistance genes two large families can be distinguished: CC-NB-LRR proteins that have a N-terminal coiled-coil (CC) domain and TIR-NB-LRR proteins with a N-terminal Toll/Interleukin-1 receptor (TIR) domain (Meyers et al., 2003).

Plant resistance genes of the NB-LRR class are highly polymorphic and are among the most rapidly evolving genes in the genome (Mondragon-Palomino et al., 2002; Cork and Purugganan, 2005). Although mutations are a major source of variation, much of the diversity within resistance gene families arises from intra- and intergenic sequence

exchanges that shuffle polymorphic sites between individual genes. While several simple *R* gene loci do exist, most *R* genes belong to gene families located at complex loci harbouring several tandemly repeated NB-LRR homologs. The occurrence of *R* gene homologs in clusters is thought to promote sequence exchange by gene conversion and unequal crossing-over. A detailed study of the Rp1 rust resistance complex of maize showed that reshuffling of sequences played a central role in the creation of genetic diversity and even lead to new specificities (Hulbert, 1997; Sun et al., 2001; Smith and Hulbert, 2005). Extensive work with the L, M, N, and P loci in flax demonstrated the role of recombination in the evolution of new recognition specificities to *Melampsora lini* strains (Ellis et al., 1999; Luck et al., 2000; Dodds et al., 2001b; Dodds et al., 2001a).

A key issue in plant pathology is the capacity of plants to generate novel resistance specificities. It has clearly been shown that the LRR domain plays a crucial role in recognizing foreign invaders and only few amino acid changes may alter the recognition specificity towards different variants of a pathogen (Farnham and Baulcombe, 2006; Takken et al., 2006). At complex loci, but also at simple loci, highly similar *R* genes have been found that recognize series of variants of a single pathogen (Hayes et al., 2004). However, the ability of NB-LRR genes to switch resistance specificities between taxonomically unrelated pathogens is largely unknown. Molecular studies to address this issue are hampered, because the vast majority of the *R* gene specificities have not been identified yet. Of all presently known NB-LRR sequences, 149 in the *Arabidopsis* genome, about 400 in poplar and over 500 in rice, relatively few can be coupled to a cognate pathogen (The Arabidopsis Genome Initiative, 2000; Meyers et al., 2003; Zhou et al., 2004; Tuskan et al., 2006). Even in clusters containing *R* genes of known specificities, the functions of adjacent paralogs are often unknown. An indication that altering resistance specificities towards widely different pathogens involves relatively few molecular changes, comes from the observation that resistance genes for downy mildew (RPP8) and for *Turnip Crinkle virus* (HRT) in *Arabidopsis* are highly homologous and are found at the same genomic position in different accessions (Cooley et al., 2000). Also random *in vitro* mutagenesis in the LRR domain of the *Rx1* gene conferring resistance to *Potato virus X* showed that extending the recognition spectrum to poplar mosaic virus, required only single amino acid changes in the LRR to recognize the related coat protein (Farnham and Baulcombe, 2006).

The *R* genes of the CC-NB-LRR and TIR-NB-LRR families have a well defined modular structure and confer disease resistance through a multistage activation process initiated by the LRR domain in the presence of the elicitor (Takken et al., 2006). Activation of the N-terminal domains leads to the transduction of a yet unknown signal that initiates the defence response. Hence, the flexibility of NB-LRR genes to generate resistance specificities to phylogenetically unrelated pathogens will not only depend on the ability to develop novel recognition specificities by the LRR domain, but also on the ability of the CC, TIR and NB domains to transduce signals that arrest the development of entirely different pathogens. The R proteins Gpa2 and Rx1 are highly homologous and located in the same *R* gene cluster of potato, *Solanum tuberosum*, but confer resistance to two different types of pathogen, the potato cyst nematode *Globodera pallida* and to *Potato virus X* (PVX), respectively (van der Vossen et al., 2000). Potato cyst nematodes penetrate the vascular tissue of the roots and fuse plant cells into multinucleate feeding cells. In resistant *Gpa2*

plants the syncytium is surrounded by necrotic cells and the reduced flow of nutrients delays the growth, and finally blocks the development of fertile adult females. PVX, however, is a single stranded RNA virus that is transmitted above ground by insects and other forms of mechanical injury, resulting in systemic infection of the aerial parts of the plant. A striking feature of Rx1 mediated resistance is the rapid arrest of PVX accumulation in the initial infected cells, resulting in symptomless resistance, so-called extreme resistance. Gpa2 and Rx1 therefore provide an excellent test system to investigate the exchangeability of recognition and signalling domains and explore the evolutionary flexibility of R proteins.

Here, we provide evidence for the hypothesis, that, via intergenic sequence exchanges and various types of mutations, NB-LRR proteins have the potential to alter resistance specificities towards taxonomically unrelated pathogens in relatively short evolutionary time periods. Both the regulatory sequences and CC-NB domains of the paralogs Gpa2 and Rx1 are non-pathogen specific and exchangeable. Remarkably, the genetic fusions of the CC-NB of *Rx1* with the LRR of *Gpa2* (*Rx1_{CN}/Gpa2_L*) and the reciprocal domain swap (*Gpa2_{CN}/Rx1_L*) were not functional when driven by the endogenous promoters or 35S promoter. Gain of wild type resistance was obtained by re-introducing the first five LRRs of Rx1 in *Rx1_{CN}/Gpa2*, restoring the compatibility between the N-terminal part of the LRR and the ARC2 domain. Decreasing the expression levels for *Gpa2_{CN}/Rx1_L* resulted in extreme resistance against PVX, indistinguishable from wild type plants. Our results indicate that not only coding sequences, but that also optimizing the expression levels may play a role in generating novel resistances.

RESULTS

The CC-NB domain of Gpa2 signals extreme resistance to PVX

To test the versatility of the various domains of Gpa2 and Rx1 in triggering defence responses to PVX and potato cyst nematodes, a chimeric gene encoding the CC-NB-ARC domain of *Gpa2* and the LRR domain of *Rx1* was created (Fig. 1A). However, under control of the double enhanced CaMV 35S promoter this construct results in a constitutive cell death response in an agroinfiltration assay on leaves of *Nicotiana benthamiana* (Rairdan and Moffett, 2006). To see if attenuating the expression level would lessen the autoactive response we introduced an out of frame start codon upstream of the original translation initiation site (Fig. 1A). Hence, translation of the *R* gene reading frame becomes dependent on leaky scanning (LS) by the ribosome (Kozak, 1995; Kozak, 1999), resulting in a strong reduction of the level of correctly translated protein. The effect of the leaky scanning (*35S_{LS}*) promoter was evaluated by the expression of GFP construct under control of the 35S promoter and the *35S_{LS}* promoter and comparing the protein levels by Western blots. As expected the *35S_{LS}* promoter showed a strong reduction of the expression of GFP (Fig. 1B). Expression of the recombinant protein *Gpa2_{CN}/Rx1_L* in an agroinfiltration assay on leaves of *N. benthamiana* under control of the *35S_{LS}* promoter showed that the protein levels were now below the autoactivation threshold (Fig. 1C). The construct, however, is at these protein levels able to induce a specific HR in the presence of the avirulent coat protein of PVX (Fig. 1D).

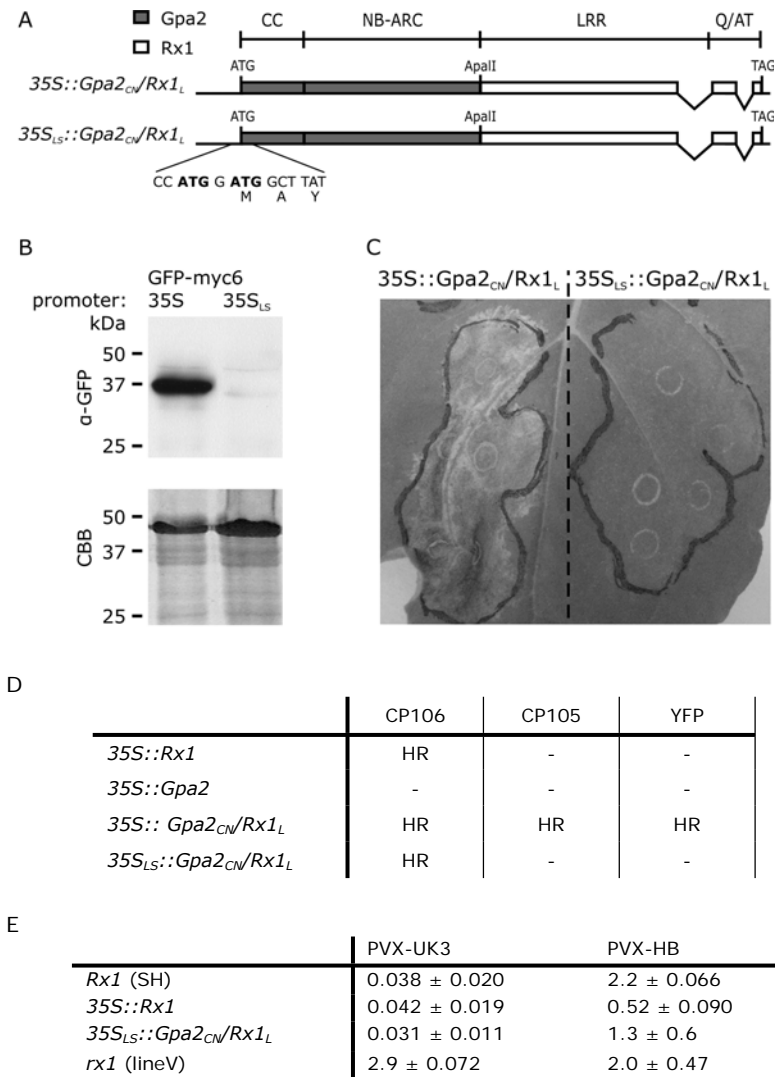


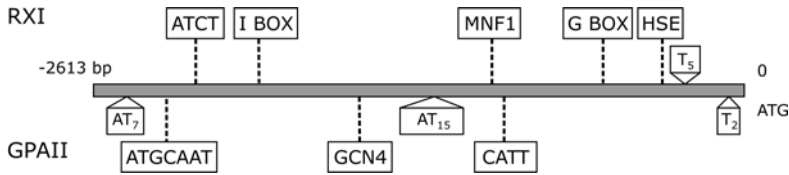
Figure 1. A. The reciprocal domain swap construct *Gpa2_{CN}/Rx1_L* was obtained by exchanging the LRR domain of *Gpa2* with the corresponding domain of *Rx1* using the *ApaI* restriction site in the context of a CaMV 35S promoter cassette for expression in plants. A second translation initiation site was introduced in p35S_{LS}::*Gpa2_{CN}/Rx1_L* to obtain leaky scanning of ribosomes (Kozak, 1995; Kozak, 1999) and a subsequent reduction of the expression levels of the protein. **B.** Comparison of the expression levels of the green fluorescent protein GFP-myc6 under control of the CaMV 35S and the leaky scanning 35S_{LS} promoter in an agroinfiltration assay. Leaf protein extracts were separated on SDS-PAGE followed by western blotting and detection of the protein by a polyclonal anti-GFP peroxidase-conjugated antibody (α-GFP) or Coomassie Brilliant Blue staining (CBB). The GFP specific band is indicated by an arrow. **C.** Agroinfiltration of *Nicotiana benthamiana* leaves with p35S::*Gpa2_{CN}/Rx1_L* results in a constitutive cell death response in the absence of the PVX elicitor, whereas no such autoactivation response was observed for 35S_{LS}::*Gpa2_{CN}/Rx1_L*. **D.** Agroinfiltration of *Nicotiana benthamiana* leaves with 35S::*Gpa2_{CN}/Rx1_L* and 35S_{LS}::*Gpa2_{CN}/Rx1_L* in the presence and absence of the virulent and avirulent PVX elicitor CP105 and CP106, respectively. Expression of the wild type *R* genes *Rx1* and *Gpa2* under control of the normal CaMV 35S promoter were included as a positive and negative control. HR = hypersensitive response. **E.** A greenhouse virus resistance assay was performed on transgenic potato plants expressing the wild type *Rx1* gene under control of the CaMV 35S promoter and the domain swap construct *Gpa2_{CN}/Rx1_L* under control of the leaky scanning 35S_{LS} promoter. The diploid potato clone SH, which contains the endogenous *Rx1* gene, was used as the resistant control and the diploid potato clone lineV, which was used for the transformation of the constructs, was used as susceptible control. Leaf material was collected from secondary infected leaves of the plant apex three weeks after infection with the avirulent strain PVX_{UK3} or the virulent strain PVX_{HB} and systemic spreading of the virus in the plants was detected by ELISA.

Transgenic potato plants harbouring the 35S_{LS}::*Gpa2_{CN}/Rx1_L* were created to test for PVX resistance. The potato clone SH containing the endogenous *Rx1* gene and a transgenic line containing the *Rx1* gene under control of the 35S promoter were included as resistant controls. Plants were inoculated with either the avirulent strain PVX_{UK3} or the virulent strain PVX_{HB} and three weeks after inoculation the compound leaves near the shoot apex were harvested for virus detection using ELISA. Figure 1E shows that no detectable amounts of the avirulent PVX strain could be observed in the transgenic plants expressing the recombinant gene 35S_{LS}::*Gpa2_{CN}/Rx1_L*, as was the case for the resistant control plants. In the susceptible control plants, however, large amounts of PVX_{UK3} could be detected, indicating that the inoculation with the avirulent strain was successful. Infection of the plants with the *Rx1*-resistance breaking strain resulted in systemic spreading of the virus in all plants, although reduced in the plants expressing *Rx1* from the 35S promoter. These results show that *Gpa2_{CN}/Rx1_L* confers extreme resistance to PVX in shoots of potato in a gene-for-gene specific manner like the original *Rx1* gene. These data support earlier findings that the recognition specificity of *Rx1* is determined by the LRR domain (Rairdan and Moffett, 2006), but more interestingly, that the CC-NB-ARC domain of the nematode resistance gene *Gpa2* is able to activate an extreme resistance response against potato virus X.

Both the endogenous promoters of Rx1 and Gpa2 are able to drive virus and nematode resistance

Comparison of the flanking sequences of *Gpa2* and *Rx1* showed that the sequences upstream of the start codon share more similarity than the sequences downstream of the ORF. Analyses of an approximately 2.6 kb DNA fragment upstream of the start codon revealed two extra TA-rich regions in the *Gpa2* sequence at -2458bp (TA₇) and at -1329 (TA₁₅), which are predicted to function as an enhancer, and two small indels just upstream of the start codons (Fig. 2A). The remaining sequences show a similarity of about 97% and only differ in a number of single base pair substitutions. For this region, a number of *cis* acting regulatory elements (CARE) was predicted including an AT-rich element (binding site for AT rich DNA binding protein ATBP-1) and AT-rich sequence (for maximal elicitor-mediated activation), an ethylene (ERE), auxin (TGA) and wound (WUN) responsive element. Analysing the genomic sequence +298 bp downstream of the stop codon revealed that the 3'UTR regions of *Rx1* and *Gpa2* were identical until +160 followed by a more variable region containing 8 single base pair substitutions and two small indels of 1 and 2 nucleotides in case of *Gpa2* and one indel of 3 nucleotides for *Rx1* (Fig. 2B).

A



B

Cis acting regulatory elements	<i>Rx1</i>	<i>Gpa2</i>	Function
AT-rich element	- 1915	- 2458	Binding site for AT rich DNA binding protein ATBP-1
AT-rich sequence	- 1354	- 1329	Element for maximal elicitor-mediated activation
ERE	- 1240	- 1235	Ethylene-responsive element
TC-rich repeats	- 85	- 85	Involved in stress and defence responsiveness
TGA element	- 1707	- 1833	Auxin-responsive element
WUN motif	- 529	- 624	Wound responsive element

Figure 2. A. Schematic representation of the Rx1 and Gpa2 promoter region, which are highly homologous in a ~2600 bp region upstream of the ATG start codon (97.3 % identity). Two extra TA-rich regions are present in the Gpa2 promoter (AT₇ at - 2458 bp and AT₁₅ at -1329 bp) and two small indels (T₅ and T₂) are located just upstream of the start codon (-207 bp and -6 bp, respectively). Various single base pair substitutions are distributed over the promoter region, resulting in the prediction of several additional *cis* acting regulatory elements (PlantCARE) for either the Rx1 or Gpa2 promoter. Most elements have a function in light responsiveness, except for HSE, and are therefore most likely not directly involved in the regulation of Gpa2 and Rx mediated resistance. **B.** PlantCARE prediction of several *cis* acting regulatory elements involved in plant defence and stress in the *Rx* and *Gpa2* promoter regions (-2573 bp and -2613 bp, respectively).

The homologous DNA fragment upstream of the start codon of *Rx1* (2571 bp) and *Gpa2* (2613 bp) were tested for promoter activity in an agroinfiltration assay in leaves of *N. benthamiana*. Expression of *Rx1* under control of either the *Rx1* promoter region (*pRXI*) or the *Gpa2* promoter region (*pGPAII*) in combination with the corresponding terminator sequences (298 bp) resulted in the detection of an HR within 2 dpi in the presence of the avirulent elicitor CP106. A similar response was observed for the original genomic BAC clone harbouring *Rx1* (Fig. 3A). Subsequently transgenic potato plants harbouring *pGPAII::Rx1* and *pRXI::Rx1* were tested in a virus resistance assay. No systemic spreading of the avirulent strain PVX_{UK3} was detected, whereas an accumulation of the virulent strain PVX_{HB} was observed using ELISA (Fig. 3B). From this experiment, it was concluded that the original *Rx1* promoter activity was retained in the selected DNA fragment and that the *Gpa2* promoter is able to drive *Rx1*-mediated extreme resistance against PVX in the shoots of potato.

The activity of the putative *Gpa2* promoter was tested in transgenic potato plants harbouring *pGPAII::Gpa2* upon nematode infection in the greenhouse. This resulted in almost a complete reduction of the number of females on roots of transgenic plants infected with the avirulent nematode population D383 compared to plants infected with the virulent nematode population Rookmaker (Fig. 3C). A more mild resistance response was obtained for the potato clone SH containing the endogenous *Gpa2* gene, which shows that the original *Gpa2* promoter activity is retained in the selected DNA fragment. In addition,

transgenic plants harbouring pRXI::*Gpa2* were included in the nematode resistance test, to see whether the *Rx1* promoter is able to drive *Gpa2*-mediated nematode resistance. This resulted in a similar reduction in the number of cysts on plants infected with the avirulent nematode population D383, whereas normal nematode development was observed on roots infected with the virulent population Rookmaker. These data show that the *Rx1* and *Gpa2* promoter and terminator sequence are interchangeable and able to drive either nematode resistance in the roots as well as virus resistance in the shoot of potato. Apparently, transcriptional regulation of separate members from a single R gene cluster can be highly conserved and independent of their recognition specificity (virus vs. nematode) or target tissue (shoots vs. roots).

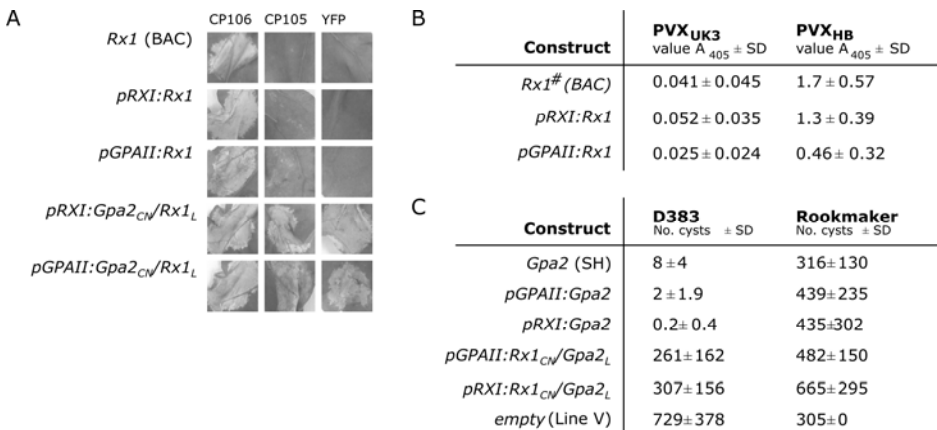


Figure 3. A. Agroinfiltration assay on *N. benthamiana* leaves of *Rx1* and *Gpa2_{CV}/Rx1_L* when coexpressed with the *Rx1* elicitor CP106, the virulent control CP105 and YFP as negative control. The chimeric constructs were expressed from the endogenous *RX1* and *GPAII* promoters. The original BAC clone harbouring *Rx1* was used as a positive control. Images were taken 7 days post infiltration. **B.** Greenhouse virus resistance assay: mean absorbance values (A₄₀₅) are shown of homogenate of secondary compound leaves in ELISA of transgenic potato plants. Genes were expressed from the 2.8 kb of 5'-UTR sequence of the wt *Rx1* gene (*pRXI*) or 2.8 kb of 5'-UTR sequence and 0.5 kb of 3'-UTR sequence of the wt *Gpa2* gene (*pGPAII*). Leaves were harvested three weeks after primary leaf inoculation with PVX_{UK3} or PVX_{HB}. Between 4 to 12 plants from 2 to 4 independent lines were assayed per construct. **C.** Greenhouse nematode resistance assay on transgenic potato plants harbouring the *Gpa2* gene and the domain swap construct *Rx_{CV}/Gpa2_L* under control of the endogenous *Rx1* or *Gpa2* promoter and terminator. Plants were tested with the avirulent Pa₂-D383 population and the virulent population Pa₃-Rookmaker of the potato cyst nematode *Globodera pallida*. Three independent transgenic lines were assayed in multiple replicates for each transgene. Cysts were counted on these plants at 16 weeks post inoculation and the average number ± SD are shown. Plants were scored resistant when the number of cysts found on the roots of the plants was < 20.

To test whether the nematode resistance responses mediated by the transgenic lines harbouring the *GPAIL::Gpa2* and *35S::Gpa2* were indistinguishable from the wild type, roots of *in vitro* grown transgenic plants were infected with avirulent pre-parasitic second stage juveniles of *G. pallida* D383 for microscopic observations (Fig. 4). As a control, roots were infected with the virulent population Rookmaker resulting in normal nematode development on all roots (Fig. 4A). In the wild type resistant roots of SH, however, a variable and mild resistance response was observed resulting in the arrest of nematode development also in later stages of their life cycle, although occasionally some avirulent nematodes were able to develop on resistant roots (Fig. 4B). The encapsulation of the induced feeding cell by a layer of necrotic cells resulted in the starvation of the developing nematodes and subsequently, the appearance of a small number of translucent undeveloped adult females (Fig. 4C). The majority of the infective juveniles, however, were arrested by a hypersensitive-like response at the feeding site (Fig. 4D). This response explains also the detection of a low number of adult females on the roots of SH in the greenhouse resistance assay, whereas hardly any females were detected on roots of the transgenic resistant plants. Normal development of adult females was observed on the roots of the susceptible potato Line V, which was used for transformation (Fig. 4E). It was noticed that nematode development was completely inhibited by a local cell death response at the feeding sites in transgenic roots when Gpa2 was expressed under control of the native GPAIL promoter (Fig. 4F) or the 35S promoter (Fig. 4G and H). These data suggest that expression of Gpa2 in the background of the potato genotype line V is more effective than in the potato clone SH.

The CC-NB-ARC domain of Rx1 signals mild nematode resistance to G. pallida

For the chimeric *Gpa2_{CN}/Rx1_L* construct we showed that it is autoactive when driven by the 35S promoter, but regains its wild-type phenotype when lowering the expression level by reducing its translation efficiency. To study the effect of expression levels on the exchangeability of the Gpa2 and Rx1 CC-NB and LRR domains in more detail, the domain swap construct was expressed under control of its endogenous promoter and terminator regions. When expressed under control of the native regulatory sequences of both Gpa2 and Rx1 the *Gpa2_{CN}/Rx1_L* construct exhibited a constitutive cell death response in a transient assay on leaves of *N. benthamiana* (Fig. 3A). Consequently, no stable transgenic potato plants could be generated for this construct in an *Agrobacterium*-mediated plant transformation assay. Apparently, the expression level of this domain swap construct under control of its native promoter sequence was still above the activation threshold.

To test the hypothesis if the CC-NB domain of Rx1 was able to mediate nematode resistance to the potato cyst nematode *G. pallida*, a reciprocal construct consisting of the CC-NB region of Rx1 and the LRR region of Gpa2 was constructed (*GPAIL::Rx1_{CN}/Gpa2_L*). As the cognate elicitor of *Gpa2* is unknown, the functionality of the recombinant gene product was tested in a nematode resistance assay. When *Rx1_{CN}/Gpa2_L* was expressed from the native Gpa2 and Rx1 promoter and terminator sequences, the chimeric gene lost its ability to mediate nematode resistance in transgenic potato plants (Fig. 3C). Proper expression of the transgene was confirmed by RT-PCR (data not shown). Real time RT-PCR was performed, showing that expression from the GPAIL promoter leads to lower transcript levels when compared with expression from the 35S promoter (Fig. 5A). This was confirmed by the detection of a

50 to 100 times lower expression level of the 27-kD green fluorescent protein (GFP) on anti-GFP Western blots (Fig. 5B). Furthermore, the discrepancy between the functionality of the wild-type Gpa2 protein and the chimeric Rx1_{CN}/Gpa2_L protein could not be explained by a difference in protein stability. Western blot analyses of GFP-Gpa2 and GFP-Rx1_{CN}/Gpa2_L showed that both proteins accumulated to similar levels (Fig. 5C). These data show that the endogenous regulatory sequences are not able to drive functional expression of this chimeric protein.

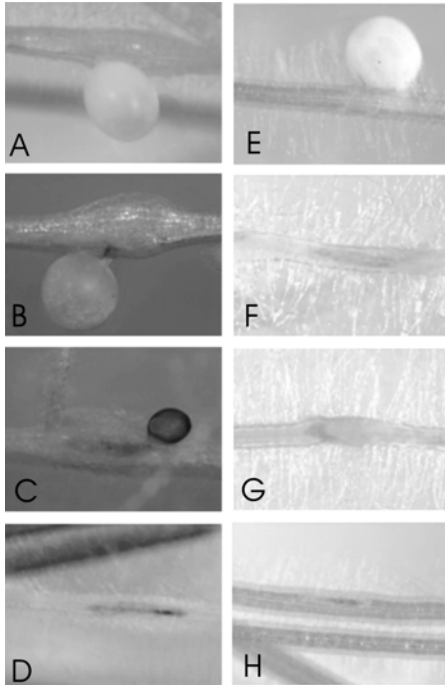


Figure 4. A. Normal development of adult females of the virulent *Globodera pallida* population Rookmaker on *in vitro* grown plants of the diploid potato cone SH harboring the endogenous *Gpa2* gene. **B.** A small number of the infective nematodes from the avirulent *G. pallida* population D383 develops into normal adult females on resistant roots of SH. **C.** The mild *Gpa2* resistance response in SH results in an arrest in nematode development resulting in typical undeveloped translucent females for D383. **D.** The majority of the infective nematodes of D383 is blocked in SH by a local cell death response at the onset of parasitism. **E.** Infective nematodes from D383 develop into normal adult females on susceptible roots of transgenic control plants (line V) harboring an empty vector. On transgenic potato roots harboring either the GPAlI::Gpa2 (**F**) or 35S::Gpa2 (**G** and **H**) constructs, nematode development was also inhibited by a local hypersensitive response at the feeding site.

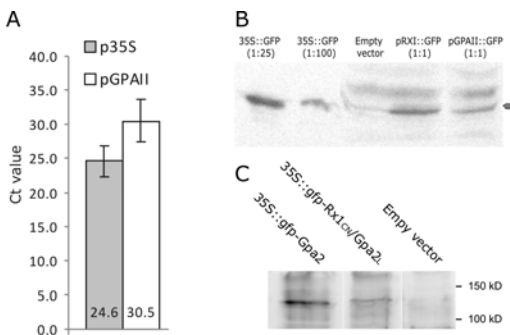


Figure 5. A. Real-time RT-PCR was performed to compare transgene expression under control of the CaMV 35S promoter or the endogenous *Gpa2* promoter upon agroinfiltration in *N. benthamiana* leaves resulted in a Δ Ct of about 6. Results were obtained in two independent experiments. **B.** Comparison of protein production under control of the 35S promoter and the GPAlI promoter on Westernblot shows that the amount of protein is significant lower (about 100 fold) for constructs driven by the endogenous promoter. **C.** Detection of the chimeric protein GFP-Rx1_{CN}/Gpa2_L on Westernblot with anti-GFP antibody after agroinfiltration shows that it is produced in similar amounts as the wild type GFP-Gpa2 protein *in planta*.

To investigate whether we could regain wild type nematode resistance by increasing the expression level of the chimeric protein Rx1_{CN}/Gpa2_L, its function was tested under control of the stronger CaMV 35S promoter and the Tnos terminator (Fig. 6A). Greenhouse experiments showed that transgenic plants expressing the *Gpa2* gene under control of the CaMV 35S promoter and the Tnos terminator were resistant to the avirulent population D383 of the potato cyst nematode *Globodera pallida*, but susceptible to the virulent *G. pallida* population Rookmaker. Similar results were obtained for the wild-type resistant potato clone SH (*Solanum tuberosum* ssp. *andigena*) harbouring the *Gpa2* gene (Fig. 6B). The potato line V, a susceptible diploid potato clone used to create the transgenic resistant plants, was susceptible in all cases as expected. For transgenic potato plants harbouring *Gpa2* under control of the 35S and native GPAPII promoter, a similar wild type resistance response was observed when tested under *in vitro* conditions (Fig. 6C) However, transgenic plants harbouring an N-terminal GFP fusion with Rx1_{CN}/Gpa2_L under control of both the constitutive CaMV 35S and GPAPII promoter resulted in a similar loss of function phenotype as observed for this chimera under control of the endogenous regulatory sequences. Apparently, enhancing the expression levels could not compensate for the lack of functionality of this chimera.

Previously, it was demonstrated for an autoactive construct identical to *Gpa2*_{CN}/Rx1_L that restoration of the compatibility between the N-terminal end of the LRR of *Gpa2* and its ARC2 domain was essential for proper gene function (Rairdan et al 2006). Therefore, the first five LRRs of Rx1 (L5) were re-introduced to see if we could restore wild type nematode resistance. Infection of transgenic *in vitro* plants harbouring the construct Rx1_{CN}L5*Gpa2*_{L6-15} under control of the 35S promoter and Tnos terminator resulted in wild type resistance to the avirulent *G. pallida* population D383 (Fig. 6D). A similar phenotype was obtained for *Gpa2* under the same experimental conditions. These data demonstrate that the compatibility between the N-terminal end of the LRR of Rx1 with its ARC2 domain is required for functionality of the chimera. Furthermore, it is concluded that the remaining part of the LRR region of *Gpa2* is the sole determinant of nematode recognition. Fusion of this moiety to the CC-NB-ARC-L5 region of Rx1, normally involved in the activation of extreme resistance against PVX, results in a functional recombinant R protein conferring pathotype-specific nematode resistance.

A



B

Potato line	D383 (No. cysts \pm SD)	Rookmaker (No. cysts \pm SD)
SH (resistant)	8 \pm 4	316 \pm 130
35S:: <i>Gpa2</i>	0.1 \pm 0.2	599 \pm 293
Line V (susceptible)	729 \pm 378	305 \pm 0

C

Potato line	D383 (No. cysts \pm SD)
<i>GPALII::Gpa2</i>	0
35S:: <i>Gpa2</i>	0
<i>GPALII::Rx<sub>CN</sub>/Gpa2<sub>L</sub></i>	43 \pm 15
<i>GPALII::GFP-Rx<sub>CN</sub>/Gpa2<sub>L</sub></i>	49 \pm 29
35S:: <i>GFP-Rx<sub>CN</sub>/Gpa2<sub>L</sub></i>	18 \pm 20
Line V (susceptible)	31 \pm 21

D

Potato line	D383 (No. cysts \pm SD)
<i>GPALII::Gpa2</i>	0
35S:: <i>Gpa2</i>	0
35S:: <i>Rx<sub>CNLS</sub>Gpa2<sub>L6-15</sub></i>	0.9 \pm 1.1
Line V (susceptible)	22 \pm 6.5

Figure 6. A. The domain swap construct *Rx1_{CN}/Gpa2_L* was obtained by exchanging the LRR domain of *Rx1* with the corresponding domain of *Gpa2* via the *ApalI* restriction site. **B.** Greenhouse nematode resistance assay on transgenic potato plants harbouring the *Gpa2* gene under control of the CaMV 35S promoter and Tnos terminator. The diploid potato clone SH, which contains the wild type *Gpa2* gene, was used as a resistant control plant. The diploid potato clone line V, which was used to create the transgenic plants, was used as a susceptible control. Plants were tested with the avirulent Pa₂-D383 population and the virulent population Pa₃-Rookmaker of the potato cyst nematode *Globodera pallida*. Three to five independent transgenic lines were assayed in multiple replicates for each transgene. Cysts were counted on these plants at 16 weeks post inoculation and the average number \pm SD are shown. Plants were scored resistant when the number of cysts found on the roots of the plants was $<$ 20. **C.** In vitro nematode resistance assay on transgenic potato plants harbouring the chimera *Rx1_{CN}/Gpa2_L* under control of the native *Gpa2* promoter or the 35S promoter. Transgenic plants harbouring the full length *Gpa2* gene were used as resistant controls. Roots were inoculated with the avirulent Pa₂-D383 population of the potato cyst nematode *Globodera pallida*. Three independent transgenic lines were assayed in multiple replicates for each transgene. Adult females were counted on these plants at 16 weeks post inoculation and the average number \pm SD are shown. **D.** In vitro nematode resistance assay on transgenic potato plants harbouring the chimera *Rx1_{CNLS}/Gpa2_{L6-15}* under control of the CaMV 35S promoter. Transgenic plants harbouring the full length *Gpa2* gene were used as resistant controls. Roots were inoculated with the avirulent Pa₂-D383 population of the potato cyst nematode *Globodera pallida*. Three independent transgenic lines were assayed in multiple replicates for each transgene. Adult females were counted on these plants at 16 weeks post inoculation and the average number \pm SD are shown.

Discussion

Exchanging the recognition specificity determining LRR domain of Gpa2 and Rx1 showed that the CC and NB domains could both mediate extreme virus resistance in the shoots and a mild nematode resistance response in the roots of potato. To our knowledge, this is the first example of the formation of functional bidirectional chimeric R proteins between two members of a single *R* gene cluster that confer resistance to two completely unrelated pathogens with distinct modes of parasitism and different routes of invasion. Most *R* genes are located in clusters in the plant genome and evolve via single base substitutions, small deletions/insertions, and intra- and intergenic sequence exchanges (Baumgarten et al., 2003; Kuang et al., 2004; Leister, 2004). The exchange of functional domains between two *R* genes without disturbing pathogen recognition and disease signalling as shown in this paper provides experimental evidence for the hypothesis that divergent selection at complex *R* gene loci may result in resistance specificities to radically different pathogens, irrespective the recognition specificity of the parental *R* genes. Apparently, the structural backbone of these modular proteins forms a framework in which intergenic sequence exchange is allowed, but our experiments also point out the functional constraints that act on the generation of effective R proteins by intergenic recombination. The observation that the CC and NB-ARC domains of Gpa2 and Rx1 are versatile modules that can mediate resistance to widely different pathogens is corroborated by the recent finding that the *Arabidopsis* resistance gene *RPP13* is able to confer resistance to transgenic strains of *Pseudomonas syringae* and turnip mosaic virus carrying the cognate effector *ATR13* from the oomycete *Hyaloperonospora parasitica* (Rentel et al., 2008).

Both CC-NB-ARC domains of Rx1 and Gpa2 are able to facilitate an extreme and a mild resistance response. This striking difference between the two resistance phenotypes suggests that the pathogen determines to some extent the outcome of the resistance response. The effectiveness of *Globodera pallida*'s secreted effectors in suppressing plant immunity provides a plausible explanation. Pathogens like bacteria, fungi, oomycetes and nematodes secrete an impressive array of proteins of which many are thought to be involved in suppressing plant defences (Gurlebeck et al., 2005; Ridout et al., 2006; Thomas, 2006; Truman et al., 2006; da Cunha et al., 2007; He et al., 2007). However, not all resistance responses to feeding cell-inducing nematodes are mild. The resistance proteins Hero and Mi-1.2 respond with a fast HR upon nematode infection (Sobczak et al., 2005; Williamson and Kumar, 2006). Another explanation for the milder Gpa2 response is that both the concentration of the elicitor and the efficiency of the recognition by the LRR domains play a role in eventual response levels.

Our data indicate that no pathogen-specific barriers on the level of pathogen-specific responses may exist within *R* gene clusters and that *R* gene clusters may generate resistances to novel pathogens in relatively short evolutionary time scales.

Assays on transgenic potato plants showed that the endogenous promoters of Gpa2 and Rx1 are exchangeable and that the resistance phenotypes were indistinguishable from their wild types. This means that the regulatory sequences for both genes allow for proper expression in under- and aboveground plant tissues, and pose no limitation to the formation of new specificities against pathogens with diverse lifestyles. This could also explain why most *R* genes are constitutively expressed at low levels throughout the plant, even in tissues that are

normally not invaded by the cognate pathogens. Exchangeability of regulatory sequences of R gene homologs in one cluster provides additional versatility to adapt quickly to a wide range of pathogens. For example, the R gene *Mi-1.2* is expressed constitutively at low levels in all plant parts and confers resistance against nematodes and aphids and whitefly.

Recently, it was shown that RanGAP2 binds to the CC domains of both Rx1 and Gpa2 (Sacco et al., 2007). Its presence is necessary for full Rx1-mediated PVX resistance (Tameling and Baulcombe, 2007), and its overexpression has an activating effect on Rx1 (Sacco et al., 2007). The N-terminal domain of several R proteins has been shown to be the binding place of a guarded host protein, a role that links it to pathogen recognition (Mackey et al., 2003; Mucyn et al., 2006; Ade et al., 2007). If Rx1 and Gpa2 are actually guarding RanGAP2, this would imply that exchanging specificities by exchanging the LRRs was only possible because they guard the same host protein which is targeted by both pathogens. Although a role for RanGAP2 in Gpa2-mediated resistance has not been shown yet, it could be a virulence target for *G. pallida*. Both PVX and potato cyst nematodes recruit the plant cell machinery for their own benefit and reproduction. In that case the specific recognition by the LRR could be triggered via a pathogen specific modification of the guarded protein or a specific interaction with the elicitor-guardee complex.

It is likely that the principles we observed in this study play a prominent role during the evolution of R proteins. The constitutively active phenotypes we observed for several chimeric constructs show that sequence divergence and coevolution between domains constrain the possibilities for reshuffling sequences within R gene clusters. The autoactivity presents a strong selection factor as was illustrated in this study by the inability to regenerate transgenic potato plants with the constitutively active R gene constructs. However, regulation of transcript levels, translation efficiency or protein stability may assuage the effects of domain incompatibility in newly formed chimeras as demonstrated in this study. This presents us with a model of R gene evolution wherein recognition specificity, activation sensitivity, and protein concentration together determine the eventual resistance response.

Materials and methods

DNA constructs

For expression under the control of the double enhanced CaMV 35S promoter and Tnos terminator, Rx1 was amplified from the binary plasmid pBINRx1 (van der Vossen et al., 2000) using the primers 5GpRxbn and Rxrev (Table 1) and cloned into the NcoI-SalI sites of pUCAP (van Engelen et al., 1995), resulting in pUCAPRx1. For *Gpa2*, the proximal end was amplified from pBINRGC2 (van der Vossen et al., 2000) with the primers 5GpRxbn and GpRxSturev (Table 1) to generate a NcoI-AvrII fragment, which was cloned together with an AvrII-PstI fragment from pBINRGC2 into the NcoI-PstI digested pUCAPRx1.

The Rx1 3'UTR (transcription termination) region was amplified from pBINRx1 using the primers 5UTRkp and 3UTRrev (Table 1) and cloned into the KpnI-PacI sites of the reporter plasmid pUCAPYFP, replacing Tnos. Next, the promoter region of *Rx1* (2805 bp between the XbaI site and ATG startcodon) was cloned in two steps. First, the region between the DraIII site (-1429 bp) and the startcodon was amplified from pBINRx1 using the primers bRxAdelf and RxbnREV (Table 1) and second, the DraIII-NcoI fragment was cloned together with the 1431 bp AscI-DraIII fragment of pBINRx1 into pUCAPYFP, replacing p35S (AscI-NcoI). The *Gpa2* 3'UTR region was amplified from pBINRGC2 using the primers 5UTRkp and 3UTRrev (Table 1) for cloning in the KpnI-PacI sites of pUCAPYFP, replacing Tnos. The *Gpa2* promoter region was constructed in two steps. First, the region between the BstZ17I (SnaI) site (-2744 bp) and the startcodon was amplified from pBINRGC2 using the primers bGpaSnaIf and GPbnREV (Table 1). This BstZ17I-NcoI fragment was cloned alongside the 720 bp PacI-BstZ17I fragment of pBINRx1, fused to a PacI-AscI adapter consisting of AD1 and AD2 (Table 1), into the AscI-NcoI digested pUCAPYFP with *Gpa2* 3'UTR after digestion with AscI-NcoI. Thereafter, the YFP sequence was subsequently replaced by the coding sequence of *Rx1* and *Gpa2* via the NcoI and KpnI restriction sites.

The domain swap constructs *Gpa2_{CN}/Rx1_L* and *Rx1_{CN}/Gpa2_L* were made by exchanging the LRR fragments of *Gpa2* and *Rx1* using the unique ApaLI and PstI site, which are conserved and situated in the beginning and the end of the LRR encoding region of the genes, respectively.

The N-terminal GFP fusion constructs were created by first providing GFP with NcoI and SstI-KpnI sites and cloning of this fragment in pUCAP. Then the AscI-SstI (35S::GFP) was cloned with a 12 amino acids encoding linker (-GGGSGGGSGGG-) into the pGPAlI driven Rgene constructs.

The leaky scan construct 35S_{LS}::Gpa2_{CN}/Rx1_L was created following the same procedure as for 35S::Gpa2_{CN}/Rx1_L, but in this case the *Gpa2* sequence was amplified with Gpa2LSFor instead of 5GpRxbn as forward primer. For the leaky scan GFPmyc6 construct 35S_{LS}::GFPmyc6, GFP was amplified with the primer pair 5nGFP and 3CFP. The PCR fragments were transferred as NcoI-SstI fragments into pRAPmyc6, pGPAlImyc6 and pRXImyc6. The 6 fold myc-tag, present in these vectors was built from 3 tandem repeats generated by triple fusion of the NheI-SpeI fragments of the annealed oligos mMYC1 and mMYC2 (Table 1).

The PVX coat proteins CP106 and CP105 were amplified from the PVX amplicons pGR106 (Jones et al., 1999) containing cDNA of the *Rx1*-avirulent PVX strain UK3 and pGR105 containing cDNA of the *Rx1*-resistance breaking strain HB (Goulden et al., 1993), respectively, using the primers 5UK3cp and 3UK3CP (Table 1) for CP106 and 5HBcp and 3HBCP (Table 1) for CP105. The products were cloned into the NcoI-KpnI sites of pUCAP between the CaMV 35S promoter and the Tnos terminator.

For agro-infiltration assays and *Agrobacterium tumefaciens* – mediated plant transformation, the expression cassettes containing the constructs were cloned into the AscI and PacI sites of the binary vector pBINPLUS (van Engelen et al., 1995) and transformed to *A. tumefaciens* (pMOG101).

Agroinfiltration assays

Agrobacterium tumefaciens strain pMOG101 were cultured for agroinfiltration as described earlier (Van der Hoorn et al., 2000). For co-infiltration experiments, cultures were mixed prior to infiltration. Leaves were infiltrated of 6 weeks old *Nicotiana benthamiana* plants grown in the greenhouse at 20°C and 16 hours of light. Each combination was tested at least *in duplo* on two different plants in at least two independent experiments.

Plant transformation

The susceptible diploid potato line V was used for *Agrobacterium*-mediated plant transformation as described (van Engelen et al., 1994). Genomic DNA was extracted using the Dneasy plant mini kit (Qiagen) for PCR to analyse the incorporation of the transgene in the plant genome. RNA was extracted using Trizol LS Reagent (Life Technologies) for RT-PCR using the Superscript TM First strand synthesis system (Life Technologies) to test expression of the transgene with gene specific primers.

Virus resistance test

To obtain infectious virus particles, leaves of *Nicotiana benthamiana* were agroinfiltrated with the PVX amplicons pGR106 and pGR105. Systemically infected leaf material was homogenized in 10 ml of 50 mM NaPO₄ buffer pH 7 and 20 µl was used for inoculation by rubbing four leaves per plant of 4 weeks old transgenic potato plants with carborundum powder. At least 3 plants per construct were used. As a control for each construct one plant was mock inoculated. Infected plants were grown in the greenhouse at 23°C and 16 hours of light. Three weeks after infection 10 leaf discs were taken from compound leaves of the apex and homogenized as described above. The relative virus concentration was determined using DAS-ELISA (Maki-Valkama et al., 2000). ELISA plates were coated with a 1:1000 dilution of a polyclonal antibody against PVX to bind the antigen and an alkaline phosphatase conjugated version of this antibody against PVX conjugated with alkaline phosphatase was used for detection (a kind gift of J. Saaijer).

Table 1. Primer, adapter and linker sequences

	Primer / adapter /linker sequence 5' > 3'
Fh1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTTTGTTCATTTTCATACTGAGAG
Fh2	GGGGACCACTTTGTACAAGAAAGCTGGGTGGTAGCTCAGACCAAC
5GpRxbn	TTTTTGGATCCATGGCTTATGCTGCTGTTACTTCCC
Rxrev	GATAGCGTCGACCACCTTAACTACTCGCTGCA
GpRxSturev	CAAAGAAAGAAGGCCATAGGAGTAC
3CCNot	GTGGTACCTTAAGCGGCCGACCAACCATTATATTCTCGGGCTGC
5CFPsbn	TCGACGGATCCATGGTGAACAAGGGCGAGGAGCTGTTT
3CFP	AGGTACCTTAGCTCATGACTGACTTGTAGAGCTCGTCCATGCCGAGAG
5nGFP	CGGATCCATGGATGGTGAACAAGGGCGAGGAG
5UTRkp	TGGTACCTTCGACGCGAGTAGTTAAGGTGTTCTGAGGAC
3UTRrev	CTTAATTAACCCGGGAGATTGAGACTCCCAAGAAAGG
bRxAdelf	GAGATTCACATATGTGCATCACCCAC
RxbnREV	AGCATAAGCCATGGATCCAAAAAATAGAAATATCTCT
bGpaSnalf	CAATTGTATACTTTCTTGCC
GPhnREV	AGCATAAGCCATGGATCCAAAAAATAGAAATATCTCT
AD1	CGCGCCACCGGTTCTAGAT
AD2	CTAGAACC GGTTG
Gpa2LSFor	TACGACCATGGATGGCTTATGCTGCTGTTAC
NBSeRev	TGGTACCTTAAGAAATTCATGTTTTCGAGCTTCCCTCAAACAG
For-LRRrx-1	CTCGACATTATTGCGGCAAGAAGC
Rev-LRRrx-1	ATGAATTTTGTGAATGTTATCAGAGG
5UK3cp	TCCATGGGCGGTGGAGTCATGAGCGCACCAGCTAGCACAAACACAGCC
3UK3CP	AGGTACCTGCGGTTATGGTGGTGTAGAGTGACAAACAGC
5HBcp	TCCATGGGCGGTGGAGTCATGACTACGCCAGCCAACACCCTC
3HBcp	AGGTACCTGCGGTTATGGTGGGGGTAGTGAGATAACAGC
L12for	AGCTCTACAAGGGCGGGGAAGTGGAGGCGGATCCGGGGGAGGCAGCATG
L12rev	CTGCCTCCCCGGATCCGCCTCCACTTCCGCCGCCCTTGATG
Nb.actinF	CCAGGTATTGCCGATAGAATG
Nb.actinR	GAGGGAAGCCAAGATAGAGC
Gpa.LRR-F A	GGTCCATACCTCGTTATCTTTATCG
Gpa.LRR-R B	TCATCTTCATCTTCATCTGTTGTC
Rx.C-F	GACAACAGATGAAGATGATGATG
Rx.D-R	CCTCAGAACACCTTAACTACTCC
mMYC1	GGCCGCTAGCGAGCAAAGCTCATTAGTGAGGAAGACTTAGGTGAACAGAAGCTAATCTCTGAAGAGGATCT TACTAGTTAAT
mMYC2	CTAGATTAAGTAGTAAAGTCTCTTCAGAGATTAGCTTCTGTTACCTAAGTCTTCTCACTAATGAGCTTTTG CTCGCTAGC

Nematode resistance test

For the nematode resistant tests, the avirulent *Globodera pallida* population D383 and the virulent population Pa3-Rookmaker were used for infection of transgenic potato lines. The resistant diploid potato clone SH harboring the Gpa2 gene (van der Vossen et al., 2000) was used as a control. Stem cuttings of *in vitro* potato plants were grown on agar plates and after three weeks, roots were infected with approximately 300 surface sterilized second stage juveniles per plate as described (Goverse et al., 2000). For each construct three independent transformed lines were used. After 21 days and 8 weeks nematode development was monitored by microscopic inspection. For the resistance test in soil, transgenic potato plants were transferred from *in vitro* cultures and grown under greenhouse condition for two months and then inoculated with 10.000 eggs per pot of *G. pallida* Rookmaker or D383. Three and a half month after inoculation cysts were isolated from the roots and counted.

Real time RT-PCR

Leaves of *N. benthamiana* were infiltrated with *Agrobacterium tumefaciens* (pMOG101) carrying constructs of interest. At 48 hours after inoculation leaves were collected and frozen in liquid nitrogen. For RNA extraction, 60 mg of leaf tissue was used for the isolation of total RNA with the RNeasy Plant Mini Kit from Qiagen, including extra DNase treatment. The total RNA concentration was measured using a NanoDrop spectrophotometer (Isogen) and all samples were adjusted to the same concentration. For cDNA synthesis, Super Script III (Invitrogen) and random hexamer primers were used. For

real-time PCR reactions, primers were designed for Gpa2 and Rx1 using the Beacon 4.0 software. Actin was used as a reference gene. The following primers were used: Nb.actinF , Nb.actinR, Gpa.LRR-F A, Gpa.LRR-R B, Rx.C-F, and Rx.D-R (Table 1). The iQ SYBR Green Supermix (Bio-Rad) was used in a reaction volume of 25 μ l (7.5 μ l water, 2x 1 μ l primer (5 mM), 3 μ l template, 12.5 μ l Supermix). The annealing temperature for the actin and Gpa.LRR primers was 64 °C and for the Rx primers 63 °C. The applied PCR program was 98 °C for 3 minutes followed by 50 cycles of 95 °C for 10 sec and 63°C for 20 sec and 70 °C for 30 sec.

Protein analysis

Total protein extract of *A. tumefaciens* transformed *N. benthamiana* leaves was made by grinding leaf material in protein extraction buffer (50 mM Tris, 10% glycerol, 150 mM NaCl, 1 mM EDTA, 20mg/ml polyclar-AT PVPP, 1 mg/ml PEFA bloc+, 5 mM DTT) on ice. The soluble fraction was analyzed by SDS-PAGE and subsequent visualisation by Coomassie Brilliant Blue staining or Western blotting and protein detection with 1:5000 diluted HRP conjugated Rabbit polyclonal anti-GFP (Novus Biologicals). HRP activity was visualised using the Pierce ECL substrate.

References

- Ade, J., Deyoung, B.J., Golstein, C. and Innes, R.W.** (2007). Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease. *Proc Natl Acad Sci U S A* **104**, 2531-2536.
- Baumgarten, A., Cannon, S., Spangler, R. and May, G.** (2003). Genome-level evolution of resistance genes in *Arabidopsis thaliana*. *Genetics* **165**, 309-319.
- Cooley, M.B., Pathirana, S., Wu, H.J., Kachroo, P. and Klessig, D.F.** (2000). Members of the Arabidopsis HRT/RPP8 family of resistance genes confer resistance to both viral and oomycete pathogens. *Plant Cell* **12**, 663-676.
- Cork, J.M. and Purugganan, M.D.** (2005). High-diversity genes in the *Arabidopsis* genome. *Genetics* **170**, 1897-1911.
- da Cunha, L., Sreerekha, M.V. and Mackey, D.** (2007). Defense suppression by virulence effectors of bacterial phytopathogens. *Curr Opin Plant Biol* **10**, 349-357.
- Dodds, P.N., Lawrence, G.J. and Ellis, J.G.** (2001a). Contrasting modes of evolution acting on the complex N locus for rust resistance in flax. *Plant J* **27**, 439-453.
- Dodds, P.N., Lawrence, G.J. and Ellis, J.G.** (2001b). Six amino acid changes confined to the leucine-rich repeat beta-strand/beta-turn motif determine the difference between the P and P2 rust resistance specificities in flax. *Plant Cell* **13**, 163-178.
- Ellis, J.G., Lawrence, G.J., Luck, J.E. and Dodds, P.N.** (1999). Identification of regions in alleles of the flax rust resistance gene L that determine differences in gene-for-gene specificity. *Plant Cell* **11**, 495-506.
- Farnham, G. and Baulcombe, D.C.** (2006). Artificial evolution extends the spectrum of viruses that are targeted by a disease-resistance gene from potato. *Proc Natl Acad Sci U S A* **103**, 18828-18833.
- Goulden, M.G., Kohm, B.A., Cruz, S.S., Kavanagh, T.A. and Baulcombe, D.C.** (1993). A Feature Of The Coat Protein Of Potato Virus-X Affects Both Induced Virus-Resistance In Potato And Viral Fitness. *Virology* **197**, 293-302.
- Goverse, A., Overmars, H., Engelbertink, J., Schots, A., Bakker, J. and Helder, J.** (2000). Both induction and morphogenesis of cyst nematode feeding cells are mediated by auxin. *Mol Plant Microbe Interact* **13**, 1121-1129.
- Gurlebeck, D., Thieme, F. and Bonas, U.** (2005). Type III effector proteins from the plant pathogen *Xanthomonas* and their role in the interaction with the host plant. *J Plant Physiol* **27**, 27.
- Hayes, A.J., Jeong, S.C., Gore, M.A., Yu, Y.G., Buss, G.R., Tolin, S.A. and Maroof, M.A.S.** (2004). Recombination Within a Nucleotide-Binding-Site/Leucine-Rich-Repeat Gene Cluster Produces New Variants Conditioning Resistance to Soybean Mosaic Virus in Soybeans. *Genetics* **166**, 493-503.
- He, P., Shan, L. and Sheen, J.** (2007). Elicitation and suppression of microbe-associated molecular pattern-triggered immunity in plant-microbe interactions. *Cell Microbiol* **9**, 1385-1396.
- Holt, B.F., Hubert, D.A. and Dangl, J.L.** (2003). Resistance gene signaling in plants - complex similarities to animal innate immunity. *Curr Opin Immunol* **15**, 20-25.
- Hulbert, S.H.** (1997). Structure and evolution of the rp1 complex conferring rust resistance in maize. *Annu Rev Phytopathol* **35**, 293-310.
- Jones, D.A. and Takemoto, D.** (2004). Plant innate immunity - direct and indirect recognition of general and specific pathogen-associated molecules. *Curr Opin Immunol* **16**, 48-62.
- Jones, L., Hamilton, A.J., Voinnet, O., Thomas, C.L., Maule, A.J. and Baulcombe, D.C.** (1999). RNA-DNA Interactions and DNA Methylation in Post-Transcriptional Gene Silencing. *Plant Cell* **11**, 2291-2302.
- Kozak, M.** (1995). Adherence to the first-AUG rule when a second AUG codon follows closely upon the first. *Proc Natl Acad Sci U S A* **92**, 7134.
- Kozak, M.** (1999). Initiation of translation in prokaryotes and eukaryotes. *Gene* **234**, 187-208.
- Kuang, H., Woo, S.S., Meyers, B.C., Nevo, E. and Michelmore, R.W.** (2004). Multiple Genetic Processes Result in Heterogeneous Rates of Evolution within the Major Cluster Disease Resistance Genes in Lettuce. *Plant Cell* **16**, 2870-2894.
- Lam, E., Kato, N. and Lawton, M.** (2001). Programmed cell death, mitochondria and the plant hypersensitive response. *Nature* **411**, 848-853.
- Leister, D.** (2004). Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance genes. *Trends Genet.* **20**, 116-122.
- Luck, J.E., Lawrence, G.J., Dodds, P.N., Shepherd, K.W. and Ellis, J.G.** (2000). Regions outside of the leucine-rich repeats of flax rust resistance proteins play a role in specificity determination. *Plant Cell* **12**, 1367-1377.
- Mackey, D., Belkhadir, Y., Alonso, J.M., Ecker, J.R. and Dangl, J.L.** (2003). Arabidopsis RIN4 is a target of the type III virulence effector AvrRpt2 and modulates RPS2-mediated resistance. *Cell* **112**, 379-389.
- Maki-Valkama, T., Valkonen, J.P., Kreuze, J.F. and Pehu, E.** (2000). Transgenic resistance to PVY(O) associated with post-transcriptional silencing of P1 transgene is overcome by PVY(N) strains that carry highly homologous P1 sequences and recover transgene expression at infection. *Mol Plant Microbe Interact* **13**, 366-373.
- Meyers, B.C., Kozik, A., Griego, A., Kuang, H. and Michelmore, R.W.** (2003). Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. *Plant Cell* **15**, 809-834.
- Mondragon-Palomino, M., Meyers, B.C., Michelmore, R.W. and Gaut, B.S.** (2002). Patterns of positive selection in the complete NBS-LRR gene family of *Arabidopsis thaliana*. *Genome Res* **12**, 1305-1315.

- Mucyn, T.S., Clemente, A., Andriotis, V.M., Balmuth, A.L., Oldroyd, G.E., Staskawicz, B.J. and Rathjen, J.P.** (2006). The tomato NBARC-LRR protein Prf interacts with Pto kinase in vivo to regulate specific plant immunity. *Plant Cell* **18**, 2792-2806.
- Rairdan, G.J. and Moffett, P.** (2006). Distinct Domains in the ARC Region of the Potato Resistance Protein Rx Mediate LRR Binding and Inhibition of Activation. *Plant Cell* **18**, 2082-2093.
- Rentel, M.C., Leonelli, L., Dahlbeck, D., Zhao, B. and Staskawicz, B.J.** (2008). Recognition of the *Hyaloperonospora parasitica* effector ATR13 triggers resistance against oomycete, bacterial, and viral pathogens. *Proc Natl Acad Sci U S A* **105**, 1091-1096.
- Ridout, C.J., Skamnioti, P., Porritt, O., Sacristan, S., Jones, J.D. and Brown, J.K.** (2006). Multiple avirulence paralogs in cereal powdery mildew fungi may contribute to parasite fitness and defeat of plant resistance. *Plant Cell* **18**, 2402-2414.
- Sacco, M.A., Mansoor, S. and Moffett, P.** (2007). A RanGAP protein physically interacts with the NB-LRR protein Rx, and is required for Rx-mediated viral resistance. *Plant J* **52**, 82-93.
- Smith, S.M. and Hulbert, S.H.** (2005). Recombination events generating a novel Rp1 race specificity. *Mol Plant Microbe Interact* **18**, 220-228.
- Sobczak, M., Avrova, A., Jupowicz, J., Phillips, M.S., Ernst, K. and Kumar, A.** (2005). Characterization of susceptibility and resistance responses to potato cyst nematode (*Globodera* spp.) infection of tomato lines in the absence and presence of the broad-spectrum nematode resistance *Hero* gene. *Mol. Plant-Microbe Interact.* **18**, 158-168.
- Sun, Q., Collins, N.C., Ayliffe, M., Smith, S.M., Drake, J., Pryor, T. and Hulbert, S.H.** (2001). Recombination between paralogs at the Rp1 rust resistance locus in maize. *Genetics* **158**, 423-438.
- Takken, F.L., Albrecht, M. and Tameling, W.I.** (2006). Resistance proteins: molecular switches of plant defence. *Curr Opin Plant Biol* **9**, 383-390.
- Tameling, W.I. and Baulcombe, D.C.** (2007). Physical association of the NB-LRR resistance protein Rx with a Ran GTPase-activating protein is required for extreme resistance to Potato virus X. *Plant Cell* **19**, 1682-1694.
- The Arabidopsis Genome Initiative** (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796-815.
- Thomas, J.H.** (2006). Adaptive evolution in two large families of ubiquitin-ligase adapters in nematodes and plants. *Genome Res* **6**, 6.
- Truman, W., de Zabala, M.T. and Grant, M.** (2006). Type III effectors orchestrate a complex interplay between transcriptional networks to modify basal defence responses during pathogenesis and resistance. *Plant J* **46**, 14-33.
- Tuskan, G.A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R.R., Bhalerao, R.P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G.L., Cooper, D., Coutinho, P.M., Couturier, J., Covert, S., Cronk, Q., Cunningham, R., Davis, J., Degroove, S., Dejardin, A., Depamphilis, C., Detter, J., Dirks, B., Dubchak, I., Duplessis, S., Ehling, J., Ellis, B., Gendler, K., Goodstein, D., Gribskov, M., Grimwood, J., Groover, A., Gunter, L., Hamberger, B., Heinze, B., Helariutta, Y., Henrissat, B., Holligan, D., Holt, R., Huang, W., Islam-Faridi, N., Jones, S., Jones-Rhoades, M., Jorgensen, R., Joshi, C., Kangasjarvi, J., Karlsson, J., Kelleher, C., Kirkpatrick, R., Kirst, M., Kohler, A., Kalluri, U., Larimer, F., Leebens-Mack, J., Leple, J.C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D.R., Nelson, C., Nieminen, K., Nilsson, O., Pereda, V., Peter, G., Philippe, R., Pilate, G., Poliakov, A., Razumovskaya, J., Richardson, P., Rinaldi, C., Ritland, K., Rouze, P., Ryaboy, D., Schmutz, J., Schrader, J., Segerman, B., Shin, H., Siddiqui, A., Terry, F., Terry, A., Tsai, C.J., Uberbacher, E., Unneberg, P., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**, 1596-1604.**
- Van der Hoorn, R.A.L., Laurent, F., Roth, R. and De Wit, P.** (2000). Agroinfiltration is a versatile tool that facilitates comparative analyses of Avr9/Cf-9-induced and Avr4/Cf-4-induced necrosis. *Mol. Plant-Microbe Interact.* **13**, 439-446.
- van der Vossen, E.A.G., van der Voort, J., Kanyuka, K., Bendahmane, A., Sandbrink, H., Baulcombe, D.C., Bakker, J., Stiekema, W.J. and Klein-Lankhorst, R.M.** (2000). Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: a virus and a nematode. *Plant J.* **23**, 567-576.
- van Engelen, F.A., Molthoff, J.W., Conner, A.J., Nap, J.P., Pereira, A. and Stiekema, W.J.** (1995). pBINPLUS: an improved plant transformation vector based on pBIN19. *Transgenic Res* **4**, 288-290.
- van Engelen, F.A., Schouten, A., Molthoff, J.W., Roosien, J., Salinas, J.s., Dirkse, W.G., Schots, A., Bakker, J., Gommers, F.J., Jongma, M.A., Bosch, D. and Stiekema, W.J.** (1994). Coordinate expression of antibody subunit genes yields high levels of functional antibodies in roots of transgenic tobacco. *Plant Molecular Biology (Historical Archive)* **26**, 1701-1710.
- Williamson, V.M. and Kumar, A.** (2006). Nematode resistance in plants: the battle underground. *Trends Genet.* **22**, 396-403.
- Zhou, T., Wang, Y., Chen, J.Q., Araki, H., Jing, Z., Jiang, K., Shen, J. and Tian, D.** (2004). Genome-wide identification of NBS genes in japonica rice reveals significant expansion of divergent non-TIR NBS-LRR genes. *Mol Genet Genomics* **271**, 402-415.

Chapter 5

A docking model for the NB-ARC and LRR domains of the CC-NB-LRR Resistance protein Gpa2

Erik Sloopweg¹, Laurentiu Spiridon², Jan Roosien¹, Patrick Butterbach¹, Rikus Pomp¹, Geert Smant¹, Andrei Petrescu², Jaap Bakker¹, Aska Goverse¹

¹ *Department of Nematology, Wageningen University, Wageningen, The Netherlands.*

² *Institute of Biochemistry of the Romanian Academy, Bucharest, Romania*

To be submitted

Abstract

The modular plant R proteins belonging to the NB-LRR class depend for functioning on extensive interdomain interactions. We present a structural model for the NB-ARC and LRR domains of the potato CC-NB-LRR R protein Gpa2 based on advanced homology modelling. On the basis of sequence exchanges between Gpa2 and the highly similar R protein Rx1 surface regions were defined on both the ARC2 and LRR domains that are required to match in the full-length proteins. Together with known autoactivating mutations these were used as spatial constraints in the computational docking of the NB-ARC and LRR structures. The resulting interaction is characterized by two main components, a conserved electrostatic interaction and hydrophobic interactions between the NB, ARC2 and the N-terminal half of the LRR. Support for this model was found through coevolution analysis and site-directed mutagenesis of positions predicted to have a specific role in the interdomain contact. The potential of this NB-ARC-LRR docking model as a framework for the interpretation of known empirical data and the design of new experiments to test R protein operational mechanisms is discussed.

Introduction

In evolution functional protein structures are often reused in new domain combinations, while retaining their basic mechanistic role (Koonin et al., 2002; Basu et al., 2008). The complex multidomain proteins belonging to the NB-LRR class of plant resistance proteins form a good example of this principle. Their central NB-ARC (van der Biezen and Jones, 1998; Albrecht and Takken, 2006) domain is a functional and structural component that can be found in various domain combinations in a broad group of proteins (Leipe et al., 2004; Ammelburg et al., 2006). The most similar counterparts of the NB-LRR proteins in the animal kingdom are proteins that have roles in innate immunity and programmed cell death, analogous to the function R proteins have in plants. These metazoan NB-ARC containing proteins can have different N-terminal domains (CARD, PYRIN,) or C-terminal domains (WD40, LRR), but are thought to operate in a similar fashion (Rairdan and Moffett, 2007; Takken and Tameling, 2009). Recognition of a signal via the C-terminal domain leads to a conformational change in the central NB-ARC domain, often followed by oligomerization and eventually the activation of subsequent signalling via the N-terminal domain. The conformational change of the NB-ARC domain coincides with the exchange of the bound ADP for an ATP, probably stabilising the active conformation. Hydrolysis of the ATP would return the domains to their inactive state.

To our knowledge no complete structure for NB-LRR R proteins is available at the moment. The availability of crystal structures for several proteins that share structural similarity with NB-LRR R proteins or their domains, however, offers insights in the activational and signalling mechanisms. Threading the sequence of the NB-ARC domains of the CC-NB-LRR I-2 on the Apaf-1 and CED-4 crystal structures, helped to place the phenotypes of many known point mutations in these domains into a structural context (Takken et al., 2006; Tameling et al., 2006). Similarly, the structural properties of the LRR domain of various R proteins have been deduced from the analogy with the 3D model of the LRR domain of the ribonuclease inhibitor (McHale et al., 2006). The LRR forms a horseshoe-like structure characterized by a parallel Beta-sheet on its concave surface. Regularly patterned leucines in

this beta-sheet point inwards to the hydrophobic core of the structure and are tightly packed in a stabilizing scaffold. The residues in between these leucines are often hydrophilic and are exposed to the solvent. The large non-globular conformation maximizes the potential interaction surface and makes this domain well suited for protein interactions.

The potato R protein Gpa2, conferring resistance against the nematode *Globodera pallida*, is highly similar to the thoroughly studied R protein Rx1, which confers resistance to Potato virus X (PVX) (van der Vossen et al., 2000). Rx1 and Gpa2 are encoded in a small cluster of 4 R gene homologs, introgressed into the genome of modern potatoes from *Solanum andigena*. The CC, NB-ARC, and LRR domains of these proteins interact with each other and recognition of the elicitor by the LRR is translated in an activation of the NB-ARC (Moffett et al., 2002; Rairdan et al., 2008). The intradomain interactions enable Rx1 and Gpa2 to form functional proteins even if their domains are coexpressed as separate protein fragments (Moffett et al., 2002). If the ARC domain of Gpa2 is combined with the LRR domain of Rx1 in a chimeric protein, the resulting construct becomes constitutively active (Rairdan and Moffett, 2006), Chapter 4 of this thesis). This indicates that intramolecular coevolution between the domains in diverging R proteins is necessary to prevent interdomain incompatibility (Rairdan and Moffett, 2006). A large set of Rx1 / Gpa2 homologs from many different *Solanum* species is available and allows insight in the evolutionary processes shaping these proteins (Butterbach *et al.*, in preparation). For both Gpa2 and Rx1 elicitors are known, which can be used to test the functionality of these proteins in rapid transient assays (Bendahmane et al., 1995; Sacco et al., 2009).

Therefore, the available functional and evolutionary data make Gpa2 and Rx1 suitable candidate R proteins for structural modelling of the individual domains and their intradomain interactions. A structural model of the NB-ARC/LRR interaction could function as a framework for the interpretation of known empirical data and the design of new experiments to test R protein operational mechanisms (Zhang, 2009). Hence, computer aided modelling of the NB-ARC domain of Gpa2 after APAF-1 Δ WD-40 (Riedl et al., 2005) was conducted, followed by modelling of the Gpa2 LRR by searching structural homology for each repeat in a structural LRR database (Spiridon *et al.*, in preparation). This provided the 3D structure domain models used as basis for a domain docking study of the NB-ARC and LRR. The functional interaction between the domains was studied via a detailed analysis of their incompatibility in chimeric Gpa2 and Rx1 proteins (Rairdan et al., 2006; Chapter 4 this thesis). A large set of refined sequence exchanges between the two proteins was created for that purpose and used as constraints in domain docking computation to limit the potential search space. The resulting docking model indicated an important role in the NB-ARC-LRR interaction for electrostatic and hydrophobic interactions.

In our opinion the docking model is supported by the site-directed mutagenesis studies, interdomain coevolution, and the sequence exchange experiments. A correlation analysis of the NB-ARC and LRR subdomains resulted in the detection of coevolution between the interacting surfaces. Furthermore, site-directed mutagenesis was used to test the role of surface features that might play an important role in the interdomain docking interface, showing that electrostatic forces play a role in the domain interaction. However, the mutations in the acidic loop did not abolish the NB-ARC to LRR interaction as shown by co-

immunoprecipitation. A predicted hydrophobic interaction between the N-terminal half of the LRR and the ARC2 domain was shown to be important in both *in trans* and *in cis* activity of the NB-ARC and LRR domains. Finally, two amino acid substitutions in the ARC2 domain, could be shown to affect either the elicitor-dependent activation or the interdomain incompatibility triggered autoactivation. These roles in Gpa2 functioning were in accordance with their position in the docking model. The implications of this model for the activation mechanism and evolution of NB-LRR R proteins will be discussed.

Results

Modelling the Gpa2 NB-ARC and LRR domains

The aim of this study is to provide structural models of the subdomains of the NB-ARC-LRR type R protein Gpa2 and a model of the interaction between the NB-ARC and LRR. No experimentally determined R protein structures have been published yet, but the NB-ARC and LRR subdomains share enough structural homology with known domain structures to make homology modelling of the Gpa2 domains feasible. In a first step suitable templates are identified for the individual subdomains.

The Gpa2 NB-ARC domain modelled after Apaf-1

Both sequence homology and automated fold recognition returned a single high confidence match for the NB-ARC region of Gpa2 in 3D databases, the human apoptosis regulator Apaf-1 (pdb entry: 1z6t). The e-values for BLAST ($4e-6$) and Phyre ($3e-24$) correspond to over 95% probability that these stretches of Gpa2 and Rx1 sequences belong to the same protein fold as human Apaf-1. The similarity between the Apaf-1 and NB-LRR R protein NB-ARC domains has been noted before and is reflected in the name of the domain (ARC: Apaf-1, R protein, CED-4)(van der Biezen and Jones, 1998; Albrecht and Takken, 2006). Although the sequence identity is low (19.3%), the similarity is significant (49.1%) and the secondary structure prediction of the Gpa2 NB-ARC matches the secondary structure of Apaf-1, making the remote homology modelling procedures possible for assessing the 3D structure of NB-ARC region starting from this template. An advanced alignment of Gpa2 to Apaf-1 was performed using ClustalW (Thompson et al., 1994) and further refined by incorporating secondary structure predictions, contact propensity profiling and locking critical amino acids (Fig. 1A).

The structure of the stretches that span the sequence conserved regions (SCR) was obtained by coordinate transfer, while the sequence variable regions (SVR), i.e. the loop regions, were randomly generated and filtered by steric constraints, followed by successive rounds of simulated annealing and energy minimization.

The final NB-ARC domain models of Gpa2 and Rx1 deviate less than 3 Å from the template structure, conserving the overall Apaf-1 five subunit structure of the fold after stability test simulations. The Apaf-1 subunits are formed by the N-terminal CARD domain, the α/β fold (NB), helical domain I (ARC1), the winged-helix domain (ARC2) and helical domain II (ARC3), all stacking against each other through extensive interdomain interactions (Riedl et al., 2005)(Fig. 1B). The central three subdomain module is known to form a nucleotide-

binding pocket and are known as NB, ARC1 and ARC2 (NB-ARC) in NB-LRR R proteins. The CARD domain and the second helical domain (ARC3) are not found in NB-LRR R proteins (Albrecht and Takken, 2006).

The amino acids forming the ADP/ATP binding pocket are conserved throughout all the NB domains and are very well aligned into the Gpa2 / Rx1 model. Here these comprise (AA Apaf-1/AA Gpa2): R129/R145), K160/K176 (P-loop), I294/L302, P321/P332 (GxP motif), L322/L333 (GxP motif), S422/S440 and H438/H459 (MHD motif). The reliability of the model is also shown by the natural way a salt bridge (R267 – D194) in the Gpa2 NB-ARC model replaces a hydrophobic contact in the Apaf-1 template structure. The shift from hydrophobic contact to a salt bridge suggests a correlated mutation in the potato resistance genes with respect to the template.

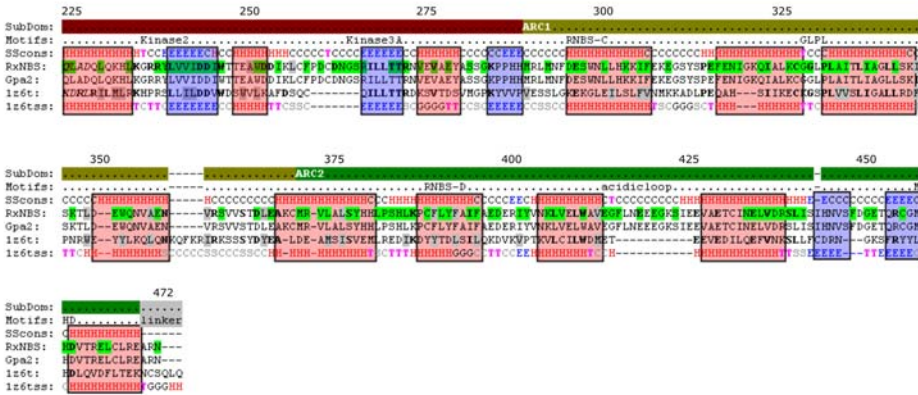
Modelling the Gpa2 LRR domain

As expected the C-terminal domain of Gpa2 was identified by motif detection and by threading as belonging to the leucine-rich repeat (LRR) fold. Precise motif detection, however, proved to be a difficult task due to the overlap and ambiguity of the motifs. Due to the large size of the LRR sequence and the structural diversity within the leucine-rich repeat fold, a one-to-one comparison with known structures was not possible. To be able to analyze the Gpa2 LRR sequence, a relational database consisting of sequence and structural information related to the LRR fold was built. There was a very large sequence departure of the LRR domain in Gpa2 from the proteins in our comprehensive structural (3D) LRR database. As a consequence, for generating structural models of the LRR region of Gpa2 a complex procedure had to be designed which is generically designated 'optimized joint fragments remote homology modelling' (Spiridon *et al.*, in preparation).

The overall sequence homology level is very low in all templates, but was shown to be better over short stretches of LRR repeats. With the optimized joint fragment procedure the short LRR fragments that show homology with templates in the database are modelled separately. The Gpa2 LRR sequence was divided in 4 overlapping fragments of about 4-5 turns by similarity with different templates, and these fragments were modelled separately. The template stretches were chosen by comparing the inter-LRR repeat lengths with those of the target, the secondary structure with the corresponding secondary structure prediction and by sequence similarity. The LRR structures of the RanGAP rna1p (1yrg)(Hillig *et al.*, 1999), the proteoglycan decorin (1xku)(Scott *et al.*, 2004), and Toll-like receptor 3 (2a0z)(Bell *et al.*, 2005), showed the best homology and were used as template for the Gpa2 LRR fragments.

The Gpa2 fragment models were generated by coordinate transfer. The 4 fragments were then joined using the superposition of the used templates as a framework (Fig. 2). The joining operation was followed by loop modelling using repeated rounds of simulated annealing and energy minimization. In this way a rough model was generated.

A



B

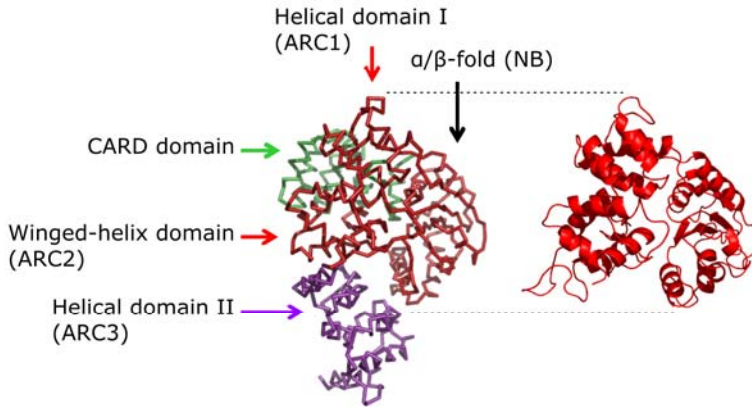


Figure 1. A. The advanced alignment of the Gpa2 and Rx1 NB-ARC domains with Apaf-1 (pdb 1z6t). Matching secondary structures are shown by red (helices) or blue (beta-strands) boxes. Numbering follows the Gpa2/Rx1 sequence. **B.** The Apaf-1 structure (pdb 1z6t), with the NB-ARC region used for homology modelling the Gpa2 NB-ARC domain shown in red. The N-terminal CARD domain (green) and second helical domain (ARC3, purple) are not present in NB-LRR R proteins.

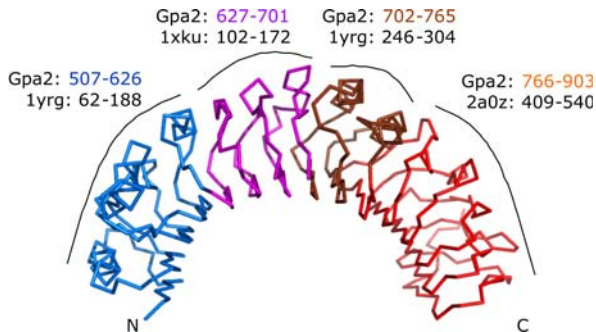


Figure 2. The 'optimized joint fragments remote homology model' of the LRR domain in Gpa2. The used structural templates are indicated by different colours in the LRR structure. The matching sequence stretches in the templates and the Gpa2 LRR are depicted in corresponding colour codes.

The next step was to assess the overall structural features of the LRR model, such as the curvature and the twist. As the inter-LRR repeat lengths and the secondary structure were similar in the templates and in the target, the curvature did not need to be corrected by other techniques. No twist was assigned to the rough model, because neither one of the 3 templates was twisted. The model was then refined by repeated rounds of energy minimization giving a fine tuned adjustment of the slight overall curvature and twist.

Characterization of the Gpa2 NB-ARC and LRR domain models

Surface electrostatic potential

The electrostatic potential of the Gpa2 NB-ARC and LRR structures was calculated using both Poisson-Boltzmann and Coulomb methods. In contrast to the NB-ARC domain that has a relatively uniform negative charge distribution over the entire surface, mapping the electrostatic potential onto the LRR 3D model revealed a marked separation of charge distribution along the structure. Judging by surface charge, the LRR domain may be divided in 2 segments. The N-terminal half, comprising repeats 1-9, is mainly basic. The C-terminal half, comprising repeat 10-15, is mainly acidic (Fig. 3A). Such a marked difference in charge distribution might be indicative of functional differences between the two halves.

Variability mapping

From a set of 75 Rx1/Gpa2 homologs (Butterbach *et al.*, in preparation) 40 full-length homologs could be used to calculate the variability per position. The other 35 were truncated sequences and therefore eliminated from the comparison. Based on the consensus sequence, the variability at a given position in the sequence was defined as the average of the Blosum62 substitution matrix values between every sequence and the consensus (equation 1).

Equation 1

$$\sum_i \frac{M(S_{ij}, C_j)}{i},$$

Where: S_i - sequence i , C - consensus sequence, j - position

These variability values were mapped on a colour scale from blue (most conserved) via green (intermediate variability) to red (maximum variability), and used to colour code the 3D structures (Fig. 3 B and C). Although overall the NB-ARC is not as variable as the LRR, variable positions distant in the primary sequence are brought close in space forming a hypervariable hotspot on the surface located around the groove formed by the NB and the ARC2 (Fig. 3B). The variability of the LRR shows an interesting distribution pattern on both the sagittal and transversal axis. Repeat 9, separating the N-terminal and the C-terminal halves of the LRR is the most variable repeat. At this position also structural differences can be seen between the Gpa2 LRR model and the Rx1 LRR model (data not shown). It is interesting to note that the separation of the LRR by repeat 9 occurs at the position where the surface charge of the repeat shifts from positive to negative (Fig. 3A). Per repeat most variability can be seen in the turn on the C-terminus of the conserved LRR motif. Combined, these variable spots form a highly variable region on one side of the LRR (Fig. 3C).

Rx1/Gpa2 interdomain incompatibility studied to support docking

Incompatibility as measure for interdomain cooperation

The NB-ARC and LRR domains of CC-NB-LRR R proteins have been shown to physically interact. Especially for Rx1 this intramolecular interaction has been studied in much detail (Moffett et al., 2002). The physical interaction underlies a fine-tuned functional interaction as was seen in domain swap experiments and mutational studies (Bendahmane et al., 2002; Rairdan and Moffett, 2006). Because Gpa2 and Rx1 are highly similar we assume that information from Rx1 can be used in modelling the Gpa2 domain interactions. Next to known autoactivating mutations in the Rx1 LRR (Bendahmane et al., 2002), we decided to use functional information derived from a detailed domain sequence exchange experiment as input for a NB-ARC/LRR docking model.

For the activation of NB-LRR R proteins, the recognition of the elicitor through the LRR domain has to be translated to the activation of the N-terminal domains responsible for downstream signalling. This intramolecular signal transduction appears to comprise more than only releasing an inhibitory interaction, as the deletion of the LRR in itself does not fully activate the N-terminal domains (Bendahmane et al., 2002; Rairdan et al., 2008; Takken and Tameling, 2009). Even autoactivating mutations like the Rx1 D460V in the MHD motif depend on the presence of the LRR for signalling and some autoactivating mutations in the LRR can transactivate coexpressed CC-NB-ARC domains (Rx1 Y712H) (Rairdan and Moffett, 2006; van Ooijen et al., 2008b).

A study by Rairdan et al. (Rairdan and Moffett, 2006) has shown that a mismatch between the Gpa2 ARC2 and Rx1 LRR domains can lead to constitutive activation. We have shown that the reverse construct containing the LRR domain of Gpa2 and the CC-NB-ARC domains of Rx1 is severely weakened in its elicitor-dependent response (chapter 4, this thesis). Because we assume these phenotypes are informative on which surfaces in the NB-ARC and LRR are minimally needed for this intramolecular communication, we decided to explore the required intercompatibility of the Gpa2 and Rx1 domains in more detail. The resulting matching domain regions were used as input for docking the Gpa2 NB-ARC and LRR domains.

Gpa2/Rx1 domain incompatibility explored via the exchange of eleven sequence fragments

A set of chimeric Gpa2/Rx1 constructs was created starting from the autoactive G13R45 (CC-NB-ARC of Gpa2 and LRR of Rx1) and inactive R13G45 (LRR of Gpa2 and CC-NB-ARC of Rx1). Both in the LRR and the CC-NB-ARC region smaller fragments of Gpa2 and Rx1 were exchanged to determine in detail which residues differing between Gpa2 and Rx1 caused the observed domain incompatibilities.

To be able to exchange sequences between Rx1 and Gpa2 several new restriction sites were introduced and several unwanted restriction sites removed in the encoding genes. Care was taken not to alter the encoded amino acid sequence and the altered genes were found functionally indistinguishable from the wild type Rx1 and Gpa2 sequences (data not

shown). In Figure 4A an overview is presented of the positions in the Gpa2 and Rx1 proteins corresponding to the sites at which sequences were exchanged between them. For comparison the positions used in Rairdan et al. 2006 are given as well, and an overview of the Gpa2 domain architecture, secondary structure prediction and amino acid variation between Rx1 and Gpa2.

An schematic overview of the thus created constructs is presented in Figures 4B and 4D. All resulting constructs were expressed under control of the CaMV 35S promoter in transient agroinfiltration assays and tested for autoactivity and their response to viral (avirulent CP106, virulent CP105) or nematode elicitors (eliciting RBP D383-1, non-eliciting RBP Rook-4). The strength of the observed hypersensitive response was ranked at a scale of zero to five, from no visible signs of HR to full necrosis. The HR was scored after two and seven days, and repeated at least 3 times. The average HR strength after 7 days is depicted in Figure 4B and 4D in the panel to the right of the schematic constructs. The autoactive response of G13R45 showed itself as a rapid cell death within two days after infiltration when the construct was not coexpressed with its elicitor. Coexpression with CP106 still intensified this response (see also Chapter 4).

Determining the Rx1 LRR region incompatible with the Gpa2 NB-ARC

Starting from the G13R45 constructs, fragments of the Rx1 LRR were stepwise replaced for the Gpa2 sequences from the junction between the ARC and LRR towards the C-terminus. A gradual loss of the autoactive response could be seen when the first three repeats of the LRR (fragments 4a1, 4a2 and 4a3) were replaced by the Gpa2 sequence (Fig. 4B), while these constructs could still be activated by the avirulent CP106. However, when the fragments encompassing LRR repeat 4 to 9 and 10 to 14 were replaced no specific activation by either the viral or nematode elicitors could be seen. Both these construct show some autoactive response, but less than the mild autoactive response of full Gpa2 expressed under identical circumstances. These results point to the LRR region containing the first three repeats as the sequence that needs to be compatible with the CC-NB-ARC. This was further explored by replacing in the background of the full Gpa2 the individual repeats and combinations thereof by their Rx1 counterparts.

Replacing only the first or only the second repeat by Rx1 sequences does not result in constructs that are phenotypically different from full Gpa2 (Fig. 4B). Exchanging both the first and second LRR repeat on the other hand gives a construct that is strongly autoactive. The addition of the third repeat from Rx1 to this exchange does not further increase the autoactivity. Exchanging in a Gpa2 background fragment 4b containing repeat 4 to 9 by the Rx1 sequence does not result in a strongly autoactive construct. The response of this construct to the nematode elicitor D383-1 however is weakened.

Altogether we conclude that the incompatibility observed for the Rx1 LRR and Gpa2 CC-NB-ARC is caused by the amino acids differing with Gpa2 that reside in the first two or three LRR repeats of Rx1. In fragment 4a1, Rx1 and Gpa2 differ in only 5 positions and in fragment 4a2 in only four positions.

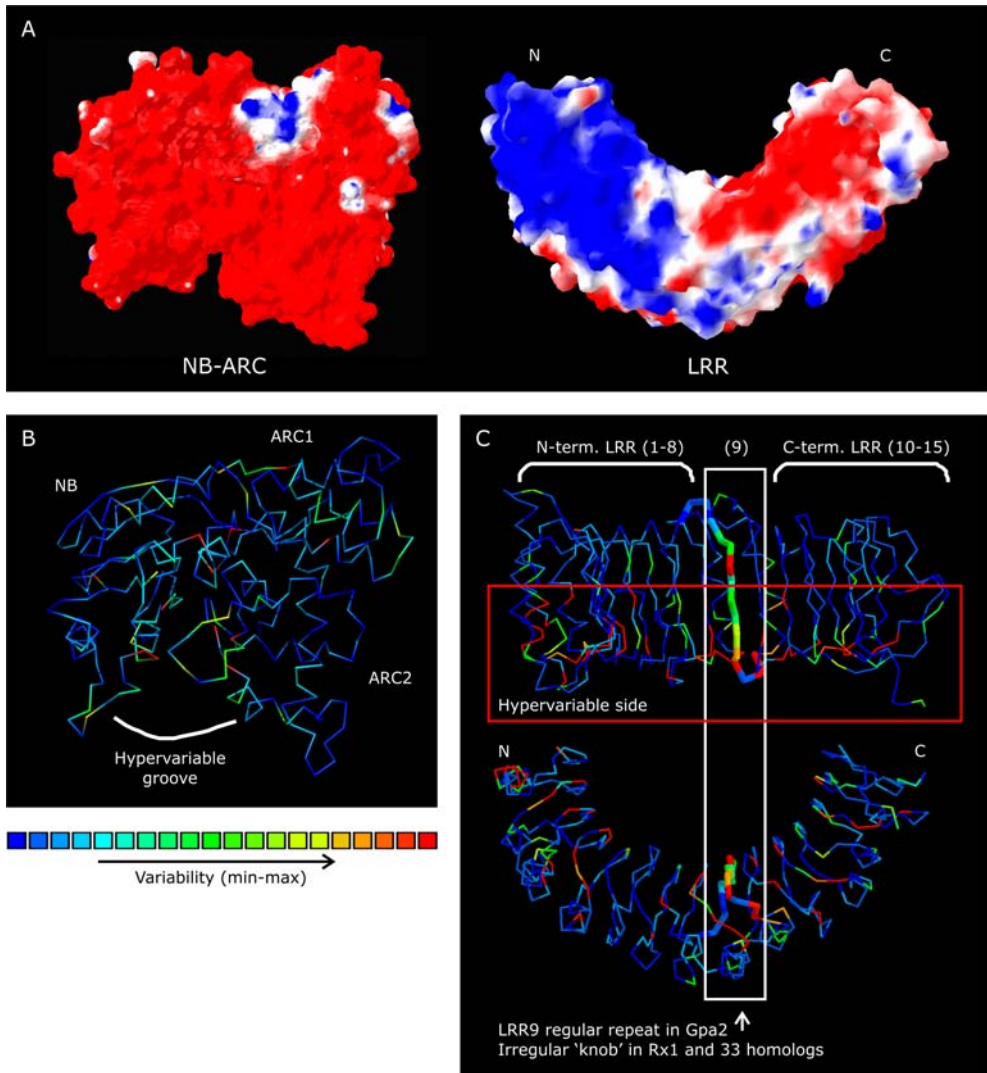
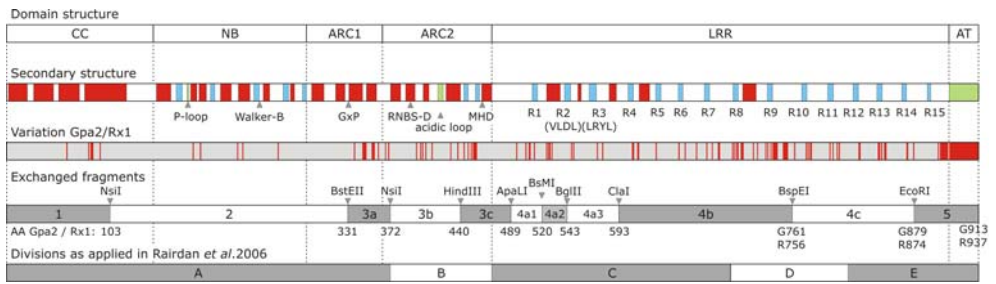


Figure 3 Characterization of the Gpa2 NB-ARC and LRR structural models. **A.** Coulomb electrostatic potential mapped on the Gpa2 NB-ARC and LRR domains. The negative charged surface is shown in red and positive charged surface areas are shown in blue. **B** and **C:** Variability as determined in a set of 40 Rx1/Gpa2 homologous sequences mapped on the backbone structure of the Gpa2 NB-ARC and LRR models. Variability is colour-coded from blue (most conserved) to red (most variable). **B.** Gpa2 NB-ARC model. The groove between the ARC2 and NB subdomains is indicated. In this groove a concentration of hypervariable sites was found on both the NB and the ARC2 side of the structure. **C.** The Gpa2 LRR model showing the concave side (upper image) and the variable side of the structure. The variable repeat 9 is boxed with a white line. In Rx1 and several homologs this repeat is irregular and is not expected to form a canonical LRR repeat (data not shown).

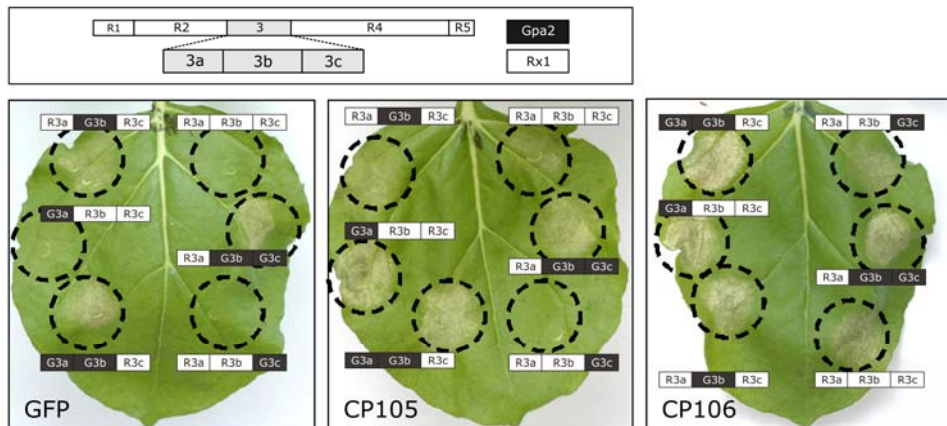
A



B



C



D

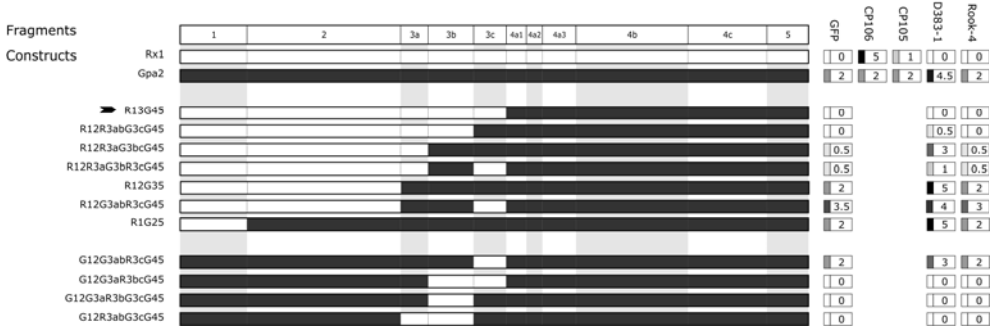


Figure 4. A. Overview of the fragments exchanged to create chimeric Gpa2 / Rx1 constructs in the context of Rx1 and Gpa2 sequence characteristics. The first row shows the domain architecture with the coiled coil (CC), nucleotide binding domain (NB), ARC1 and ARC2 domain, the LRR domain and the C-terminal domain which is extended in Rx1 with an acidic tail (AT). The second row shows the predicted secondary structure with alpha helices in red, beta-strands in blue and the P-loop, acidic loop and acidic tail in green. Below, important sequence motifs are named; the phosphate binding loop (P-loop) and Walker B (Kinase 2) motifs in the NB, the GxP, RNBS-D and MHD motifs in the ARC domain. In the LRR the central beta-strands of the LRR repeats are numbered (R1-R15) and the two conserved repeats R2 and R3 are shown with their conserved sequence motif. In the third row the amino acid positions differing between Rx1 and Gpa2 are indicated in red. For the exchanged sequence fragments the break points in the sequence in Gpa2 and Rx1 are shown below. If the numbering of the positions differed between Rx1 and Gpa2 both are given. The last row shows the divisions made in the sequence exchange experiment as published by Rairdan, et al. (2006). **B.** Gpa2 / Rx1 sequence exchange constructs. The constructs shown in these two sets are created to explore the domain incompatibility underlying the constitutive activity of G13R45. The naming of the constructs is based on the numbering of the exchanged fragments as shown in Figure 4A. The first set of constructs, indicated by CC-NB-ARC on the left is aimed at delineating a minimal Gpa2 region in the N-terminal half of the protein incompatible with the Rx1 LRR. The second set of constructs, indicated by LRR, is aimed at finding a minimal fragment of the RX1 LRR incompatible with the Gpa2 CC-NB-ARC. On the right hand side the hypersensitive phenotype of these constructs is given for combinations with GFP (to detect elicitor-independent activity), CP106 (the avirulent PVX CP), CP105 (the CP of the PVX breaker strain), D383-1, an *G. pallida* secreted specific elicitor of Gpa2, and Rook-4, a homologous protein that does not activate Gpa2 (Sacco et al., submitted). The strength of the HR phenotypes after 7 days are given by a scale from 0 (no HR) to 5 (full necrosis). **C.** The response of the ARC1-ARC2 RX1/Gpa2 sequence exchanges in the background of the Rx1 sequence. In the upper panel the position of the fragment 3a, 3b and 3c in the domain structure of Rx1 is given (R1 is the Rx1 CC, 2 the NB, 3 the ARC, R4 the LRR and R5 the acidic tail). Fragment 3a comprises all positions differing between Rx1 and Gpa2 in the ARC1 domain. The differences in the ARC2 domain are more or less equally distributed over fragment 3b and 3c. On the leaves small panels are shown which indicate the sequence composition of the constructs used in the agroinfiltration assay (Rx1 in white, Gpa2 in black). The chimeric constructs were co-infiltrated in an agroinfiltration assay with GFP, the virulent CP105 and the avirulent CP106. Images were taken five days after infiltration. All constructs responded with an HR to co-infiltration with CP106. Combined with CP105 only R3aR3bG3c did not respond with an HR. Co-infiltration with GFP identifies three of the constructs as autoactive; the ones containing R3aG3bR3c, R3aG3bG3c and G3aG3bR3c in the Rx1 sequence background. **D.** Gpa2 / Rx1 sequence exchange constructs based on R13G45. The chimeric construct R13G45 contains the Gpa2 LRR, but does not respond to the Gpa2 specific elicitor D383-1. In this set of constructs varying sequences of the CC-NB-ARC were replaced by Gpa2 sequences. The response to the D383-1 and the non-eliciting homolog Rook-4 was tested in agroinfiltration experiments. Co-infiltration with GFP was used to assess the autoactivity of the constructs. The weak autoactivation normally seen when Gpa2 is expressed under control of the CaMV 35S promoter in these transient assays is noted as HR strength 2 (on a scale of 0 (no HR) to 5 (full necrosis)).

A short fragment in the Gpa2 ARC2 domain determines the incompatibility with the Rx1 LRR domain

To define regions in the Gpa2 CC-NB-ARC incompatible with the Rx1 LRR, the Gpa2 CC-NB-ARC in the G13R45 background was replaced stepwise by the corresponding Rx1 sequence. Replacing the C-terminal part of the ARC2 (fragment 3c) resulted in a slight reduction of autoactivity and elicitor dependent activity (Fig. 4B). This fragment contains half of the positions that differ between Rx1 and Gpa2 in the ARC2. When the complete ARC2 domain (fragment 3b and 3c) was replaced by the Rx1 sequence (G12G3aR3bcR45) the autoactive response was strongly reduced. Exchanging also the ARC1 domain (G12R35) gave a construct that completely lost autoactivity and responded to CP106 in a way indistinguishable from Rx1.

When only the ARC domain in Rx1 is replaced for the Gpa2 sequence (R12G3R45) this creates a constitutively active construct which is indistinguishable from G13R45 in response. Replacing the single fragments of the ARC domain in a Rx1 background with the Gpa2 sequences only leads to a clearly autoactive construct when the N-terminal half of ARC2 (fragment 3b) is replaced. However, combining this part of the ARC2 with either the N-terminal or C-terminal flanking fragments (G3ab, G3bc) gives autoactivity as strong as when the full ARC domain is derived from Gpa2 (Fig. 4B and 4C). The complementary construct, where in the G13R45 background only G3b is replaced by R3b, shows no autoactivity at all. The ARC2 region corresponding to fragment 3b (AA 372-440) was therefore concluded to be the part of ARC2 that functionally interacts with the LRR. Within this region the Rx1 and Gpa2 differed in only 7 amino acid positions, and these differences are assumed to be responsible for the observed domain incompatibilities.

The constructs in which the CC (G1R25), the CC-NB (G12R35) or the CC-NB-ARC1 (G12R3aG3bcR45) were derived from Gpa2 and the rest of the protein from Rx1, showed a consistently stronger response to the virulent CP105 than to GFP (Fig. 4B). This response was also stronger than the response of wild type Rx1 to the virulent CP105. For G12R3aG3bcR45 this response was almost as strong as the response to the avirulent CP106. The broader response could be caused by a change in specificity or by an increase of sensitivity of the R protein.

The Gpa2 LRR has to be combined with the Gpa2 ARC2 to gain functionality in chimeric constructs

In R13G45, the CC-NB-ARC of Rx1 is combined with the LRR of Gpa2. The incompatibility between these domains manifests itself as a loss-of-function. This construct shows no autoactivation, and shows no response to nematode elicitor D383-1. We assumed replacing the incompatible region in the CC-NB-ARC part of the construct by the complementary Gpa2 sequence would lead to a reconstitution of the Gpa2 specific elicitor-dependent activation. When in a series of constructs the fragments of the CC-NB-ARC in R13R45 were replaced by Gpa2 sequences we could see a gradual regaining of elicitor-dependent activation in response to D383-1 (Fig. 4D). Replacing only the second half of the ARC2 resulted in a weak elicitor-dependent activation. Replacing the complete ARC2 (fragments 3b and 3c) resulted in an intermediate HR response to the nematode elicitor. When both ARC1 and ARC2 were exchanged for their Gpa2 counterparts, the resulting elicitor-dependent response became indistinguishable from full Gpa2 (R12G35). In a Gpa2 background the exchange of the second half of the ARC2 domain (fragment 3c) by the Rx1 is allowed without a complete loss-of-function. The exchange of the first half of the ARC2 (fragment 3b) by the Rx1 sequence results in a construct that does not show elicitor-dependent activation, nor the Gpa2 specific weak autoactivation (Fig. 4D).

Both sequence exchange experiments, starting from G13R45 or from R13G45, support a model in which the ARC2 and the first repeats of the LRR form a functional unit. In the ARC2 domain the region cooperating with the N-terminal part of the LRR could be pinpointed to a region in which only 7 amino acids differ between Gpa2 and Rx1. These are used as input for computing the NB-ARC to LRR docking.

Docking the Gpa2 NB-ARC and LRR domains

Experimental constraints as input for docking

For determining the most likely interdomain docking for the Gpa2 NB-ARC and LRR structure models the Haddock was used (Dominguez et al., 2003). This software allows the definition of *a priori* ambiguous restraints to guide the docking. Two sets of positions were used as restraints in the docking, because phenotypical evidence showed that they are linked to the functional interaction of the NB-ARC and the LRR, and are therefore likely to be located in the interface between the domains. The first set of positions is derived from the sequence exchange experiments described above. Within the minimal Gpa2 ARC2 region still able to cause constitutive activation if combined *in cis* with the Rx1 LRR, only 7 amino acids differed with the corresponding Rx1 sequence. The five residues located on the ARC2 surface in the structure were chosen as docking restraint for the NB-ARC structure (Fig. 5). The second set consists of the positions of three constitutive activating mutations in the LRR as determined by Bendahmane *et al.* (2002) and Farnham & Baulcombe (2006). The first two (S516, D543) (Fig. 5) are located in the N-terminal part of the LRR which is required to match the ARC2 domain for functionality in the sequence exchange experiments. The 3rd (H738) is located near repeat 9 (Fig. 5), which based on the electrostatic surface potential and variability mapping (Fig. 3) divides the LRR in two halves with varying properties.

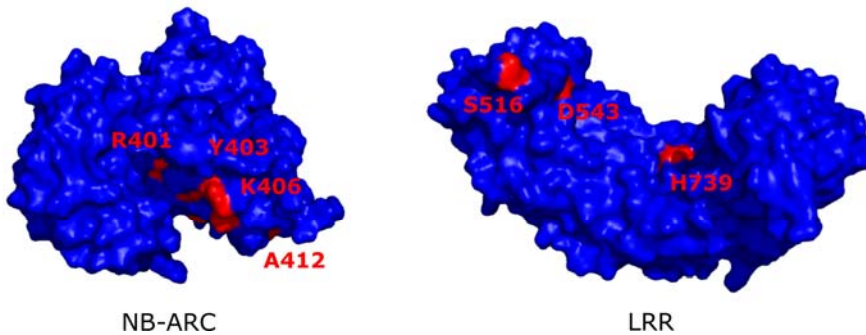
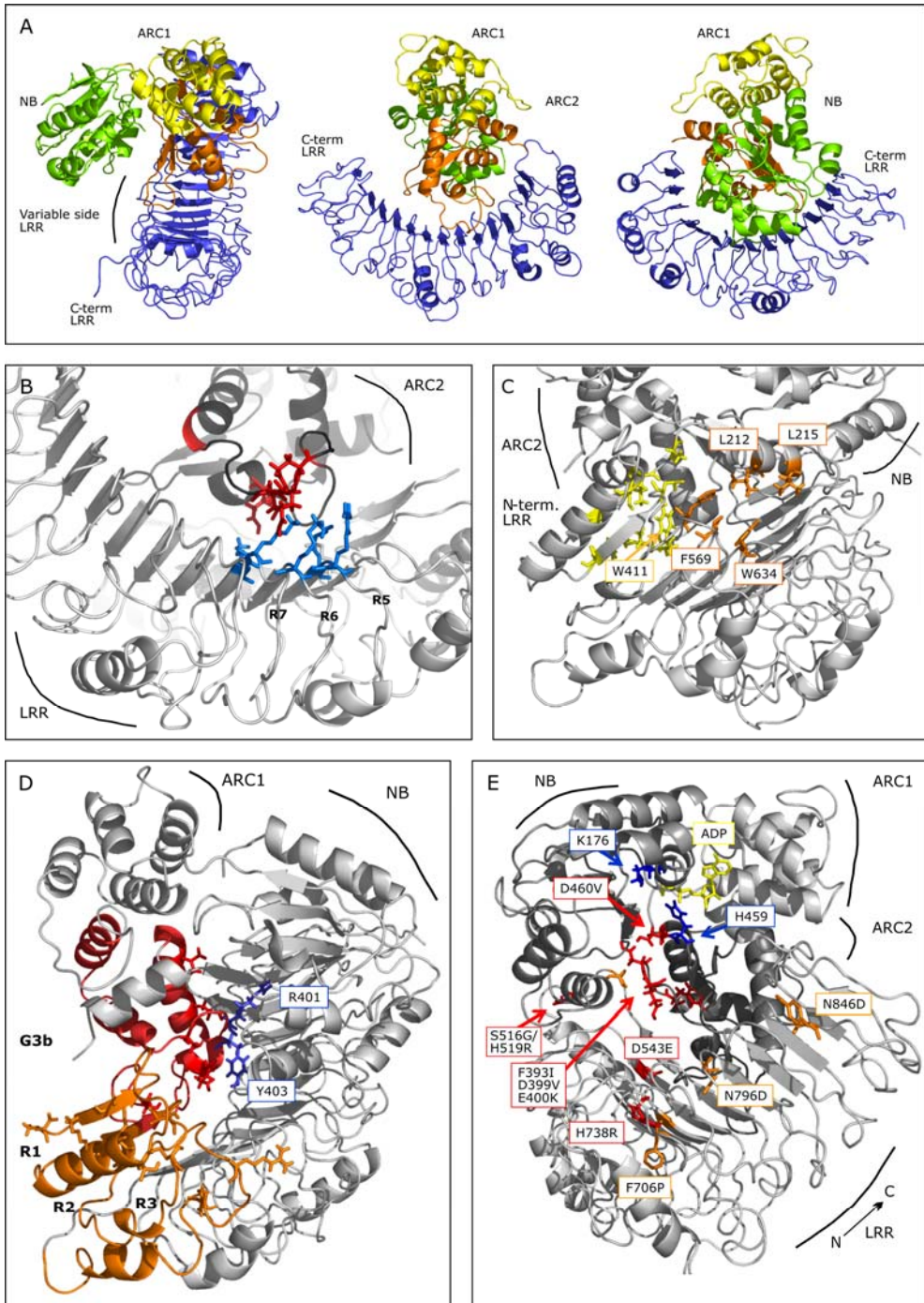


Figure 5. Residues marked in red are defined as active for the AIR restraints table in the docking software Haddock. A397, R401, Y403, K406 and A412 are surface positions differing between Rx1 and Gpa2 in the minimal Gpa2 ARC2 fragment incompatible with the Rx1 LRR in sequence exchange experiments (Fig. 4B and 4C). The Rx1 S516G, D543E and H739R were identified as constitutive active mutations in mutagenesis studies by Bendahmane et al. (2002) and Farnham & Baulcombe (2006).

Figure 6 (right page). **A.** General overview of the docking model with the subdomains coloured (NB green, ARC1 yellow, ARC2 orange, LRR blue). The first view shows the structure from the top looking into the concave side of the LRR. The second view shows the structure from the side of the ARC2-LRR contact. The last view shows the structure from the side where the NB contacts the variable side of the LRR. **B.** The five basic residues in repeat 5 (R5; R628, K629), repeat 6 (R6; K650) and repeat 7 (R7; K675, K676) of the Gpa2 LRR are depicted in blue. The acidic residues in the acidic loop of the ARC2 closest to the basic patch (E414 and EEE 419-421) are depicted in red. Additional acidic residues in the acidic loop region are coloured red in the cartoon. **C.** Close-up view of the B-ARC/LRR docking showing the residues involved in hydrophobic interactions between the NB and LRR (L212, L215, F569, P608, W634, in orange) and between the ARC2 (I395, A397, W411, A412, V413 in yellow) and the LRR (V511, C513, V541, L546 in yellow). **D.** Illustration of the regions in the ARC2 and LRR that are shown by the sequence exchange experiments to form a functional unity. The first three LRR repeats are shown in orange. The minimal incompatible domain the ARC2 is shown in red. At the positions were Rx1 and Gpa2 differ in these regions the residues are shown as sticks. R401 and Y403 are shown in blue. **E.** The position of autoactivating mutations known from Rx1 are shown at the corresponding positions on the Gpa2 NB-ARC/LRR docking model. The positions of the mutations used as constraints in the domain docking are shown in red (S516, D543, H738). Additional mutations (Bendahmane et al., 2002, Farnham & Baulcombe, 2006) are shown in orange.



Docking and description of resulting model

An ensemble of 1,000 structural models for the complex was generated by docking the NB-ARC and LRR domains. From a total of 100 structures left after the refinement stage only two major clusters within a 7.0 Å rms threshold were produced. One of the clusters contained structures with binding interfaces where the linker between the domains should be located, and was therefore discarded. From the other cluster, the best decoy (-119.105 Haddock score) was taken for further analysis. The chosen decoy (Fig. 6A) displaced ~2300 Å² from the solvent accessible area. Detailed interface analysis revealed two major players in the interaction: a hydrophobic component and an electrostatic component. The electrostatic component consists of the opposite attraction between a highly charged cluster (Gpa2 K604, R628, K629, K650, K675, K676 in LRR repeat 5-7) and a highly acidic loop region in the ARC2 subdomain (Gpa2 409- ELWAVEGFLNEEEGKSIEEVAETCINE-435) (Fig. 6B). The hydrophobic amino acids comprise an important area from the interface. L212 and L215 in the NB contact and F569, P608, and W634 in the LRR. I395, A397, W411, A412, and V413 in the ARC2 come in contact with V511, C513, V541, and L546 of the LRR (Fig. 6C).

The docking model of the complex between the two domains matched several properties inferred from the structural models alone. In this docking model, the C-terminal end of the NB-ARC is oriented towards the N-terminal part of the LRR domain (Fig. 6A). Also, the hyper-variable groove between the NB and ARC2 contacts the variable side of the LRR.

Support and validation of the docking model

Coevolution

A complementary bioinformatic analysis was performed on the sequences of a set of Rx1/Gpa2 homologs (Butterbach *et al.*, in preparation). In order to find further support for the domain docking, a coevolution study was performed on the surfaces of the domains according to the Sato method (Sato *et al.*, 2006). Following structure and variability analyses, the LRR domain surface was divided by two planes: a transversal plane passing through the 10th repeat and a sagittal plane (containing the cylinder axis) over the length of the structure. The transversal plane delimits two halves: an N-terminal half and a C-terminal half. The sagittal plane, according to variability mapping divided the surface of the domain into a variable half and a conserved half. These divisions defined four surfaces: N-terminal conserved, N-terminal variable, C-terminal conserved and C-terminal variable (N-LRRc, N-LRRv, C-LRRc and C-LRRv). The NB-ARC domain was divided according to its structural domains: NB, ARC1 and ARC2. Sequences corresponding to the NB, ARC1, ARC2, N-LRRv and N-LRRc surfaces were subjected to coevolution analysis using the Pearson correlation coefficient as a measure (Table 1). The results show that the variation in the conserved part of the N-terminal part of the LRR correlates strongest with the ARC2 surface, and that the variation in the variable part of the LRR N-terminal half correlates most with the variation in the NB domain. These results are in agreement with the interdomain contacts found in the docking model.

Table 1. Pearson correlation coefficient as measure of intradomain coevolution. The NB-ARC was divided in its subdomains (NB, ARC1, ARC2). The N-terminal half of the LRR (repeat 1-9) was divided in two over the sagittal plane in its more conserved side (N-LRRc) and its variable side (N-LRRv). The highest correlation coefficient can be seen for the combination of N-LRRv and NB, and for the combination of N-LRRc and ARC2. This corresponds to the intradomain contacts found in the docking.

	N-LRRv	N-LRRc
NB	<u>0.232</u>	0.074
ARC1	0.211	0.164
ARC2	0.108	<u>0.211</u>

Mutagenesis acidic loop & hydrophobic patch

To assess the importance of the hydrophobic and electrostatic interaction predicted by the docking model the involved sets of residues were altered via targeted mutagenesis in a series of constructs. Two mutations were introduced in the Gpa2 NB-ARC domain to test the importance of the hydrophobic patch on the ARC2 surface: Gpa2 I395N and Gpa2 V413N (Fig. 7A). In both a hydrophobic residue is replaced by the polar asparagine (N). To test the role of the acidic loop, a construct was made in which the three glutamic acid residues (419-EEE-421) in the centre of the acidic loop were replaced by the neutral serines and an alanine (SAS) (Fig. 7A).

The functionality of these constructs was tested both in the context of the full-length Gpa2 constructs and by expressing the CC-NB-ARC domain containing the mutations (G13) *in trans* with the transactivating Rx1 LRR domain containing mutation Y712H (R45 Y712H) (Rairdan and Moffett, 2006). Because functionality *in trans* depends fully on the interaction between the domains, it was expected that mutations influencing the interaction would have a greater impact on the functionality of *in trans* coexpressed domains than on functionality *in cis*. Elicitor-dependent activation of the full-length Gpa2 mutants was only reduced in the case of Gpa2 I395N, but both Gpa2 V413N and Gpa2 EEE-SAS responded to the elicitor D383-1 like the wild-type Gpa2 (Fig. 7B). Coexpression of the mutated Gpa2 CC-NB-ARC domains combined with the Rx1 LRR containing the Y712H mutation resulted in a completely abolished HR response for G13 I395N, a strongly reduced HR response for G13 EEE-SAS and almost no reduction in HR response for G13 V413N when compared to the wild-type G13 construct (Fig. 7C). Expression of Myc-tagged versions of these constructs and subsequent western blotting verified that all mutants were expressed stably and present in the plant cells at similar levels (Fig 7D).

Co-immunoprecipitations of Myc-tagged CC-NB-ARC and HA-tagged LRR domains with anti-Myc antibodies were performed to study if the hydrophobic patch and acidic loop mutations affected the physical interaction between the LRR and CC-NB-ARC domains. Surprisingly, all of the hydrophobic patch and acidic loop mutants could still co-immunoprecipitate the 4HA-LRR domain in proportions not significantly different from the wild-type Gpa2 G13-4Myc (Fig. 7E). Although the loss-of-function caused by the I395N mutation in the *in trans* expressed G13 I395N, combined with only a reduction in response in the full-length Gpa2 I395N seems in line with the phenotype expected for a loss of interaction, the physical interaction as studied by co-immunoprecipitation was not affected.

It is possible that the precise orientation of the domains in relation to each other has changed, without changing the affinity. In addition, there is a possibility that the mutations were not disruptive enough. In the acidic loop only 3 out of 9 acidic residues were changed to neutral residues. In the hydrophobic patch only single positions were changed to polar, but neutral, residues.

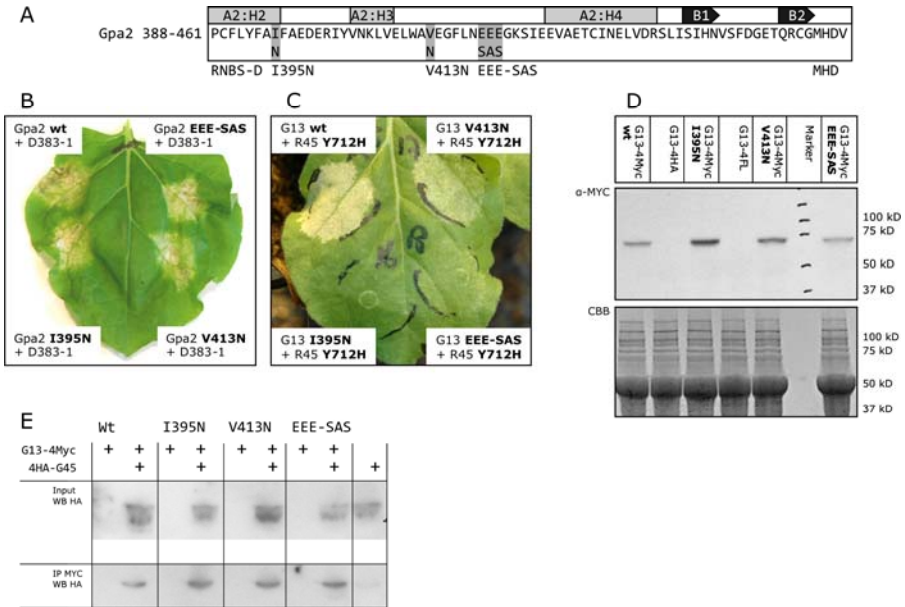


Figure 7. Targeted mutagenesis of positions in the hydrophobic patch and the acidic loop of the ARC2 domain. **A.** The position and identity of the point-mutations in the Gpa2 ARC2 sequence. Above the sequence the secondary structure is indicated (A2:H2, H3, H4; A-helices 2-4 of the ARC2. B1, B2: predicted β -strands preceding the MHD motif. **B.** Full-length Gpa2 with hydrophobic patch and acidic loop mutations tested for activation by the elicitor D383-1 in an agroinfiltration assay. In some repetitions of this experiment Gpa2 I395N would respond to the D383-1 elicitor with a slower response than wild type Gpa2. Image taken 5 days after infiltration. **C.** Assessment of the in trans functionality of the Gpa2 CC-NB-ARC (G13) hydrophobic patch and acidic loop mutants. The CC-NB-ARC of wild type Gpa2 is readily activated in trans by the transactivating Rx1 LRR mutant Y712H (R45 Y712H). Coexpression of the mutant versions of the Gpa2 CC-NB-ARC with R45 Y712H show a full cell death response for wild type Gpa2, and a gradual reduction for G13 V413N and G13 EEE-SAS. G13 I395N does not show any response in this combination. **D.** Anti-Myc immunoblot showing the 4Myc-tagged versions of the Gpa2 CC-NB-ARC hydrophobic patch and acidic loop mutants. The proteins were expressed in leaves for 2 days and leaf protein extract was run on SDS-PAGE. The Coomassie Brilliant Blue stained gel shows equal loading of the samples (CBB). CC-NB-ARC tagged with 4HA or 4FLAG were run as negative controls. No breakdown products of the Myc-tagged products were detected. **E.** Co-immunoprecipitation of 4HA tagged Ga2 LRR (4HA-G45) with MYC-tagged wild type and mutant versions of the Gpa2 CC-NB-ARC (G13-4Myc). Immunoblots with anti-HA of the input and of the results from anti-Myc immunoprecipitation are shown. The 4HA-G45 construct is shown to co-immunoprecipitate with each CC-NB-ARC construct, but not in the absence of CC-NB-ARC protein (last lane).

Mutagenesis Basic patch

The role of the basic patch in the LRR was assessed by a series of mutations altering the basic residues in LRR repeat 5 (R5: R628Q/K629T), repeat 6 (R6: K650T) and repeat 7 (R7: K675Q/K676A) (Fig. 8A). Constructs of the full-length Gpa2 and of the separate LRR domain (G45) were made containing the individual mutated repeats or combinations thereof. The full-length mutant constructs were tested via coexpression with the Gpa2 elicitor D383-1 and the LRR constructs by coexpression with the Gpa2 CC-NB-ARC containing the D460V mutation.

All full-length Gpa2 basic patch mutants (R5, R56, R6, R67, R567) showed an HR when coexpressed with D383-1 indistinguishable from the wild-type Gpa2 (Fig. 8B). However, when the separate LRR constructs containing the basic patch mutations were coexpressed with G13 D460V, only G45 R6 showed an HR response similar to the combination with the wild-type G45 construct. G45 R5, R56, R67 and R567 completely lost the ability to mediate the activation of G13 D460V (Fig. 8C). The difference between the effect on the *in cis* and *in trans* functionality of the basic patch mutation is much more pronounced than the difference seen for the hydrophobic patch and acidic loop mutations. The mutation of the basic residue in repeat 6 (K650T) does not affect *in cis* and *in trans* functionality of Gpa2. To assure that loss-of-function *in trans*, seen for the mutant LRRs was not caused by a loss of protein stability, 4HA-tagged versions of the wild type Gpa2 LRR and the LRR containing all basic patch mutations (G45 R567), were expressed in leaves and detected after SDS-PAGE with anti-HA antibodies. No loss of stability was seen (Fig. 8D).

Mutagenesis R401 / Y403

Based on both the series of detailed NB-ARC sequence exchange constructs starting from the constitutively active G13R45 or the inactive R13G45 we concluded that the amino acid differences between Rx1 and Gpa2 in fragment 3b are the main factors determining the intramolecular incompatibility with the LRR domain. For that reason five of these were used as input in the NB-ARC/LRR docking calculations (Fig. 5).

In sequence exchange fragment 3b the ARC2 domain of Rx1 and Gpa2 differ in only 7 amino acids (Fig. 9A). From these 7 positions differing between Rx1 and Gpa2 two stand out. The amino acid at position 401 is a glutamine in Rx1 and an arginine in Gpa2. In the Gpa2 NB-ARC/LRR docking model this R401 lies in the variable groove between the ARC2 and the NB domain, in close proximity of the nucleotide binding pocket (Fig. 6D). A large and positively charged residue at this position could influence nucleotide binding and the NB-ARC conformational switch.

The second notable difference between Gpa2 and Rx1 is the serine at position 403 in Rx1 which is a tyrosine in Gpa2. In the docking model this Y403 in Gpa2 is positioned at the interacting surface of the ARC2 domain and the β -strands of LRR repeat 6-8 (Fig. 6D). The change from a small serine (van der Waals volume 73 Å³) to a bulky aromatic tyrosine (141 Å³) in the ARC2/LRR interface is likely to affect the interface.

To study if the amino acids at position 401 and 403 in Rx1 and Gpa2 played a role in the domain incompatibility, these position were mutated, either as single changes or the two

combined, from the Gpa2 identity to the Rx1 identity or *vice versa*. The point mutations were made in the background of the constitutively active G13R45 (G13R45 R401Q, Y403S, RY/QS), Gpa2 (Gpa2 R401Q, Y403S, RY/QS), and Rx1 (Rx1 Q401R, S403Y, QS/RY). The resulting constructs were tested for elicitor-dependent activation and elicitor-independent autoactivation in transient agroinfiltration assays (Fig. 9B and C).

The mutations at position 401 and 403 resulted in distinct phenotypes. Both Gpa2 R401Q and RY/QS show a clear reduction in elicitor-dependent activation which was not seen for Gpa2 Y403S (Fig. 9B). G13R45 Y403S and G13R45 RY/QS on the other hand showed a clear reduction in their autoactive response, but not in their elicitor-dependent activation (phenotypes of the G13R45 versions autoactive response shown in Fig. 9C).

However, neither the Q401R and S403Y on their own nor the combination of the two in the Rx1 background resulted in an autoactivation as observed for the Rx1 in which fragment 3b was exchanged the Gpa2 sequence (data not shown). Apparently, additional amino acid residues are required to reproduce this phenotype.

It appears that R401 has a role in Gpa2's elicitor-dependent activation, but not in the autoactive response caused by the incompatibility of the Gpa2 ARC2 and Rx1 LRR. The tyrosine at position 403 on the other hand appears to play a role in the autoactivation of G13R45. This is in agreement with the spatial position both residues have in the Gpa2 docking model, where Y403 lies in the interface between the ARC2 and the LRR whereas R401 lies in the groove between the NB and ARC2 subdomains, close to the nucleotide-binding pocket.

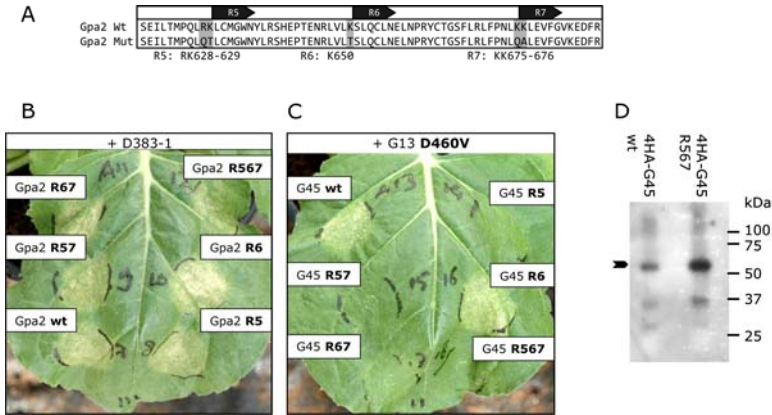


Figure 8. Mutagenesis of the Gpa2 LRR basic patch residues. **A.** Alignment showing the basic patch residues in the sequence of repeat 5, 6 and 7 (R5-R7) and the nature of the point-mutations. **B.** Functionality of full-length Gpa2 containing different combinations of basic patch mutations. Mutations were made in single repeats (R5, R6), combinations of two repeats (R67, R57), and all combined (R567). The mutant constructs were coexpressed via an agroinfiltration assay with the Gpa2 elicitor D383-1. This image was taken 5 days after infiltration. Only Gpa2 R567 consistently showed a weaker HR in this type of experiments. **C.** In trans functionality of the basic patch mutants was tested by coexpression of the mutant Gpa2 LRRs (G45) with the Gpa2 CC-NB-ARC containing the D460V mutation in the MHD motif (G13 D460V). Expressed without LRR this construct does not show activity. Coexpression of the Gpa2 LRR leads to an elicitor-independent HR. Coexpression of G13 D460V with G45 R6 results in an HR indistinguishable from wild-type. Coexpression of any of the other mutant LRRs did not result in an HR. This image was taken 5 days after infiltration. **D.** Anti-HA immunoblot showing the presence of 4HA-tagged mutant Gpa2 LRRs in a leaf extract after 3 days expression.

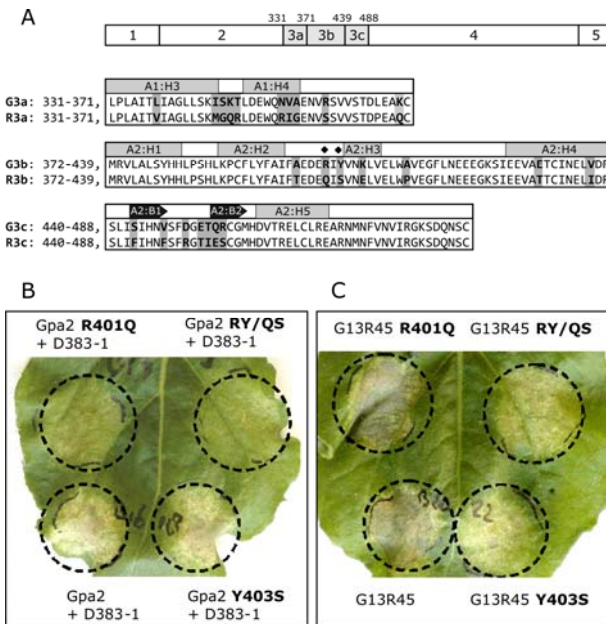


Figure 9. **A.** Alignment of the amino acid sequences of Rx1 and Gpa2 in swap fragments 3a, 3b and 3c. The secondary structure as used in the 3D modelling, is shown in the bar above the sequences. Fragment 3a overlaps the 3rd and 4th α -helices of the ARC1 domain (A1:H3, A1:H4). Gpa2 and Rx1 do not differ in the first two α -helices of the ARC1 domain, but do vary in 10 amino acid positions in the sequence contained by fragment 3a. Fragment 3b consists of the first 4 α -helices of the ARC2 domain (A2:H1-H4). Rx1 and Gpa2 vary in 7 amino acid positions in this sequence. The two positions studied by targeted mutagenesis are marked with stars. Fragment 3c contains the conserved MHD motif, the two β -strands preceding it, and the 5th ARC2 α -helix. In this region Rx1 and Gpa2 differ in 7 amino acid positions. **B.** Agroinfiltration assay of Gpa2 and mutated versions thereof combined with the Gpa2 elicitor D383-1 from *G. pallida*. Gpa2 R401Q and Gpa2 RY/QS show a reduced HR response in comparison with wild type Gpa2 and Gpa2 Y403S. **C.** Agroinfiltration assay of the autoactive G13R45 construct and mutated versions thereof. G13R45 R401Q and G13R45 wild type show a full HR, whereas G13R45 RY/QS and G13R45 Y403S show a reduced HR response.

Discussion

In this study, a structural model is proposed for the interaction between the Gpa2 LRR domain and the ADP-bound conformation of the NB-ARC domain. In our opinion, this model can form a theoretical template for forming new hypotheses and designing the experiments to test them, as no crystal structures have been published for full NB-LRR proteins yet. The individual domain structures, which were both modelled by threading after known homologous structures, revealed that the electrostatic potential and the variability seen among homologous Gpa2/Rx1 proteins are both not evenly distributed over the surface. The surface of the N-terminal half of the LRR has an overall basic charge, whereas the C-terminal half is acidic like observed for the NB-ARC domain. Remarkably, on the NB-ARC the variability is concentrated around the groove between the NB and ARC2 subdomains. On the LRR, the positions immediately C-terminal to the repeat motif were shown to be hypervariable resulting in a polar distribution of the variability on the LRR surface. Repeat 9 differs markedly from the knob like structure predicted for the Rx1 LRR (data not shown) and marks the position in the LRR where the charge is shifted from basic to acidic.

Extensive sequence exchanges between Rx1 and Gpa2 showed in detail which NB-ARC and LRR sequences are functionally interdependent. Five positions identified in this way in the ARC2 and three positions earlier reported as constitutive activating mutations in the Rx1 LRR were applied as restraints in the computational docking. In the final docking model two components appear to be important factors in the interaction; the contact of hydrophobic residues in the NB, ARC2 and the LRR, and the electrostatic interaction between a highly acidic loop in the ARC2 and a highly basic patch on the conserved side of the LRR. Mutational analysis of these features resulted in phenotypes supporting the proposed function in the interaction. However, mutations in the acidic loop and two positions in the ARC2 hydrophobic patch did not block the physical interaction with the LRR. An interdomain coevolution analysis based on a large set of sequences from Rx1/Gpa2 homologs further supported the docking of the ARC2 and LRR domain.

Electrostatic interactions play a role in ARC2/LRR interaction

A interesting result of the 3D modelling of the Gpa2 domains was that it showed clearly that the electrostatic charges are distributed unevenly along the surfaces of the NB-ARC and LRR 3D models, which was not obvious from the amino acids sequence alone. The R protein Bs2 from pepper is a relatively close homolog to Rx1 and Gpa2 (Tai et al., 1999)(Mazourek, et al., in press, Genetics). The Bs2 interdomain interactions have been shown to be similar to those of Rx1, but with the difference that in the presence of the elicitor the interaction between the CC-NB-ARC and LRR was not disrupted (Leister et al., 2005). Comparing the amino acid sequences of Rx1, Gpa2 and Bs2 (AAF09256 in genbank) at the position of the acidic loop and the basic patch shows that both are conserved features. The residues in repeat 5 of Gpa2 (R628-K629) can be aligned to R653 and H654 in Bs2. The Lysine in repeat 6 (K650) is not conserved, which is interesting, because site-directed mutagenesis of this position in Gpa2 did not influence the *in trans* functionality of the LRR. Both the lysines in repeat 7 (K675-K676) are also conserved in Bs2 (K704-K705). The acidic loop is present in Bs2 as well as the concentration of six acidic residues between positions 431 and 446. It will be

interesting to see if other R proteins have similar concentrations of basic residues in the N-terminal half of the LRR and whether these charged surfaces have a conserved role in the NB-LRR structure.

The specific role of electrostatic interactions in protein-protein interactions has been widely studied, but its contribution to inter- and intramolecular R protein interactions is unknown. Many known interactions are stabilised by electrostatic interaction, in addition to the van der Waals interactions, hydrogen bonds and hydrophobic interactions. However, burying charged residues in interaction interfaces require their energetically unfavourable desolvation for the formation of salt bridges, which might even lead to a netto destabilising effect on the interaction (Sheinerman and Honig, 2002; Sulea and Purisima, 2003). It is often observed that instead of being part of the buried protein-protein interface, the oppositely charged residues are more involved in the initial encounter and orientation of the interactors. The formation of an interaction can be described as a two-step process, first a loose encounter complex is formed after which desolvation and the formation of hydrogen bonds lead to a tighter final complex. The Coulombic forces work over larger distances than the forces that eventually stabilise the protein-protein interface. Strong electrostatic attraction therefore heightens the association rate of the interaction, rather than the dissociation rate, which is dominated by the hydrophobicity, van der Waals interactions and hydrogen bonds (Sheinerman et al., 2000; Kiel et al., 2004a).

The interface characteristics match those of experimentally determined protein-protein interfaces

In the full-length NB-LRR protein the linkers connecting the domains offer only a limited amount of free movement. This spatial constraint will strongly boost the association rate in comparison to the situation where the domains are not physically linked, like in the *in trans* coexpression experiments. For that reason we expected the mutations influencing the association of the domains to have a greater phenotypical impact in the *in trans* expressed domains than in the *in cis* expressed domains. Such a difference was exactly what we observed for the basic patch mutations of repeat 5 and repeat 7. *In cis* no effect of mutating the basic patch on the functioning of Gpa2 was observed, whereas *in trans* all functionality was lost. The effect of the acidic loop mutations was smaller than expected; some functionality was retained *in trans*. This might be due to the limited number of residues changed in the acidic loop mutation; only 3 out of 9 acidic residues were replaced by neutral residues. Another factor worth consideration is the CC domain, which was not modelled in this docking study, but which is acidic in nature and might contribute to the interaction between the NB-ARC and LRR.

In Gpa2, several residues were identified that formed complementary hydrophobic patches on the surfaces of the NB, the ARC2 subdomain and the LRR. The amino acid residue composition of the surfaces buried in the interface matches documented protein-protein interfaces. In addition, the total amount of molecular surface of the complex subtracted by the docking ($\sim 2300 \text{ \AA}^2$) is consistent with values observed in empirically documented protein-protein complexes. Residues known to contribute above average to protein-protein interactions are arginine, tryptophan and tyrosine, and to a lesser extend asparagine,

histidine, lysine, isoleucine and proline (Bogan and Thorn, 1998; Kiel et al., 2004b; Moreira et al., 2007). Two tryptophan residues, one proline and one isoleucine were found to contribute to the hydrophobic interaction in Gpa2. One arginine and lysines were found to contribute via the electrostatic interactions. Additional site-directed mutagenesis of positions involved in the interaction could be guided by the predicted contributions these positions make to the binding.

In our site-directed mutagenesis study only V413 and I395 were replaced in separate constructs by the polar Asparagine (N). V413N did not have any effect on Gpa2 functionality *in cis* and not a large effect *in trans*. In the docking it is positioned at the border of the interface, in contrast to I395 which has a more central position in the interface. Accordingly, the I395N mutation had a much stronger effect on the functioning of Gpa2, leading to a partial loss-of-function *in cis* and a complete loss-of-function *in trans*. Surprisingly, no difference could be seen in the interaction between the LRR and CC-NB-ARC as shown by co-immunoprecipitation. Apparently the effect on the interdomain interaction was large enough to affect the functionality of the domains, but not enough to be visualised by co-immunoprecipitation. It could have a specific effect on the interdomain communication, but less on affinity. Similar effects have been seen in other studies (Rairdan and Moffett, 2006; van Ooijen et al., 2008a).

Intramolecular coevolution supports the docking model

Intramolecular coevolution is a sign of functional or structural dependency often in the form of an interdomain interaction (Gloor et al., 2005; Fares and Travers, 2006). Mutations in one position disrupting an interaction with another position can be counterbalanced if the other position changes as well. Especially if driven by functional and structural constraints whole groups of positions can coevolve together (Buck and Atchley, 2005; Gloor et al., 2005). Even coevolution between separate proteins is used in some studies to predict interaction networks (Choi et al., 2005; Halperin et al., 2006). The modular architecture of R proteins implicate that such interdomain coevolution may have occurred, resulting in specific adaptation of the interacting surfaces in CC-NB-ARC-LRR proteins. The correlation seen between varying residues at sides located on the interface between the LRR and the ARC2 or NB in the docking model can therefore be seen as support for the model. For several other R proteins large sets of homologous sequences are available and it would be highly informative to study the interdomain coevolution in these proteins as well, especially if they are supposed to have a similar interdomain docking structure.

The ARC domain and N-terminal part of the LRR are finely tuned to each other in evolution and small changes in either of them can result in autoactivation or inactivation (Bendahmane et al., 2002; Farnham and Baulcombe, 2006; Rairdan and Moffett, 2006). In our experiments we delimited the areas within the ARC and LRR domains that need to be matched to short stretches of amino acids in both domains. The amino acids in the ARC2 domain that need to match the LRR were used as input in the creation of the docking model. In the final docking model the matching area in the LRR turned out to be in close proximity to the ARC2 positions. Interestingly, the protein regions that are required to match, overlap with the coevolving regions detected in the set of Rx1/Gpa2 homologs. Only a limited

number of residues vary between Gpa2 and Rx1 in the minimal incompatible regions. In the minimal autoactive ARC2 fragment only seven amino acids differ, as many as in the bordering ARC1 and ARC2 exchanged fragments. In the first three repeats of the LRR a total of 13 residues have changed (5 in the first repeat, 4 in the second and another 4 in the third)(Fig. 6D).

Sequence exchange and site-directed mutagenesis show distinct roles for the NB-ARC interface and the ARC2-LRR interface in NB-LRR activation

Incompatibility between these regions does not only lead to constitutive activity when the Gpa2 ARC2 is combined with the Rx1 LRR, it also leads to a strongly reduced activity when the ARC2 domain of Rx1 is combined with the LRR of Gpa2. For both scenarios wild type functionality can be reconstituted if the construct is made to contain matching domain sequences from either Gpa2 or Rx1 (Rairdan et al., 2006; Chapter 4, this thesis). This indicates that apparently the ARC domain of Gpa2 is more sensitive to activation or the LRR domain of Rx1 is a more potent activator. If expressed *in trans* the Gpa2 and Rx1 N- and C-terminal domains show a similar asymmetric activation pattern. The LRR of Rx1 can complement the Gpa2 CC-NB-ARC D460V mutant, resulting in an HR. The LRR of Gpa2 on the other hand complements the Rx1 CC-NB-ARC D460V much less efficient than it complements the Gpa2 CC-NB-ARC D460V (data not shown).

The finding that one of the residues differing between Rx1 and Gpa2 specifically influences the elicitor-dependent activation shows that the Gpa2 ARC2 evolved to be more sensitive for activation than the Rx1 ARC2. The predicted position of R401 in the structure, close to the nucleotide binding pocket, might influence the interactions of the surrounding subdomains with the nucleotide and thereby the nucleotide-dependent conformations. This could explain the effect of mutation R401Q on elicitor-dependent activation. The Gpa2 Y403 was shown to have mainly an effect on the autoactivation caused by ARC2/LRR incompatibility. In the Gpa2 docking model this residue is positioned in the interface of the ARC2 and the LRR. According to the model the mutation Y403S would modify the domain interaction, which fits the observed phenotype.

The NB-ARC/LRR binding surface in Gpa2 is similar to the NB-ARC/ARC3 binding surface in Apaf-1

In our docking model the N-terminal seven repeats of the LRR interact with the ARC2 and NB at the fold where the domains are closed around the bound ADP. The position of the charged patches and hydrophobic residues on the ARC2 and LRR suggests that in this model the ARC2-LRR interface contributes more to the final affinity than the NB-LRR interface. In the 2.2 Å resolution crystal structure of the ADP-bound WD40-deleted Apaf-1 (Riedl et al., 2005), the Apaf-1 Helical Domain II (ARC3) contacts the NB and Winged Helix Domain (ARC2) in approximately the same position as the LRR does in the Gpa2 docking model. The interaction of Helical Domain II with the NB domain is disrupted when Apaf-1 assumes its open conformation and the NB, WH and HD2 move away from each other (Riedl et al., 2005; Diemand and Lupas, 2006; Riedl and Salvesen, 2007). Too little is known

about the exact function of the Helical domain II to speculate if there are functional analogies between the ARC3 domain and the N-terminal half of the R protein LRR. In the closed conformation of Apaf-1 the WD40 repeats fold over the rest of the protein (Riedl and Salvessen, 2007) through an interaction with the NB domain and thereby inhibiting oligomerization (Hu et al., 1998; Adrain et al., 1999).

The role of the ARC1 domain in the interaction

In the study by Rairdan *et al.* (Rairdan and Moffett, 2006) it was shown that the ARC1 domain is necessary for the interaction of the Rx1 LRR with the CC-NB-ARC. In our docking model it was not possible with the chosen constraints to bring the LRR and ARC1 domain in close proximity to each other. An explanation might be that the ARC1 and NB can contact the LRR if the NB-ARC has a conformation different from the ADP bound Apaf-1 conformation used in the Gpa2 docking. The deletion of the ARC2 domain could have influenced the orientation of the CC, NB and ARC1 relatively to one other, changing the interaction surface. Mutations of positions on the surface of the Rx1 ARC1 domain (FE307/318AA), did not alter the Rx1 functioning *in cis* or *in trans* and did not disrupt the CC-NB-ARC/LRR interaction (Rairdan and Moffett, 2006). A second explanation for the discrepancy between the co-immunoprecipitation data and our docking model could be that the CC and NB domain contribute more to the interaction than we see now in the docking model and lose this interaction in the absence of the ARC1 domain. The autoactivity shown for the stabilised CC-NB domain (Rairdan et al., 2008) could indicate a change in intermolecular interactions in the absence of the ARC1 domain.

In the same study (Rairdan and Moffett, 2006), it is shown that in a co-immunoprecipitation of the Rx1 CC-NB-ARC with N- and C-terminal deletion constructs of the Rx1 LRR only the relatively small N-terminal region aa473-497 or the C-terminal aa904-937 could be removed without losing the functionality *in trans* and the interdomain interaction. In our docking model the N-terminal half (repeat 1-8) is closely connected to the NB-ARC, but the C-terminal half does not participate in the interaction. A functional differentiation within the LRR is supported by the sequence exchange experiments that separate recognition specificity from interdomain communication / compatibility (Chapter 4 of this thesis). Furthermore, the coevolution data and the overall charge distribution over the LRR seem to point to different roles for N- and C-terminal parts. If the curvature of the LRR is stronger than we predicted, then the C-terminal halve might bind the NB-ARC as well in the conformation shown in the docking model. The CC domain, for which we cannot predict the exact position in the model, could also participate in the interaction with the LRR and explain why the full-length LRR is needed for the interaction. The LRR is necessary for the interaction of the CC to the NB-ARC, and several mutations in the NB-ARC have been shown to disrupt the interaction between the CC and the NB-ARC-LRR (Rairdan et al., 2008).

Conserved motifs and mutations in the interface

Four point-mutations in the Rx1 LRR influencing activation sensitivity or recognition specificity were described in a recent study by Farnham and Baulcombe (Farnham and Baulcombe, 2006). The mutations M1 (N846D) and M2 (N796D) broadened specificity and are located in the C-terminal half of the LRR (Fig. 6E). In our LRR model N796 is part of the LRR scaffold stabilising the turn following the beta-strand in repeat 11. Mutation N796D would lead to a structural change of this repeat. N846 is located on the surface downstream the beta-strand of repeat 13. Both are located in the area that is shown to be involved in recognition by sequence exchange experiments (Rairdan and Moffett, 2006)(Chapter 4, this thesis). M3 (repeat 8) and M4 (repeat 1) are located in the N-terminal half of the LRR and caused sensitized activation or autoactivation. S516G (M4) is located in the interface with the ARC2 domain opposite the first helix downstream of the MHD motif (D460-A470). The position of mutation M3 in repeat 8 places it in the interface with the ARC2 opposite the loop upstream of the MHD motif. The autoactive Y712H mutation in Rx1 (S712 in Gpa2) is located near repeat 8 as well (Rairdan and Moffett, 2006).

In the mutation study by Bendahmane et al. (Bendahmane et al., 2002) several gain-of-function mutations were found in the ARC and LRR domains. Three are grouped together in the RNBS-D motif which forms the second helix of the ARC2 subdomain and is close in space to the first two repeats of the LRR (Fig. 6D). F393I falls within the helix itself. D399 and E400 form a conserved charged linker between the ARC2 helices 2 and 3. The linker between the helices (398-EDE-400) interacts with the NB domain. E400K could either affect the ARC2/NB interaction or the ARC2/LRR interaction. The mutation D460V in the MHD motif probably affects the interaction between the conserved histidine H459 and the bound nucleotide as has been studied in several other R proteins (Howles et al., 2005; van Ooijen et al., 2008b). H519R is located in the first LRR repeat close to S516G described by Farnham *et al.* (2006). D543E changes a highly conserved residue in the second LRR repeat (VLDL), located on the concave surface directly opposite the ARC2 RNBS-D motif in which other gain-of-function mutations were found. H738R is located in repeat 9, close in space to mutation 3 (M3) described by Farnham and contacting the same loop upstream of the MHD motif. Considering the positions of the gain-of-function mutations described in both studies, there appear to be two locations in the interface that can, if changed, cause activation. The first consists of the first conserved repeats in the LRR and the RNSB-D motif in the ARC2. The second contains the inner surface of repeat 8 and 9 (close to the basic patch in repeat 5-7) and the loop region upstream the MHD motif in the ARC2 domain (F449-T453 in Gpa2). Interdomain connections in these regions might play a role in the normal activation of the protein, translating recognition into the conformational change of the NB-ARC that leads to further downstream signalling.

Material and methods

Structure modelling

Sequence analysis comprised of the prediction of intrinsically disorder, interdomain linker propensity, turn-forming propensity, secondary structure, contact forming tendency, and the prediction of motives and patterns. To increase the prediction reliabilities several methods were used in parallel for every type of analysis.

- For *domain delimitations*, the DLP (domain linker prediction) program was used (Miyazaki et al., 2002).

- For *intrinsically disorder prediction*, DisEMBL (Linding et al., 2003), IUPRED (Dosztanyi et al., 2005) and DISOPRED (Ward et al., 2004) were used. Intrinsically disorder predictions methods use three main sequence properties: similarity with known sequences of which 3D coordinates are missing or have high B-factors from structures determined by X-ray crystallography, propensity for making loops or propensity of making contacts. These different approaches (DisEMBL, DISOPRED2, IUPRED) gave similar results – the domain has a high probability to be structured.

- For *secondary structure prediction* the best ranked methods by CASP2004 were used: GOR IV (Garnier et al., 1996), Jpred (Cuff et al., 1998), nnPredict (McClelland and Rumelhart, 1988), Porter (Pollastri and McLysaght, 2005), PSA (Stultz et al., 1993), PSIPRED (Jones, 1999), SOPMA (Geourjon and Deleage, 1995), SCRATCH programs (also used for contact and accessibility predictions)(Cheng et al., 2005).

- For *turn prediction*, BETATPRED2 (Kaur and Raghava, 2003) and COUDES (Fuchs and Alix, 2005) programs were used.

Molecular modeling including side chain and constrained *ab initio* loop generation, and the simulations for stability tests were performed using MODELER, Homology and Discover modules of the InsightIII platform from Accelrys (Accelrys Software Inc., San Diego).

Docking was performed using HADDOCK due to its facilities for using constraints and the good results reached with it in CAPRI experiments (Dominguez et al., 2003; van Dijk et al., 2005). Images of structures were created using PyMol, www.pymol.org.

Construction of expression vectors

All constructs used in transient expression assays were cloned in the pBINPLUS binary vector system (van Engelen et al., 1995). The transcription of the constructs used in this study was controlled by the CaMV 35S promoter and Tnos terminator sequences as described earlier (Chapter 4, this thesis). Primary full-length Gpa2 and Rx1 constructs and the construction of the sequence exchange constructs based on the ApaLI site (corresponding to AA position 489), ClaI (AA position 593), BspEI (Gpa2 761 or Rx1 756), and EcoRI (Gpa2 879 or Rx1 874) were described before (Chapter 4, this thesis). Constructs harboring expression cassettes with the Rx1 elicitors CP106 (avirulent) and CP105 (virulent) have been described in Chapter 4 and 6 of this thesis, and constructs harboring the Gpa2 elicitors D383-1 (eliciting) and Rook4 (non-eliciting) were described by Sacco et al., *submitted*.

Domain swaps

To construct the additional chimeric sequences, synthetic gene fragments were introduced in the Gpa2 and Rx1 coding sequencing. Thereby, unwanted restriction sites were removed and new unique sites were added. Care was taken not to change the encoded amino acid sequence. The introns in the Gpa2 and Rx1 coding sequences were retained. The following changes were made in the Gpa2 sequence (A in ATG is 1): T1179C, A1182G, T1228C, A1230C, T1320C, C1321T, C1335T, T1377C, G1383C, A1498T, G1499C, C1501A, G1503A, T1554G, C1563T, G1596C, A1559G, A1623G, C1659A, A1695C, T1696A, C1697G, T1698C, T1749C, A1782T, A1812G, C1818T, T1833C, G1857C, G2052A, T2172A, T2175C, A2244T, T2247G, C2379A, T2403C, C2406G, A2439G, T2568G, A2571G.

In the Rx1 coding sequence the following nucleotides were altered: A990G, T1179C, A1182G, T1228C, A1230C, T1377C, G1383C, A1498T, G1499C, C1501A, G1503A, T1554G, A1584G, G196C, A1599G, A1623G, T1749C, A1782T, A1812G, G1887T, A1971G, A2106G, T2172A, T2175C, T2244G, A2313G, T2388C, C2391G, C2394T, A2424G, T2553G, A2556G.

Sequences were exchanged at the restriction sites indicated in Figure 4A.

Site-directed mutagenesis

Site-directed mutagenesis was attained by inserting small sections of synthetic DNA with specific codons. In Gpa2 mutation I395N was made by T1184A. Mutation V413N was made by changing the codon GTA at 1237 to AAC. The amino acids 419-EEE in the acidic loop were replaced by SAS via the change of the coding sequence (1255-2263) from GAAGAAGAG to AGCGCTTCT. The mutation R401Q was made by changing 1201-CGG to CAG and Y403S by changing 1207-TAT to TCT. In the targeted mutagenesis of the Gpa2 basic patch the following changes were made in the Gpa2 coding sequence: Gpa2 R5 (RK628-629QT) was made by changing 1882-AGGAAG into CAGACT, Gpa2 R6 (K650T) by changing 1948AAA into ACA, Gpa2 R7 (KK675-676-QA) by changing 2023AAGAAG into CAGGCT. In the Rx1 coding sequence the mutation Q401R was made by changing 1201-CAG to CGG and the mutation S403Y was made by changing 1207-TCT to TAT. The mutation Y712H was created by changing the corresponding codon from TAT to CAT.

All DNA constructs were sequenced.

Protein tags

Multimeric c-Myc (EQKLISEEDL) and HA (YPYDVPDYA) tags were fused to the C-terminus of CC-NB-ARC constructs via a NotI site creating a short linker consisting of three alanines. Fusions to the N-terminus of the LRR were made through a NcoI site overlapping the start-codon in this construct with in between the 4HA tag and the LRR sequence a short linker (TSS).

Transient expression assays

For agroinfiltration experiments *Agrobacterium tumefaciens* strain pMOG101 was cultured in YEB medium (1L YEB: 5 g. beef extract, 1 g. yeast extract, 5 g. bactopectone, 5 g. sucrose, 2 mL MgSO₄) with the appropriate antibiotics as described earlier (Van der Hoorn et al., 2000). After growing overnight at 28°C the bacteria were pelleted by centrifugation and resuspended in infiltration medium (MMAi, 1L: 5 g. Murashi-Skoog salts, 1.95 g. MES, 20 g. sucrose, pH 5.6 with NaOH, 200 µM acetosyringone) and incubated at room temperature for 2 hrs. For agroinfiltration the bacterial suspensions were diluted to final concentrations

between OD600 0.2 and 1.0. Leaves were infiltrated of 6 weeks old *Nicotiana benthamiana* plants grown in the greenhouse at 20°C and 16 hours of light. Each combination was tested at least *in triplo* on two different plants in at least three independent experiments.

Protein methods

Total protein extract of transient transformed *N. benthamiana* leaves was made by grinding leaf material in protein extraction buffer (50 mM Tris-HCl pH7.5, 10% glycerol, 150 mM NaCl, 1 mM EDTA, 2% polyclar-AT PVPP, 0.4 mg/ml Pefabloc SC plus (Roche), 5 mM DTT) on ice. For immunoprecipitation the total protein extract was first passed over a sephadex G-25 column. The protein extract was precleared with rabbit-IgG agarose (40 µl slurry per mL protein extract). After preclearing the protein extract was mixed with 25 µl anti-Myc agarose beads (Sigma) and incubated for 2 hours at 4°C. After 6x washing (washing buffer: protein extraction buffer with 0.15% Igepal CA-630) the agarose beads were resuspended in laemli buffer and the bound protein was separated by SDS-PAGE and blotted on nitrocellulose. For detection on Western blot the following antibodies were used: 9E10 anti-Myc (Sigma) and 3F10 anti-HA (Roche).

References

- Adrain, C., Slee, E.A., Harte, M.T. and Martin, S.J.** (1999). Regulation of apoptotic protease activating factor-1 oligomerization and apoptosis by the WD-40 repeat region. *J. Biol. Chem.* **274**, 20855-20860.
- Albrecht, M. and Takken, F.L.** (2006). Update on the domain architectures of NLRs and R proteins. *Biochem Biophys Res Commun* **339**, 459-462.
- Ammelburg, M., Frickey, T. and Lupas, A.N.** (2006). Classification of AAA+ proteins. *J Struct Biol* **26**, 26.
- Basu, M.K., Carmel, L., Rogozin, I.B. and Koonin, E.V.** (2008). Evolution of protein domain promiscuity in eukaryotes. *Genome Res* **18**, 449-461.
- Bell, J.K., Botos, I., Hall, P.R., Askins, J., Shiloach, J., Segal, D.M. and Davies, D.R.** (2005). The molecular structure of the Toll-like receptor 3 ligand-binding domain. *Proc Natl Acad Sci U S A* **102**, 10976-10980.
- Bendahmane, A., Farnham, G., Moffett, P. and Baulcombe, D.C.** (2002). Constitutive gain-of-function mutants in a nucleotide binding site-leucine rich repeat protein encoded at the Rx locus of potato. *Plant J.* **32**, 195-204.
- Bendahmane, A., Kohm, B.A., Dedi, C. and Baulcombe, D.C.** (1995). The coat protein of potato virus X is a strain-specific elicitor of Rx1-mediated virus resistance in potato. *Plant J.* **8**, 933-941.
- Bogan, A.A. and Thorn, K.S.** (1998). Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* **280**, 1-9.
- Buck, M.J. and Atchley, W.R.** (2005). Networks of coevolving sites in structural and functional domains of serpin proteins. *Mol Biol Evol* **22**, 1627-1634.
- Cheng, J., Randall, A.Z., Sweredoski, M.J. and Baldi, P.** (2005). SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Res* **33**, W72-76.
- Choi, S.S., Li, W.M. and Lahn, B.T.** (2005). Robust signals of coevolution of interacting residues in mammalian proteomes identified by phylogeny-aided structural analysis. *Nature Genet.* **37**, 1367-1371.
- Cuff, J.A., Clamp, M.E., Siddiqui, A.S., Finlay, M. and Barton, G.J.** (1998). JPred: a consensus secondary structure prediction server. *Bioinformatics* **14**, 892-893.
- Diemand, A.V. and Lupas, A.N.** (2006). Modeling AAA+ ring complexes from monomeric structures. *J Struct Biol* **156**, 230-243.
- Dominguez, C., Boelens, R. and Bonvin, A.M.** (2003). HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* **125**, 1731-1737.
- Dosztanyi, Z., Csizmok, V., Tompa, P. and Simon, I.** (2005). IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* **21**, 3433-3434.
- Fares, M.A. and Travers, S.A.A.** (2006). A novel method for detecting intramolecular coevolution: Adding a further dimension to selective constraints analyses. *Genetics* **173**, 9-23.
- Farnham, G. and Baulcombe, D.C.** (2006). Artificial evolution extends the spectrum of viruses that are targeted by a disease-resistance gene from potato. *Proc Natl Acad Sci U S A* **103**, 18828-18833.
- Fuchs, P.F. and Alix, A.J.** (2005). High accuracy prediction of beta-turns and their types using propensities and multiple alignments. *Proteins* **59**, 828-839.
- Garnier, J., Gibrat, J.F. and Robson, B.** (1996). GOR method for predicting protein secondary structure from amino acid sequence. *Methods Enzymol* **266**, 540-553.
- Geourjon, C. and Deleage, G.** (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Comput Appl Biosci* **11**, 681-684.
- Gloor, G.B., Martin, L.C., Wahl, L.M. and Dunn, S.D.** (2005). Mutual information in protein multiple sequence alignments reveals two classes of coevolving positions. *Biochemistry* **44**, 7156-7165.
- Halperin, I., Wolfson, H. and Nussinov, R.** (2006). Correlated mutations: Advances and limitations. A study on fusion proteins and on the cohesin-dockerin families. *Proteins-Structure Function and Bioinformatics* **63**, 832-845.
- Hillig, R.C., Renault, L., Vetter, I.R., Drell, T.t., Wittinghofer, A. and Becker, J.** (1999). The crystal structure of rna1p: a new fold for a GTPase-activating protein. *Mol Cell* **3**, 781-791.
- Howles, P., Lawrence, G., Finnegan, J., McFadden, H., Ayliffe, M., Dodds, P. and Ellis, J.** (2005). Autoactive alleles of the flax L6 rust resistance gene induce non-race-specific rust resistance associated with the hypersensitive response. *Mol. Plant-Microbe Interact.* **18**, 570-582.
- Hu, Y., Ding, L., Spencer, D.M. and Nunez, G.** (1998). WD-40 Repeat Region Regulates Apaf-1 Self-association and Procaspase-9 Activation. *J. Biol. Chem.* **273**, 33489-33494.
- Jones, D.T.** (1999). Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* **292**, 195-202.
- Kaur, H. and Raghava, G.P.** (2003). Prediction of beta-turns in proteins from multiple alignment using neural network. *Protein Sci* **12**, 627-634.
- Kiel, C., Selzer, T., Shaul, Y., Schreiber, G. and Herrmann, C.** (2004a). Electrostatically optimized Ras-binding Ral guanine dissociation stimulator mutants increase the rate of association by stabilizing the encounter complex. *Proc Natl Acad Sci U S A* **101**, 9223-9228.
- Kiel, C., Serrano, L. and Herrmann, C.** (2004b). A Detailed Thermodynamic Analysis of Ras/Effectors Complex Interfaces. *J. Mol. Biol.* **340**, 1039-1058.
- Koonin, E.V., Wolf, Y.I. and Karev, G.P.** (2002). The structure of the protein universe and genome evolution. *Nature* **420**, 218-223.

- Leipe, D.D., Koonin, E.V. and Aravind, L.** (2004). STAND, a class of P-loop NTPases including animal and plant regulators of programmed cell death: Multiple, complex domain architectures, unusual phyletic patterns, and evolution by horizontal gene transfer. *J. Mol. Biol.* **343**, 1-28.
- Leister, R.T., Dahlbeck, D., Day, B., Li, Y., Chesnokova, O. and Staskawicz, B.J.** (2005). Molecular genetic evidence for the role of SGT1 in the intramolecular complementation of Bs2 protein activity in *Nicotiana benthamiana*. *Plant Cell* **17**, 1268-1278.
- Linding, R., Jensen, L.J., Diella, F., Bork, P., Gibson, T.J. and Russell, R.B.** (2003). Protein disorder prediction: implications for structural proteomics. *Structure* **11**, 1453-1459.
- McClelland, J.L. and Rumelhart, D.E.** (1988). *Explorations in Parallel Distributed Processing* (Cambridge, MA: MIT Press), pp. 318-362.
- McHale, L., Tan, X., Koehl, P. and Michelmore, R.W.** (2006). Plant NBS-LRR proteins: adaptable guards. *Genome Biol* **7**, 212.
- Miyazaki, S., Kuroda, Y. and Yokoyama, S.** (2002). Characterization and prediction of linker sequences of multi-domain proteins by a neural network. *J Struct Funct Genomics* **2**, 37-51.
- Moffett, P., Farnham, G., Peart, J. and Baulcombe, D.C.** (2002). Interaction between domains of a plant NBS-LRR protein in disease resistance-related cell death. *Embo J.* **21**, 4511-4519.
- Moreira, I.S., Fernandes, P.A. and Ramos, M.J.** (2007). Hot spots - A review of the protein-protein interface determinant amino-acid residues. *Proteins: Structure, Function, and Bioinformatics* **68**, 803-812.
- Pollastri, G. and McLysaght, A.** (2005). Porter: a new, accurate server for protein secondary structure prediction. *Bioinformatics* **21**, 1719-1720.
- Rairdan, G. and Moffett, P.** (2007). Brothers in arms? Common and contrasting themes in pathogen perception by plant NB-LRR and animal NACHT-LRR proteins. *Microbes and Infection* **9**, 677-686.
- Rairdan, G.J., Collier, S.M., Sacco, M.A., Baldwin, T.T., Boetrich, T. and Moffett, P.** (2008). The Coiled-Coil and Nucleotide Binding Domains of the Potato Rx Disease Resistance Protein Function in Pathogen Recognition and Signaling. *Plant Cell* **20**, 739-751.
- Rairdan, G.J. and Moffett, P.** (2006). Distinct Domains in the ARC Region of the Potato Resistance Protein Rx Mediate LRR Binding and Inhibition of Activation. *Plant Cell* **18**, 2082-2093.
- Riedl, S.J., Li, W.Y., Chao, Y., Schwarzenbacher, R. and Shi, Y.G.** (2005). Structure of the apoptotic protease-activating factor 1 bound to ADP. *Nature* **434**, 926-933.
- Riedl, S.J. and Salvesen, G.S.** (2007). The apoptosome: signalling platform of cell death. *Nat Rev Mol Cell Biol* **8**, 405-413.
- Sacco, M.A., Koropacka, K., Grenier, E., Jaubert, M.J., Blanchard, A., Goverse, A., Smant, G. and Moffett, P.** (2009). The Cyst Nematode SPRYSEC Protein RBP-1 Elicits Gpa2- and RanGAP2-Dependent Plant Cell Death. *PLoS Pathogen* **5**, e1000564.
- Sato, T., Yamanishi, Y., Horimoto, K., Kanehisa, M. and Toh, H.** (2006). Partial correlation coefficient between distance matrices as a new indicator of protein-protein interactions. *Bioinformatics* **22**, 2488-2492.
- Scott, P.G., McEwan, P.A., Dodd, C.M., Bergmann, E.M., Bishop, P.N. and Bella, J.** (2004). Crystal structure of the dimeric protein core of decorin, the archetypal small leucine-rich repeat proteoglycan. *Proc Natl Acad Sci U S A* **101**, 15633-15638.
- Sheinerman, F.B. and Honig, B.** (2002). On the role of electrostatic interactions in the design of protein-protein interfaces. *J Mol Biol* **318**, 161-177.
- Sheinerman, F.B., Norel, R. and Honig, B.** (2000). Electrostatic aspects of protein-protein interactions. *Curr Opin Struct Biol* **10**, 153-159.
- Stultz, C.M., White, J.V. and Smith, T.F.** (1993). Structural analysis based on state-space modeling. *Protein Sci* **2**, 305-314.
- Sulea, T. and Purisima, E.O.** (2003). Profiling Charge Complementarity and Selectivity for Binding at the Protein Surface. *Biophys. J.* **84**, 2883-2896.
- Tai, T.H., Dahlbeck, D., Clark, E.T., Gajiwala, P., Pasion, R., Whalen, M.C., Stall, R.E. and Staskawicz, B.J.** (1999). Expression of the Bs2 pepper gene confers resistance to bacterial spot disease in tomato. *Proc Natl Acad Sci U S A* **96**, 14153-14158.
- Takken, F.L., Albrecht, M. and Tameling, W.I.** (2006). Resistance proteins: molecular switches of plant defence. *Curr Opin Plant Biol* **9**, 383-390.
- Takken, F.L.W. and Tameling, W.I.L.** (2009). To Nibble at Plant Resistance Proteins. *Science* **324**, 744-746.
- Tameling, W.I.L., Vossen, J.H., Albrecht, M., Lengauer, T., Berden, J.A., Haring, M.A., Cornelissen, B.J.C. and Takken, F.L.W.** (2006). Mutations in the NB-ARC Domain of I-2 That Impair ATP Hydrolysis Cause Autoactivation. *Plant Physiol.* **140**, 1233-1245.
- Thompson, J.D., Higgins, D.G. and Gibson, T.J.** (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673-4680.
- van der Biezen, E.A. and Jones, J.D.G.** (1998). The NB-ARC domain: A novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. *Curr. Biol.* **8**, R226-R227.
- Van der Hoorn, R.A.L., Laurent, F., Roth, R. and De Wit, P.** (2000). Agroinfiltration is a versatile tool that facilitates comparative analyses of Avr9/Cf-9-induced and Avr4/Cf-4-induced necrosis. *Mol. Plant-Microbe Interact.* **13**, 439-446.
- van der Vossen, E.A.G., van der Voort, J., Kanyuka, K., Bendahmane, A., Sandbrink, H., Baulcombe, D.C., Bakker, J., Stiekema, W.J. and Klein-Lankhorst, R.M.** (2000). Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: a virus and a nematode. *Plant J.* **23**, 567-576.

van Dijk, A.D., de Vries, S.J., Dominguez, C., Chen, H., Zhou, H.X. and Bonvin, A.M. (2005). Data-driven docking: HADDOCK's adventures in CAPRI. *Proteins* **60**, 232-238.

van Engelen, F.A., Molthoff, J.W., Conner, A.J., Nap, J.P., Pereira, A. and Stiekema, W.J. (1995). pBINPLUS: an improved plant transformation vector based on pBIN19. *Transgenic Res* **4**, 288-290.

van Ooijen, G., Mayr, G., Albrecht, M., Cornelissen, B.J.C. and Takken, F.L.W. (2008a). Transcomplementation, but not Physical Association of the CC-NB-ARC and LRR Domains of Tomato R Protein Mi-1.2 is Altered by Mutations in the ARC2 Subdomain. *Mol Plant* **1**, 401-410.

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

Ward, J.J., McGuffin, L.J., Bryson, K., Buxton, B.F. and Jones, D.T. (2004). The DISOPRED server for the prediction of protein disorder. *Bioinformatics* **20**, 2138-2139.

Zhang, Y. (2009). Protein structure prediction: when is it useful? *Curr. Opin. Struct. Biol.* **19**, 145-155.

Chapter 6

Nuclear localization of the NB-LRR Resistance protein Rx1 is dependent on coiled-coil domain, ATP-binding, SGT1 and Rar1, and its activation is triggered in the cytoplasm

Erik Sloatweg¹, Jan Roosien¹, Laurentiu N. Spiridon⁴, Andrei-Jose Petrescu⁴, Casper van Schaik¹, Robert Dees², Jan Willem Borst³, Geert Smant¹, Arjen Schots², Jaap Bakker¹, Aska Govere¹

¹*Lab. of Nematology, Department of Plant Sciences, Wageningen University, the Netherlands*

²*Lab. of Molecular Recognition and Antibody Technology, Department of Plant Sciences, Wageningen University, the Netherlands*

³*Lab. of Biochemistry, Department of Agrotechnology and Food Sciences, Wageningen University, the Netherlands*

⁴*Institute of Biochemistry of the Romanian Academy, Bucharest, Romania*

To be submitted

Abstract

The Rx1 protein, as many resistance proteins of the NB-LRR class, is predicted to be cytoplasmic, because it lacks discernable nuclear targeting signals. Here we demonstrate that Rx1, which confers extreme resistance in potato (*Solanum tuberosum*) to *Potato virus X* (PVX), is located both in the nucleus and cytoplasm. The CC domain was found to be required for accumulation of the Rx1 protein in the nucleus, while the LRR domain promoted the localization in the cytoplasm. Analyses of fragments of the CC domain revealed no autonomous signals responsible for active nuclear import. Fluorescence recovery after photobleaching (FRAP) indicated that the CC domain binds transiently to large complexes in the nucleus. Disruption of the Rx1 resistance function by mutating the ATP-binding P-loop in the NB domain, or by silencing the co-chaperones SGT1 and RAR1, impaired the accumulation of the full Rx1 protein in the nucleus, while no effects were observed in Rx1 versions lacking the LRR domain. Manipulating the nucleocytoplasmic distributions of Rx1 and its elicitor revealed that Rx1 is activated in the cytoplasm and can not be activated in the nucleus. Our results support a model in which interdomain interactions and folding states determine the nucleocytoplasmic distribution of Rx1.

Introduction

Disease resistance (R) proteins are the central actors in the cell-based innate immune system of plants. They are highly specific in the detection of certain pathogens and can initiate an array of defense responses to prevent further spreading, often culminating in self-destruction of the attacked cell (Martin et al., 2003). In addition, other immune receptors sense highly conserved pathogen-associated molecular patterns (PAMPs), like flagellin, and are thought to represent a more ancient immune system (Navarro et al., 2004; Zipfel et al., 2004; Heese et al., 2007). Pathogen specific R proteins may have evolved to recognize pathogen effectors that neutralize the weak immune responses triggered by PAMPs (Jones and Dangl, 2006). R proteins sense the pathogen either directly or indirectly via host factors that have been modified by the pathogen (Jia et al., 2000; de Wit, 2002; Deslandes et al., 2003; Jones and Takemoto, 2004; Dodds et al., 2006).

Members of the most abundant class of R proteins, the intracellular NB-LRR class, consist of a carboxy-terminal leucine-rich repeat (LRR) with pathogen recognition as primary function, a central nucleotide-binding domain (NB-ARC), and at the N-terminus mostly either a Toll-Interleukin Receptor-like domain (TIR) or a putative coiled-coil (CC) domain. The N-terminal domains are thought to be the signaling modules, but the exact mechanism for signal transduction has yet to be resolved. Some R proteins have been shown to oligomerise (N, RPS5) (Mestre and Baulcombe, 2006; Ade et al., 2007), and ATP-binding and ATPase activity has been observed for the R proteins Mi1-2 and I-2 (Tameling et al., 2002). Studying the intramolecular interactions in the absence and presence of an specific elicitor shows that in its inactive state the R protein must have a compact form with the CC interacting with the NB-ARC-LRR domains and the LRR interacting with the CC-NB-ARC domains (Moffett et al., 2002; Leister et al., 2005; Rairdan and Moffett, 2006). The presence of the elicitor disrupts the interaction between the LRR and the N-terminal half of the protein, which is likely a first step towards an activate conformation (Moffett et al., 2002; Ueda et al., 2006).

Recent studies have placed the nucleus in the centre of attention for plant disease resistance signaling. Several R proteins, including N, MLA and RPS4, have a nucleocytoplasmic localization, and their nuclear localization is required for proper functioning (Burch-Smith et al., 2007; Shen et al., 2007; Wirthmueller et al., 2007). However, the underlying mechanism determining the distribution of the R proteins between the subcellular compartments is not well understood. Large proteins like R proteins cannot diffuse freely from one compartment to another and need to pass through selective pores to enter organelles like the nucleus, the mitochondria, the vacuole or the endoplasmic reticulum. Complex mechanisms have evolved to shuttle proteins between the cellular compartments. Trafficking between the cytoplasm and the nucleus is coordinated around the nuclear pore complexes via nuclear import and nuclear export receptors and the small GTPase Ran (Alber et al., 2007; Cook et al., 2007). The classical monopartite or bipartite nuclear localization signals (NLS) are well defined and consist of short stretches of basic residues (3-6 Lys/Arg) which interact with specific binding surfaces on the import receptor importin α , which in turn forms a heterodimer with importin β (Gorlich and Kutay, 1999). Nuclear export is directed via Leucine-rich Nuclear Export Signals (NES), which interact with the nuclear export receptor (exportin) (Haasen et al., 1999; Hutten and Kehlenbach, 2007). However, many R proteins of the NB-LRR class, including N and MLA, lack such discernable localization signals and therefore until recently it was assumed that R proteins are localized in the cytoplasm.

The first R protein shown to have a nuclear localization was the Arabidopsis RRS1-R, a chimera of TIR-NB-LRR protein and a WRKY-type transcription factor (Lahaye, 2002; Deslandes et al., 2003). RRS1-R interacts with its elicitor, the *Ralstonia solanacearum* effector PopP2 in the nucleus. According to the Rosetta stone principle (Eisenberg et al., 2000; Enright and Ouzounis, 2001), the existence of such a chimeric protein is an indication that in other instances an interaction with transcription factors is part of the R protein signaling mechanism. The discovery of the interaction between the barley MLA proteins and the WRKY1 and WRKY2 transcription factors gave evidence that there is indeed a close link between R proteins and transcriptional regulation (Shen et al., 2007). In the presence of the *Blumeria graminis* elicitor Avr10 the R protein MLA10 interacts with the WRKY2 transcription factor in the nucleus. In the same study WRKY1 and WRKY2 were shown to be suppressors of basal defense. MLA may activate the resistance response by lifting this suppression. The tobacco TIR-NB-LRR R protein N, conferring resistance against TMV, also has a nuclear localization and binds squamosa promoter-like (SPL) transcription factors via its LRR domain, whereas its N-terminal TIR domain indirectly binds the viral helicase p50 (Liu et al., 2004; Shen and Schulze-Lefert, 2007; Caplan et al., 2008a; Caplan et al., 2008b). The finding that R proteins themselves are able to enter the nucleus indicates that the pathway between R protein activation and the downstream transcriptional reprogramming may contain less components than originally expected (Burch-Smith et al., 2007; Shen et al., 2007; Wirthmueller et al., 2007).

The potato Rx1 protein confers resistance to most Potato Virus X strains and has proven to be a valuable model for understanding R protein functioning. The resistance response it triggers after recognizing the virus coat protein is fast and under normal circumstances does not require the induction of cell death to stop virus replication (Bendahmane et al., 1999). The domains of the protein function together via several intramolecular interactions.

Recognition of the coat protein has been linked to a disruption of the interaction between the LRR and the NB-ARC domain (Moffett et al., 2002). The MHD motif in the ARC2 domain, in which substitutions can lead to constitutive activity, is thought to act as a sensor integrating the recognition-mediated conformation changes and the nucleotide-binding state of the protein. It is, however, clear that although the interaction between the NB-ARC domains and the LRR are relatively aspecific, the N-terminal halve of the LRR and the bordering ARC2 subdomain need to be well tuned to each other to give a functional protein (Rairdan and Moffett, 2006). Incompatibility between these domains can lead to constitutive activity, indicating that in the wild type protein their interaction has an autoinhibitory function. Recently, two studies independently showed that the CC domain interacts with a Ran GTPase Activating Protein (RanGAP2) (Sacco et al., 2007; Taming and Baulcombe, 2007).

Here we studied the subcellular localization of Rx1 and the contribution of its functional domains to the subcellular localization pattern. As with many R proteins of the NB-LRR family, a cytoplasmic localization of Rx1 is predicted, because no classical linear nuclear localization signals can be found in the sequence and the size exceeds the limit for passive transport into the nucleus. However, fusions with fluorescent proteins showed that Rx1 is localized in both the cytoplasm and the nucleus. Expression of Rx1 domains and deletion constructs showed that the CC domain is predominantly localized in the nucleus and is required for the accumulation of the full Rx1 protein in the nucleus. Further analyzing various fragments of the CC domain revealed no autonomous linear sequences responsible for nuclear translocation of the Rx1 protein. Photobleaching experiments revealed that the nuclear accumulation of the CC protein appears to be caused by transient interactions with immobile components in the nucleus and not by strong active nuclear import signals. To differentiate between the function of the nuclear and cytoplasmic compartments in elicitor recognition and Rx-mediated signaling, both the elicitor and Rx1 were redirected to either compartment using exogenous targeting signals. Depletion of the elicitor concentration in the cytoplasm by adding a nuclear import signal showed that Rx1 is activated in the cytoplasm and that in the nucleus the elicitor cannot activate Rx1. On the other hand adding a nuclear import or export signal to Rx1 had no noticeable effect on its functioning. Furthermore, we demonstrate that disrupting the function of the Rx1 protein by mutating the ATP/ADP-binding P-loop, or by silencing the co-chaperones SGT1 and RAR1, impaired the translocation of the full Rx1 protein to the nucleus, while Rx1 versions lacking the LRR domain were not affected. The interplay between the domains and the possible role of conformational changes in R protein signaling and localization are discussed.

RESULTS

Fluorescent protein fusions of Rx1 are functional

Rx1 is a large (110 kDa) modular protein that, upon recognizing the coat protein of the plant virus PVX, mediates a strong local resistance response inhibiting the replication and systemic spreading of the virus. Based on the absence of transmembrane domains or specific subcellular targeting motifs in the amino acid sequence Rx1 is assumed to be localized in the cytoplasm. To test this assumption we generated fluorescent protein fusions of Rx1. The fluorescent proteins, GFP or YFP, were fused via a short linker of glycine and serine residues to either the N- or the C-terminus of Rx1 (Fig. 1A). To ensure that steric hindrance from the GFP-fusion did not interfere with R protein activity, we tested the functionality of Rx1 with an N-terminal fluorescent fusion in transgenic potato plants. Potato plants expressing CaMV 35S promoter driven YFP-Rx1 showed an extreme resistance (ER) phenotype to the avirulent PVX_{UK3}, but not to the PVX_{HB} breaker strain (Fig. 1B). The extreme resistance phenotype was similar to that of the resistant potato clone SH with Rx1 in its genetic background. The functionality of GFP-Rx1 and Rx1-GFP was tested in a transient assay by coexpressing them in *Nicotiana benthamiana* leaves with the avirulent (CP106) or virulent (CP105) coat protein, or with GFP as negative control. Under these conditions both the N-terminal and C-terminal GFP-fusions of Rx1 gave a strong HR to the avirulent coat protein CP106 and a weaker HR in response to the virulent coat protein (CP105). Rx1 without GFP fusion does not show an HR in response to CP105 under similar circumstances. No response could be seen after coexpression with GFP, ruling out autoactivation (Fig. 1C).

Rx1 is located in the cytoplasm and nucleus

To study the localization of the fusion proteins they were transiently expressed in *N. benthamiana* leaves by means of *Agrobacterium* transient transformation assays. Optimal fluorescence levels were reached 3 days after infiltration. Fluorescence was studied *in planta* in living cells of the abaxial leaf epidermis using confocal laser scanning microscopy. The GFP-fused Rx1 constructs accumulated to relatively low levels compared with free GFP, but stayed well above background fluorescence levels. Under similar conditions no fluorescence could be seen in cells where Rx1 was expressed without a fluorescent protein fusion (Fig. 1E). The fluorescent signal was observed in both the cytoplasm and the nucleus for both GFP-Rx1 and Rx1-GFP (Fig. 1E). Because the threshold for free diffusion to and from the nucleus is around 40 kDa in plant cells (Merkle, 2003), we did not expect a nuclear localization for the fluorescent Rx1 proteins, which have a predicted size of 140 kDa. The nuclear fluorescence can not be explained by the degradation of fluorescent Rx1 proteins, as is shown by Western blot using an antibody against GFP (Fig. 1D).

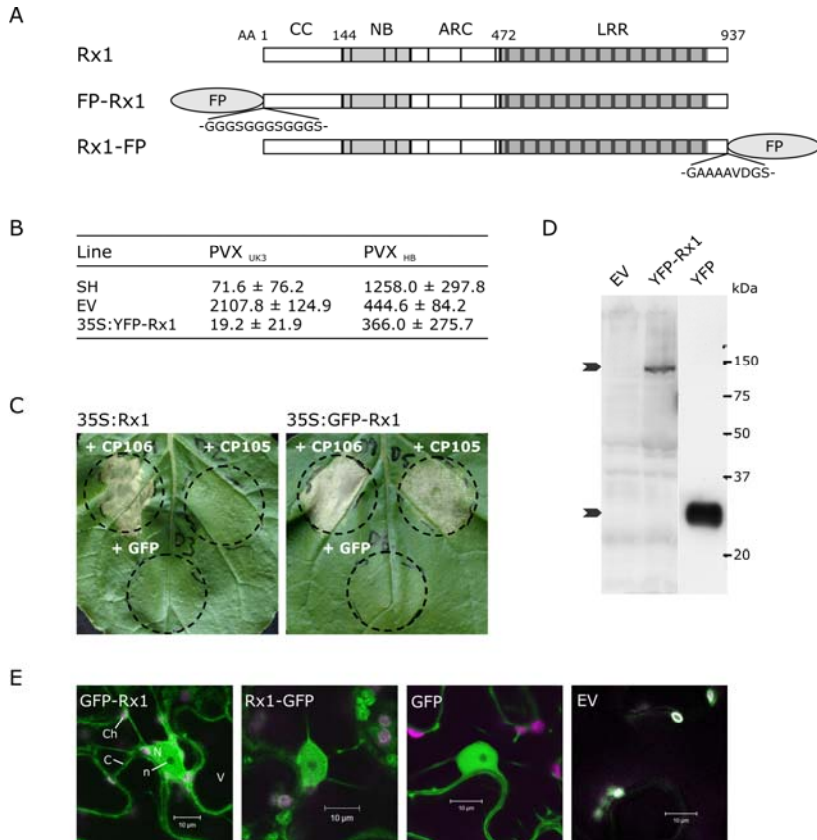


Figure 1. Subcellular localization of the full length Rx1 protein **A.** Schematic overview of the N- and C-terminal fluorescent protein (FP) fusions of full-length Rx1. Both yellow fluorescent protein (citrine YFP) and green fluorescent protein (enhanced GFP) fusions were made. The amino acid sequences of the linkers connecting the FP to Rx1 are shown. Expression of the fusion constructs was controlled by the CaMV 35S promoter. **B.** Virus resistance assay on transgenic potato lines transformed with empty pBINPLUS plasmid or Rx1 under control of the CaMV 35S promoter after mock or sap inoculation with the avirulent PVXUK3 or virulent PVXHB strain. Three plants from three independent primary transformants were assayed. SH is the Rx1 resistant potato genotype used as a positive control. Virus concentrations were measured at 21 days after inoculation in an ELISA with a PVX specific antibody (Mean absorbance values at 405nm (± SD)). **C.** Hypersensitive response (HR) obtained with Rx1 or GFP-Rx1 when co-infiltrated with the avirulent PVX coat protein (CP106), the virulent PVX coat protein (CP105) or a negative control (GFP) in an agroinfiltration assay on *N. benthamiana* leaves. Images were taken two days after infiltration. **D.** Western blot of protein extracts from *N. benthamiana* leaves transiently transformed with YFP-Rx1, an empty vector control (EV) or free YFP. Protein was detected with an anti-GFP antibody. The two arrows indicate the position of YFP-Rx1 (140 kDa) and YFP (27 kDa) on the blot. **E.** Localization pattern of GFP-Rx1, Rx1-GFP and GFP. The GFP-labelled proteins were imaged by confocal microscopy in the *N. benthamiana* epidermal cells of transiently transformed leaves. Empty vector (EV) transformed cells are shown as control for background fluorescence. Subcellular structures are indicated in the first panel (N=nucleus, n= nucleolus, C = cytoplasm, V = vacuole, and Ch = chloroplast). Images were taken three (GFP-Rx1, Rx1-GFP, EV) and two (GFP) days after agroinfiltration.

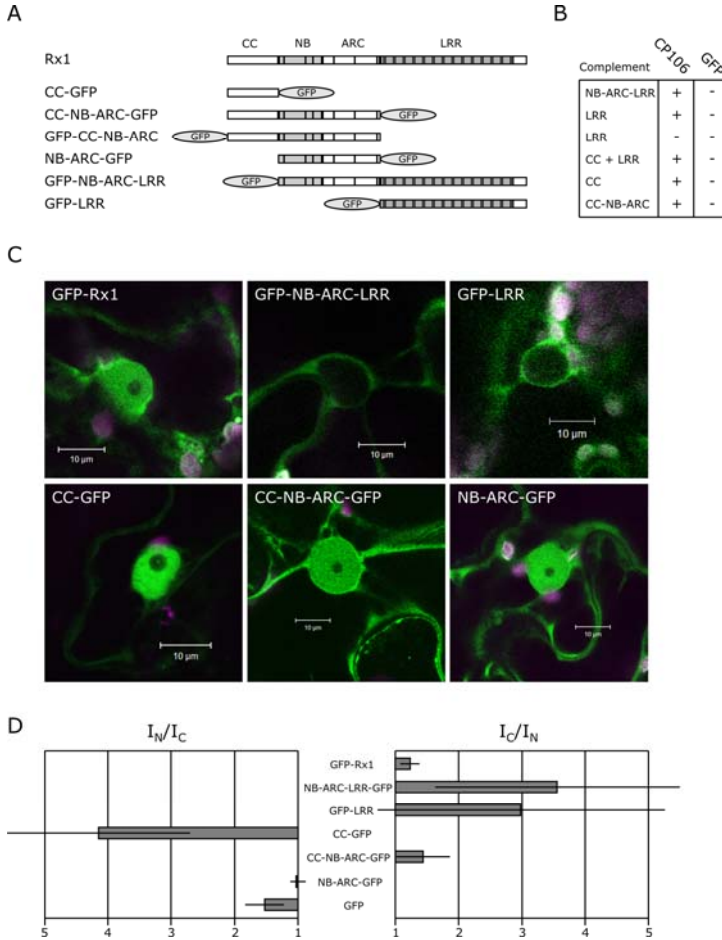


Figure 2. Subcellular localization of truncated Rx1 and its separate domains. **A.** Schematic overview of the N- and C-terminal fluorescent protein (GFP) fusions of the Rx1 subdomains and deletion constructs. **B.** Complementation in trans of the fluorescent constructs by the corresponding Rx1 domains in an agroinfiltration assay. Fluorescent fusions of the Rx1 subdomains were coexpressed with their complementary parts and the avirulent PVX CP. Where the subdomains can complement each other an HR was observed at three days after agroinfiltration (+). **C.** Subcellular localization of the GFP-labeled truncated Rx1 polypeptides and separate domains shown by confocal imaging. Full length GFP-Rx1 is shown for comparison. Images were taken two (CC-GFP, CC-NB-ARC-GFP, NB-ARC-GFP) or three days post infiltration (NB-ARC-LRR-GFP and GFP-LRR). **D.** Nucleo-cytoplasmic distribution of the fluorescent fusion proteins shown as fluorescent intensity ratios. Average fluorescence intensity ratios (\pm SD) were determined from the fluorescence intensities on the cytoplasm and nucleus in confocal images of 8 *N. benthamiana* leaf epidermal cells with the java-application ImageJ.

The CC domain accumulates in the nucleus and is required for nuclear localization of the Rx1 protein

No classical nuclear localization signals, which could explain its presence in the nucleus, are predicted in the amino acid sequence of Rx1 (PredictNLS, (Cokol et al., 2000)). To gain insight in the mechanism underlying the observed nucleocytoplasmic distribution, we constructed a series of fluorescent proteins in which specific Rx1 domains are deleted (Fig. 2A). The Rx1 protein is composed of several flexibly linked domains that upon co-expression as separate modules still function as the original Rx1 protein (Moffett et al., 2002). This characteristic allowed us to test if the fluorescent fusion constructs of the domains retained their functionality when coexpressed with the complementary parts of Rx1 in an agroinfiltration assay on leaves of *N. benthamiana*. The N-terminal half of the Rx1 protein (CC-NB-ARC) proved to be sensitive to the position of the fluorescent fusion protein. The YFP-CC-NB-ARC product lost the ability to induce an HR when expressed *in trans* with the LRR and the elicitor, whereas the full-length Rx1 tolerates an N-terminal fluorescent protein fusion. The CC-NB-ARC-GFP construct on the other hand, showed normal functionality in a complementation assay. In addition, the N-terminal fusion construct GFP-LRR was shown to be functional and also for the C-terminal fusion construct CC-GFP, a normal hypersensitive response was observed when expressed with an NB-ARC-LRR construct. Even the combination of the CC-GFP, NB-ARC-GFP and GFP-LRR constructs *in trans* could induce an HR response in the presence of the PVX coat protein (Fig. 2B). Apparently, most fusions with fluorescent protein do not disturb the assembly of the domains into a functional protein. Confocal microscopy showed that the subcellular localization patterns of various fluorescent versions of Rx1 differed markedly from the full Rx1 protein (Fig. 2C). The nucleocytoplasmic distribution was quantified by determining the ratios of the fluorescent intensities in the nucleus and cytoplasm from the confocal images (Fig. 2D). The CC-NB-ARC or the NB-ARC-GFP showed an almost equal distribution between cytoplasm and nucleus, as observed for the full-length Rx1 protein. However, the NB-ARC-LRR-GFP construct was predominantly localized in the cytoplasm, indicating that the CC domain is essential for nuclear accumulation of the full Rx. The high nuclear accumulation of the CC domain supported this interpretation. The LRR domain seemed to have an opposing role. The GFP-LRR protein was excluded from the nucleus and only observed in the cytoplasm (Fig. 2C & D).

A short helix-rich segment determines nuclear accumulation of CC domain

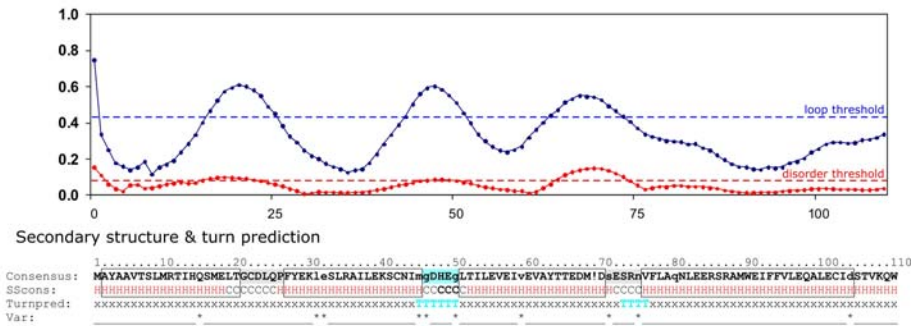
To investigate whether a distinct sequence within the CC domain determines the nuclear location of Rx1, a set of small fragments of the CC was fused to GFP. Boundaries of the CC domain were delineated with CDART, Interpro and predictors for secondary structure and intradomain loop sequences. Both CDART and Interpro indicated that the NB-ARC domain starts at amino acid (aa) 136 while the helix propensity in Rx1 stops around aa 115, as shown by the consensus secondary structure prediction. In addition the Domain Linker Predictor (DLP) indicated a high propensity for intradomain loop sequences between aa 115-135. All these suggested that in Rx1 the CC domain can be set approximately between aa 1-115. Consensus secondary structure profile indicated here the presence of four major helical regions: H1 = aa 2-20, H2 = aa 25-45, H3 = aa 51-69, H4 = aa 76-110 joined by three coils that also showed a high propensity for intrinsic disorder. In addition, more specialised

predictors such as SOPMA 5-state and BETATURN indicated a high propensity for beta-turn between H2-H3 and to a lesser extent between H3 and H4 (Fig. 3A).

The four fragments predicted to form helices were each fused to GFP, individually or in combination with neighboring helices, resulting in seven constructs (Fig. 3B). Five of these constructs showed a subcellular distribution similar to free GFP, with an almost equal intensity in the cytoplasm and the nucleus (Fig. 3C and D). Two constructs (A-GFP and G-GFP), however, showed a strikingly different distribution. Fragment A containing the most N-terminal two predicted helices (aa 1-45) was, despite its small size, almost completely absent from the nucleus. Fragment A was predominantly present in the cytoplasm where it associated with structures resembling Golgi bodies and the endoplasmic reticulum (Fig. 3D). This subcellular localization pattern was not observed for fragment B (helix 1) and C (helix 2), indicating that the combination of the two helices is essential for the cytoplasmic targeting of fragment A. Fragment G, containing the third and fourth predicted helix, had a strong nuclear accumulation, comparable with the full CC domain. No such nuclear accumulation was observed for fragment E1n (helix 3) and F (helix 4), showing that the combination of these two helices is required for the nuclear targeting of fragment G. All CC fragment constructs were shown to be stable by immunoblot with anti-GFP antibody (Fig. 3E). These results demonstrate that the CC domain includes two non-overlapping subdomains each consisting of two α -helices, and which show distinct localization patterns.

In addition, we determined the subcellular distribution of a set of three N-terminal and six C-terminal deletion constructs of the CC domain (Rairdan et al., 2008). The C-terminal serial deletion constructs encoded the amino acids 1-55, 1-67, 1-73, 1-79, 1-85 and 1-115 and the N-terminal deletion constructs coded for 14-144, 26-144 and 37-144. All constructs had a C-terminal GFP fusion. The three N-terminal deletion constructs had a localization similar to the full-length CC-constructs, with a high concentration of the fluorescence in the nucleus. The largest C-terminal deletion construct (1-115) too, showed high fluorescence intensity inside the nucleus. Fragment 1-85 however, lacking most of the fourth helix, showed a predominantly cytoplasmic localization, similar to the pattern seen for fragment A (1-45). Also the smaller constructs, 1-55, 1-67, 1-73 were almost completely absent from the nucleus. So, from these results we concluded that the sequence between amino acid 85 and 115 was necessary for the nuclear accumulation of the CC domain, but not sufficient, because fragment F, comprising aa 79-116, showed no nuclear accumulation. These observations were in line with the findings of the secondary structure based fragments (Fig. 3) and confirm that the two α -helices between amino acid position 1-45 and the two α -helices between position 45-116 represent separate subdomains with two distinct localization patterns.

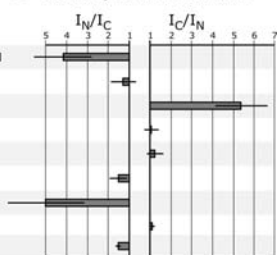
A Intrinsic disorder and loop prediction



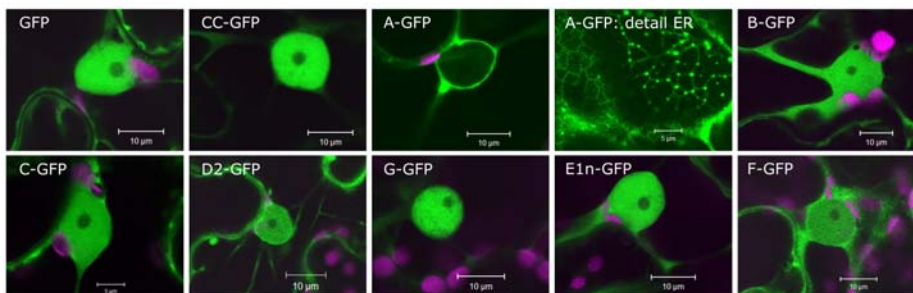
B Rx1 CC fragment GFP-fusion constructs



C Intensity distribution ratio



D



E

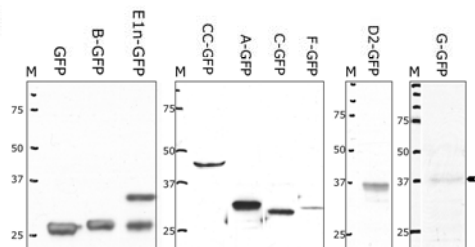


Figure 3 (left page). Subcellular localisation of fluorescent fusion proteins of secondary structure based CC domain fragments of Rx1. **A.** Consensus secondary structure propensity of the CC domain suggests the presence of four alpha-helices (H - shown in red) joined by three stretches that show a high intrinsic disorder propensity defined both by the inability to form secondary structure ('coils' - in blue) and high B-factors ('hot-loops' - in red). In addition the CC sequence shows a high propensity for beta-turn structure especially between helices 2 and 3 (T - shown in blue). **B.** Schematic representation of the constructed Rx1 CC fragments, which were all fused to GFP at the C-terminus. The predicted secondary structure is indicated in color: α -helices in red, the central β -turn in blue. The amino acid positions corresponding with the fragments sequence in the full CC domain are noted after the fragments name. **C.** Fluorescence intensity distribution ratio showing the nuclear (I_N/I_C) or cytoplasmic (I_C/I_N) localization of each CC-fragment. Average intensities were determined in confocal images using the program ImageJ. Ratios shown are the averages (\pm SD) of 8-10 cells. **D.** Confocal images showing the specific subcellular localization patterns for each CC fragment. For fragment G, a similar nuclear localisation was observed like shown for the full CC domain. For fragment A, fluorescence was exclusively located in the cytoplasm in ER and Golgi-like structures (a detail of the cytoplasmic localization pattern fragment A is shown). The other fragments were shown to be located in both the nucleus and the cytoplasm comparable to free GFP. **E.** Western blot of the fluorescent fusion proteins for each CC fragment when expressed in an agroinfiltration assay on leaves of *N. benthamiana*. Proteins of the expected size were detected with an anti-GFP antibody and free GFP (27 kDa) was used as a control. Only for E1n-GFP, an additional band was observed.

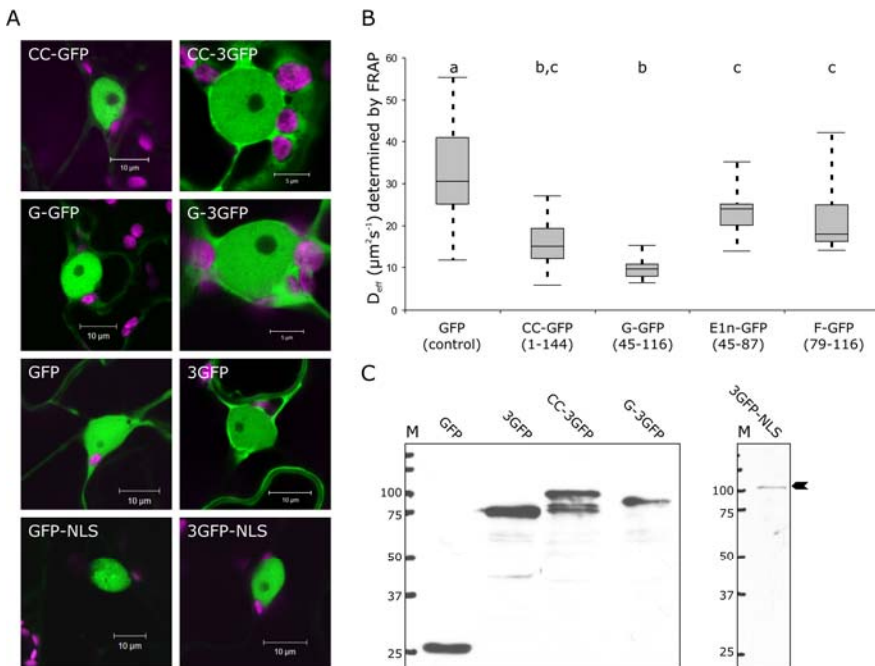


Figure 4. Diffusional behaviour of Rx1 CC fragments **A.** Confocal images of the subcellular localization of single GFP fusion constructs (upper panels: CC-GFP, G-GFP, GFP and GFP-NLS) and the equivalent triple GFP fusion constructs (lower row: CC-3GFP, G-3GFP, 3GFP and 3GFP-NLS), showing that 3GFP-NLS is still targeted to the nucleus, whereas the nuclear localization of CC-3GFP and G-3GFP is reduced compared to GFP-CC and GFP-G that coincides with an accumulation of these fusion proteins in the cytoplasm. GFP fluorescence (excitation 488 nm, band pass 505-550 nm) is displayed in green and chloroplast fluorescence (above 650 nm) in magenta. **B.** Effective diffusion coefficients (D_{eff}) of GFP, CC-GFP, G-GFP, E1n-GFP and F-GFP as derived from FRAP (fluorescence after photo bleaching) measurements in the nucleus of *N. benthamiana* cells ($n = 20$, except for GFP: $n=35$) 2 days post agroinfiltration. For fragment G and the full CC domain, a significant slower diffusion was observed compared to the other fragments tested and free GFP. The distribution of data points is depicted as quartiles and medians. The means were statistically analysed with the Tukey-test and grouped (a, b, and c) according to significant differences at alpha 0.01. **C.** Western blot of the triple GFP fusion constructs to show the integrity of the constructs. Free GFP was used as control. The constructs were transiently expressed in *N. benthamiana* leaves. Fusion proteins of the expected size were detected with an anti-GFP antibody.

The CC domain binds to large complexes in the nucleus and contains no strong autonomous nuclear import signals

The CC constructs described above are all fusions with a single GFP and none of the fusion products has a molecular mass larger than 45 kDa, leaving the question open whether the observed nuclear accumulation of the full CC and fragment G resulted from active or passive nuclear import. To limit passive diffusion through the nuclear pore, triple GFP fusions were constructed for the CC domain and fragment G, which increased the overall mass of the constructs to approximately 97 and 89 kDa, respectively. As controls, a triple GFP construct without targeting signals and a triple GFP construct containing the SV40 NLS were made. The more intense fluorescence in the layer of cytoplasm surrounding the nucleus suggested that the nucleocytoplasmic distributions of the CC-3GFP, G-3GFP, and 3GFP proteins all shifted to a more cytoplasmic localization compared with the single GFP constructs (Fig. 4A). It was noted that although the 3GFP protein (95 kDa) had a molecular mass larger 45 kDa, it was still able to diffuse passively into the nucleus, probably because the three fused GFPs did not form a single globular unit. However, the passive diffusion of the 3GFP protein into the nucleus was far less efficient than the diffusion of the single GFP protein (Fig. 4A) and could not be explained by instability of the protein (Fig. 4C). The nucleocytoplasmic distribution patterns showed that the CC domain and fragment G do not contain strong autonomous signals directing the 3GFP proteins actively into the nucleus. The CC-3GFP and G-3GFP constructs had distributions similar to the triple GFP construct, while the triple GFP construct containing the SV40 NLS for active nuclear import was almost exclusively localized in the nucleus.

If the CC protein and fragment G (45-116) moved into the nucleus by passive diffusion, then their nuclear accumulation can only be explained by a mechanism in which the protein is sequestered inside the nucleus, for example via an interaction with a nuclear component. To test this hypothesis, local photobleaching was applied to the fluorescent fusion proteins and the dynamics of fluorescence recovery were studied using FRAP (fluorescence recovery after photobleaching). A small area in the nucleus was bleached and fluorescence recovery was monitored. Subsequently, the recovery half times were used to calculate the diffusion coefficients. Free GFP is known to diffuse unrestrictedly in the nucleus (Thompson et al., 2002; Houtsmuller, 2005). As expected we observed the highest diffusion coefficients for the free GFP (Fig 4B). For the full CC domain a significantly smaller diffusion coefficient was measured (Fig. 4B). Fragment G (45-116) had even a lower diffusion coefficient than the full CC domain. The two fragments E1n (45-87) and F (79-116) gave diffusion coefficients intermediate to the values for GFP and the CC-GFP, but still diffused significantly faster than fragment G (45-116). Thus, the strong nuclear localization of the CC domain and of fragment G (45-116) coincided with an apparent slower diffusion in the nucleus. This suggested that the nuclear accumulation of the CC protein and fragment G can be explained by transient binding with very large nuclear complexes or immobile structures.

Nuclear localization of Rx1 is impaired upon silencing *NbSGT1* and *NbRAR1*

Because SGT1 and RAR1 are known to form a complex and are essential for the functioning of many R proteins, possibly as part of a so-called resistosome or as an essential chaperone complex for R proteins (Austin et al., 2002; Azevedo et al., 2002; Azevedo et al., 2006; Boter et al., 2007), we tested whether their presence affects Rx1 localization. SGT1 and RAR1 were knocked out by TRV based virus induced gene-silencing (VIGS) in *N. benthamiana* plants. Rx1 and the avirulent CP were co-expressed at low levels in a transient agroinfiltration assay. As expected the silencing of SGT1 prevented the HR completely. Silencing RAR1 partially blocked the HR. Interestingly, the subcellular localization of the fluorescent Rx1 protein was clearly affected in SGT1 and RAR1 silenced plants. In wild-type plants the fluorescence intensity observed in the cytoplasm is just slightly higher than the intensity observed in the nucleus ($I_C/I_N = 1.2$). In the SGT1-silenced plants the Rx1 protein was mostly excluded from the nucleus ($I_C/I_N = 4.4$) (Fig. 5). Silencing RAR1 enhanced the Rx1 concentration in the cytoplasm as well ($I_C/I_N = 2.5$) (Fig. 5). Constructs not containing the LRR did not exhibit a shift towards the cytoplasm in the absence of SGT1 and RAR1. The nuclear accumulation of the CC-GFP construct (data not shown) or the distribution of the CC-NB-ARC-GFP construct (Fig. 5) were not changed in SGT1 and RAR1 silenced plants. This suggests that the effect of SGT1 and RAR1 on the subcellular distribution of Rx1 is linked to its LRR domain. The empty vector control did not affect the Rx1 distribution, excluding the possibility that the phenotype was a side-effect of the silencing system.

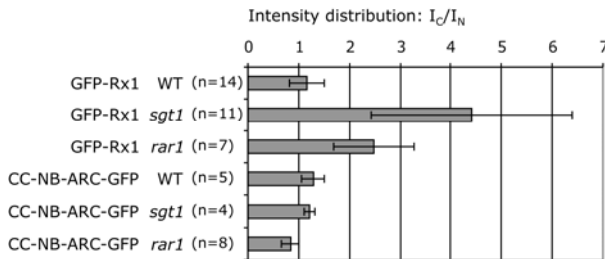


Figure 5. The effect of SGT1 and RAR1 silencing on the nucleo-cytoplasmic distribution of GFP-Rx1. Full-length GFP-Rx1 and the truncated Rx1 CC-NB-ARC-GFP constructs were expressed in wild type *N. benthamiana* (WT) and *N. benthamiana* in which SGT1 or Rar1 were silenced by VIGS using TRV (*sgt1* and *rar1*). The average intensity ratios (\pm SD) were determined for 10 cells per combination (cytoplasmic intensity I_C / nuclear intensity I_N). A shift towards a more cytoplasmic localization of GFP-Rx1 in SGT1 and RAR1 silenced plants is seen as a higher average I_C/I_N . No such shift was observed for CC-NB-ARC-GFP.

Nuclear localization of Rx1 requires ADP/ATP binding

It is hypothesized that conformational changes following recognition of the elicitor switch the R protein from an inactive form into an active form, a process which involves changes in nucleotide-binding status (Tameling et al., 2002; Tameling et al., 2006) and intramolecular interactions (Moffett et al., 2002; Rairdan and Moffett, 2006; Rairdan et al., 2008). A mutation in the P-loop of the nucleotide binding domain abolishes the activity of Rx1, and changes the interaction between the CC domain and the NB-ARC-LRR (Moffett et al., 2002). To study the effect of conformational changes on the nucleocytoplasmic distribution of Rx1, a point mutation (K176R) was made in the P-loop.

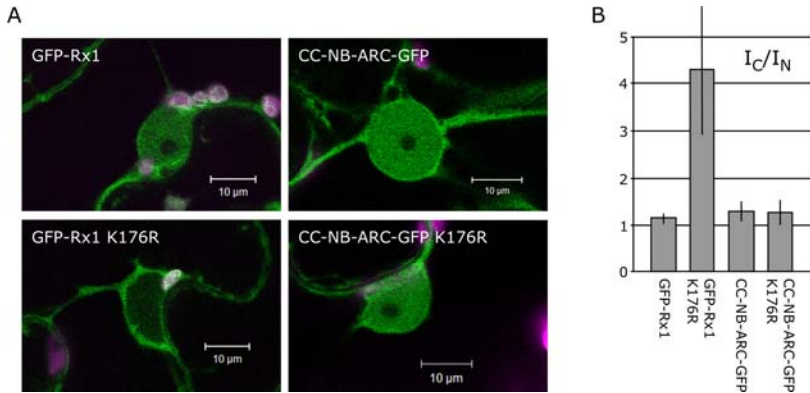


Figure 6. The effect of the P-loop mutation on the Rx1 localization. **A.** Confocal images of the subcellular localization of GFP-Rx1, GFP-Rx1 with mutation K176R (P-loop), the CC-NB-ARC-GFP and the CC-NB-ARC-GFP with mutation K176R in *N. benthamiana* cells, showing a specific more cytoplasmic localization pattern for GFP-Rx1 K176R. The CC-NB-ARC-GFP K176R construct does not show a shift towards the cytoplasm. **B.** The average fluorescence intensity ratio between the cytoplasm (I_C) and the nucleus (I_N) was determined for 10 *N. benthamiana* cells per constructs.

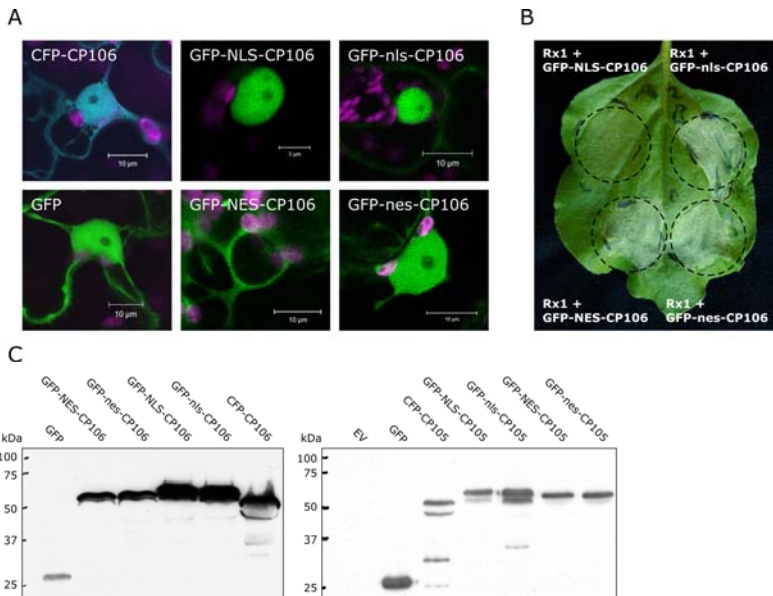


Figure 7. PVX coat protein targeted to the nucleus can not activate Rx1. **A.** A Confocal image of the subcellular localization of the PVX coat protein when expressed as a fusion to cyan fluorescent protein (CFP-CP106) is shown in the upper left panel (CFP fluorescence: blue, chloroplast fluorescence: magenta). The localization of GFP-tagged versions of CP106 with a targeting signal (SV40 NLS, PKI NES and mutated variants thereof) are shown in the lower four panels. **B.** HR phenotypes of exogenous targeted GFP-CP106 constructs coexpressed with Rx1. GFP-NLS-CP106 (SV40 NLS, PKKKRKVEDP), GFP-NES-CP106 (PK1 NES, NELALKLAGLDINK) and the mutated versions thereof (GFP-nls-CP106, PKNKRKVEDP and GFP-nes-CP106, NELALKAGADANK) were coexpressed with 35S:Rx1 in an agroinfiltration assay in *N. benthamiana* leaves. At 2 dpi, a clear HR was observed for all combinations except for the coexpression of Rx1 with GFP-NLS-CP106. The image was taken 3 dpi. The infiltrated spots are marked by a dashed circle. **C.** Western blot of free GFP, CFP-CP106, CFP-CP105, and the NLS and NES versions of GFP-CP106 and GFP-CP105, showing fusion proteins of the expected size when probed with an anti-GFP antibody.

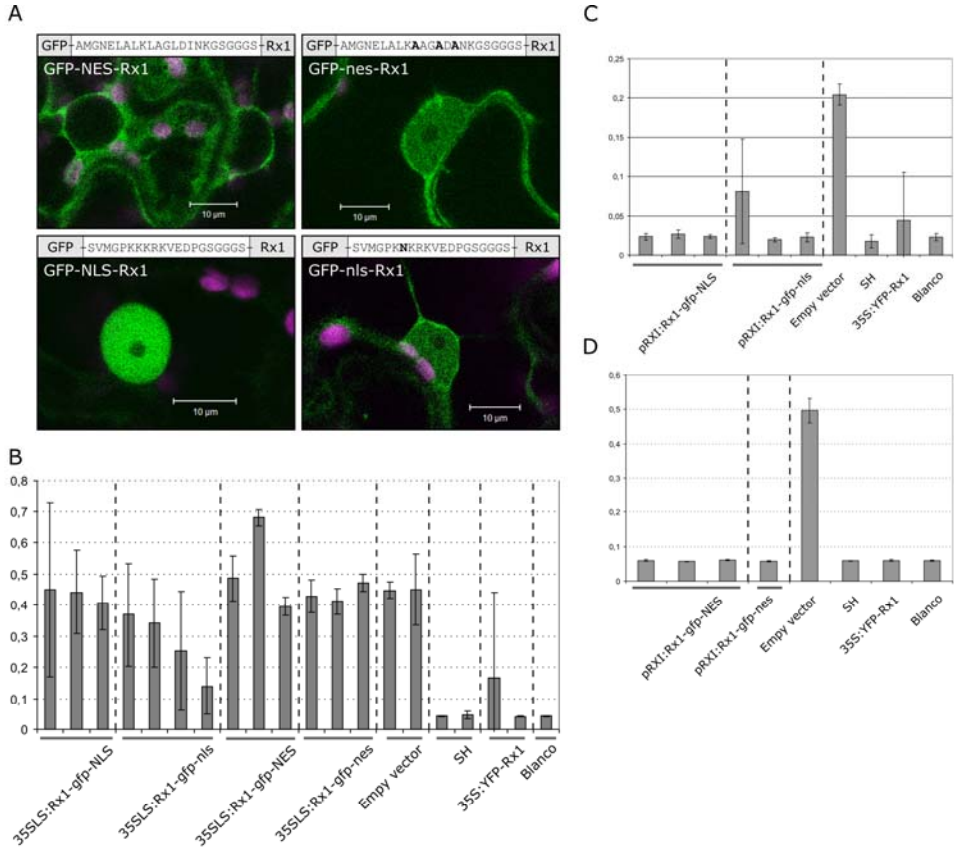


Figure 8. The effect of exogenous targeting signals on the localization of Rx1. **A.** Schematic overview of the GFP-Rx1 constructs where in between the GFP and the Rx1 sequence a nuclear localization signal (SV40 NLS) or a nuclear export signal (PK1 NES) is incorporated. As control, for both the NLS and NES mutated versions were used (nls and nes). The sequence of each targeting signal is depicted above the image showing the localization of the construct in *N. benthamiana* cells. GFP-NES-Rx1 and GFP-NLS-Rx1 show a localization pattern as expected for the corresponding targeting signal. The mutated versions of the targeting signals result in a localization pattern similar to GFP-Rx1 without targeting signal, with an equal distribution of fluorescence between the cytoplasm and the nucleus. **B.** Virus resistance assay on transgenic potato plants transformed leaky scan promoter regulated Rx1-GFP-NLS, nls, NES, and nes constructs. Two empty vector transformed potato lines, resistant line SH and two 35S:YFP-Rx1 transgenic lines served as controls. Avirulent PVX accumulation was quantified in an ELISA assay (See figure 1). Five replicates were used per line. **C and D.** Virus resistance assay on transgenic potato lines harbouring prX1:Rx1-GFP-NLS and nls (**C**) or prX1:Rx1-GFP-NES or -nes (**D**). Per construct three lines were used and per line five replicates. For prX1:Rx1-GFP-nls only one transgenic line was available.

The fluorescence intensity of the GFP-Rx1 K176R construct was similar to the intensity measured for the GFP-Rx1 construct under identical microscope settings, indicating similar protein levels. In contrast to GFP-Rx1, fluorescence of GFP-Rx1 K176R was almost absent from the nuclei (Fig. 6A). The fluorescence intensity ratio between cytoplasm and nucleus (I_c/I_n) for GFP-Rx1 K176R was 4.3 ± 1.4 (Fig. 6B), resembling the effect of silencing SGT1. Since the nuclear exclusion of GFP-Rx1 in SGT1 and RAR1 silenced plants was linked to the LRR, we tested whether the nuclear exclusion of the P-loop mutant was similarly LRR-

dependent. A truncated Rx1 version containing the K176R mutation, and lacking the LRR domain was constructed. We used the wild-type CC-NB-ARC-GFP as control. When expressed in *N. benthamiana* cells, the localization patterns of both fluorescent CC-NB-ARC versions were indistinguishable (Fig. 6A). The CC-NB-ARC-GFP K176R mutant was not excluded from the nucleus. Hence, we conclude that the nuclear exclusion of the full-length GFP-Rx1 K176R depends on the presence of the LRR domain.

Manipulating the nucleocytoplasmic distribution of the PVX coat protein shows that Rx1 is activated in the cytoplasm

To study the functional differentiation of the cytoplasm and the nucleus in initiating an HR response, we manipulated the nucleocytoplasmic distribution of the PVX coat protein. Fluorescent versions of avirulent and virulent PVX coat proteins were constructed by fusing a fluorescent protein (CFP or GFP) to the N-terminus of the coat proteins. The fluorescent fusions did not alter their recognition by Rx1 (data not shown) and have been shown not to hamper the functioning of the coat protein (Cruz et al., 1996). Confocal microscopy showed that the avirulent and virulent coat proteins do not differ in their subcellular distribution in *N. benthamiana* leaf epidermis cells. Both were found in the nucleus and the cytoplasm in equal intensities, as was expected because the mass of the fusion products (~45 kDa) does not exceed the size-exclusion limit of the nuclear pore and their sequence does not contain specific targeting signals (Fig. 7A). These observations are in accordance with the localization pattern described earlier (Batten et al., 2003).

To distinguish the role of the cytoplasm and nucleus in the recognition of PVX, we constructed versions of the coat protein that contained the SV40 nuclear localization signal (NLS)(Lanford and Butel, 1984; Haasen et al., 1999) or the PKI nuclear export signal (NES)(Wen et al., 1995) and mutated versions of these signals as controls. The NLS-versions of the virulent and avirulent coat protein are efficiently targeted to the nucleus and the NES-versions are almost completely excluded from the nucleus (Fig. 7A). All NLS and NES versions are stable and expressed at similar levels as was shown by immunodetection on western blot (Fig. 7C). The ability of these coat protein versions to activate Rx1 was tested in an agroinfiltration assay on *N. benthamiana* leaves. All avirulent coat protein constructs elicited an HR within two days after the infiltration, except the construct containing the functional nuclear localization signal (Fig. 7B). The mutated NLS control construct differs in only one amino acid (KKKRRK > KNKRRK) from the functional NLS construct, but can still fully elicit an Rx1 mediated HR. As an extra control, the virulent coat proteins were tested with GFP-Rx1, which gives a weak response to the virulent coat protein (Fig. 1C). In these combinations too, only the version of the coat protein with a functional NLS escaped recognition (data not shown). From these results we conclude that recognition of the PVX coat protein takes place in the cytoplasm and the nucleus does not provide an environment in which Rx1 can be activated. Our data also show that Rx1-mediated signaling requires no, or at least no high concentrations of PVX coat protein in the nucleus, as no effect was observed upon adding a nuclear export signal to the avirulent coat protein.

The functioning of Rx1 is not influenced by manipulating its nucleo-cytoplasmic distribution by adding exogenous targeting signals

Since adding nuclear import signals to the PVX coat protein impaired the activation of Rx1, we also studied whether diminishing the Rx1 concentration in the cytoplasm had the same effect. It was also anticipated that lowering the Rx1 concentration in the nucleus by adding nuclear export signals may block the development of an HR. Recent publications have shown that decreasing the nuclear concentration of the barley R protein MLA10 and the tobacco protein N by adding an exogenous nuclear export signal, strongly limits their ability to induce a cell death response (Burch-Smith et al., 2007; Shen et al., 2007). To test these hypotheses, we created a set of constructs of the full Rx1 protein fused to either a nuclear localization signal (NLS) or a nuclear export signal (NES) and mutated versions of these signals as controls. These constructs were transiently expressed using the endogenous Rx1 promoter. Confocal microscopy imaging of the constructs expressed in *N. benthamiana* confirmed that the NES and NLS sequences were able to redirect the localization of the GFP-Rx1 constructs, whereas GFP-Rx1 constructs with a mutated version of the targeting signal had a localization pattern identical to the original GFP-Rx1 construct (Fig. 8). To our surprise, the addition of neither targeting signal influenced the development of an HR in an agro-infiltration assay on *N. benthamiana* leaves. Also no effect was observed on the spreading of the virus as was tested by coexpressing GFP-NES-Rx1 or GFP-NLS-Rx1 with PVX:GFP (Peart et al., 2002a). Because our experiments with the PVX coat protein demonstrated that Rx1 is activated in the cytoplasm and cannot be activated in the nucleus, we concluded that despite the addition of a nuclear localization signal, the concentration of Rx1 in the cytoplasm was still above the threshold level required for triggering an HR. In order to attain a level below this threshold we lowered the *Agrobacterium* concentration used in the agroinfiltration assay to $OD_{600} = 0.01$. Furthermore we decreased the sensitivity of Rx1 by coexpressing targeting signal fusions of the Rx1 protein domains *in trans*. Both approaches revealed no difference between GFP-Rx1 and the GFP-NLS-Rx1. Therefore we concluded that with these transient expression approaches it was not possible to lower the cytoplasmic concentration enough to affect Rx1 functionality. A similar conclusion can be drawn for the GFP-NES-Rx1. Transgenic potato lines were created with NLS/nls or NES/nls tagged versions of Rx1 under either the 35S leaky scan promoter (described in Chapter 4) or the native RXI regulatory sequences. Five transgenic lines per construct were tested for virus resistance against PVX^{UK3}. Accumulation of PVX in these plants was tested in an ELISA assay. All NLS, nls, NES and nes tagged versions of Rx1 expressed from its native regulatory sequences gave full PVX resistance. All tagged versions of Rx1 expressed from the leaky scan promoter had lost the ability to confer resistance, probably because expression levels were too low. No significant differences could be seen between Rx1 constructs tagged with functional or mutated versions of the localization signals.

Discussion

In this study we have shown that the CC-NB-LRR resistance protein Rx1 is localized to both the nucleus and the cytoplasm. Our data indicate that the nuclear localization of Rx1 is not regulated by short linear sequences, but appears to be dependent on its folding state and the interplay between its domains. Although the CC domain is required for the accumulation of Rx1 in the nucleus, analyses of subdomain fragments revealed no autonomous nuclear localization signals in the CC protein. Also, the complete CC domain shows no strong activity in transporting triple GFP fusions into the nucleus, indicating that the CC domain on itself is not sufficient to translocate Rx1 to the nucleus. Fluorescence recovery after photobleaching (FRAP) demonstrated that the strong nuclear accumulation of the separate CC domain can be explained by transient binding to large complexes or immobile structures like chromatin. A small region, composed of two α -helices, within the CC was found to be responsible for transient binding to nuclear components. The LRR domain is predominantly located in the cytoplasm, while the NB-ARC domain has a dual location, which is hardly affected by addition of the CC domain (Fig. 2C & D). Our data suggest that the LRR domain has a crucial role in adopting the appropriate conformation for Rx1 to accumulate in the nucleus. Disrupting the ability of the NB domain to bind ADP or ATP results in nuclear exclusion of the full Rx, while the CC-NB-ARC and NB-ARC proteins retain their dual location. Also silencing the putative co-chaperones SGT1 or RAR1 results in a dramatic decrease of the Rx1 concentration in the nucleus, while proteins lacking the LRR were not affected by these factors. Remarkably, the forced relocation of the Rx1 protein by the addition of exogenous nuclear export or import signals did not affect the functioning of the protein in our transient assays. However, the Rx1 protein is highly sensitive to changes in the nucleocytoplasmic distribution of its elicitor. Manipulating the dual location of the PVX coat protein by adding a nuclear import signal impaired the induction of a hypersensitive response, indicating that the Rx1 protein is activated in the cytoplasm. Shifting the nucleocytoplasmic distributions of Rx1 and its elicitor also demonstrated that the cytoplasmic concentration of Rx1 required to initiate a hypersensitive response is relatively low when compared to its elicitor.

Proteins of masses above 40 kDa are limited in their passive passage through the nuclear pore complex (Merkle, 2003), and the large R proteins need to be actively transported into and out of the nucleus. A small number of NB-LRR proteins does contain predicted classical NLS sequences, as was noticed in the analysis of the R proteins encoded in the poplar genomes (Tuskan et al., 2006; Kohler et al., 2008). The *Arabidopsis* TIR-NB-LRR RPS4 has a bipartite NLS at its C-terminus which is necessary for its nuclear targeting. A functional NLS was shown to be essential for elicitor dependent and independent activation (Wirthmueller et al., 2007). The RPS4 NLS does not direct the complete pool of RPS4 to the nucleus as a cytoplasmic subpool is still present which associates with the endomembrane system, indicating a mechanism counteracting its transport into the nucleus. Like potato Rx1, the tobacco N and barley MLA proteins do not contain recognizable classical NLSs. Most classical NLSs are linear sequence motifs, but they can also be formed by the combination of residues that lie further apart. An even more complex NLS is exemplified by STAT1, for which only in the homodimer a functional NLS is formed by combining residues from two monomers (Fagerlund et al., 2002). Therefore, we cannot rule out the possibility that many R proteins, including Rx1, contain a complex NLS based on a discontinuous

stretch of residues that facilitate the import into the nucleus. Such an alternative import mechanism may explain our finding that the CC domain is required for the nuclear accumulation of the Rx1 protein and that on the other hand no autonomous import signal could be found that actively transport triple GFP fusions into the nucleus. It is also noted that in yeast only 57% of the nuclear proteins contains a classical NLS (Lange et al., 2007). The remaining 43% of nuclear proteins is thought to have passed the Nuclear Pore Complex (NPC) via alternative mechanisms like direct interactions with importin β or nucleoporins, non-classical nuclear targeting signals or interactions with proteins that do contain NLSs, the so-called piggyback mechanism (Ursula Stochaj, 1999; Christophe et al., 2000; Dostie et al., 2000; Cingolani et al., 2002; Xu and Massague, 2004; Lee et al., 2006; Chuderland et al., 2008; Lange et al., 2008).

The fact that other R proteins have been shown to interact with transcription factors provides a plausible explanation for our FRAP data and lead us to hypothesize that the CC domain of Rx1 is involved in similar processes. The strong nuclear localization of the full CC domain (aa 1-144) and fragment G (aa 45-115), composed of two α -helices, coincides with a low effective diffusion coefficients inside the nucleus. The slow diffusion cannot be explained by the size of the proteins. Molecules up to 500 kDa have been shown to diffuse unrestrictedly in the nucleus (Seksek et al., 1997; Görisch et al., 2005), and the size dependency is small (the diffusion coefficient is proportional to the inverse of the cube root of the molecular mass for globular proteins ($D \sim M^{-1/3}$)) (Reits and Neefjes, 2001). In addition it is shown, that other fusions with a similar size, but encompassing different fragments of the CC, have higher diffusion coefficients. Therefore, the most likely explanation for the diffusional behavior of the CC domain is a transient binding to a larger or less mobile component in the nucleus resulting in the sequestering of the CC domain in this cellular compartment. Similar diffusion patterns can be seen in photobleaching experiments with *Arabidopsis* HMG proteins in the nucleus, which associate transiently with chromatin and facilitate the formation of regulatory complexes (Launholt et al., 2006). Transient binding can contribute to the observed diffusion coefficient in a complex manner, dependent on binding equilibrium (k_{off} , k_{on}), relative number of binding sites and the balance with the diffusion speed (Sprague et al., 2004; Sprague and McNally, 2005; Beaudouin et al., 2006). For this reason FRAP has been used extensively to study the diffusional behavior of nuclear proteins that interact with chromatin or nuclear matrix components (Houtsmuller, 2005).

Our experiments indicate that intramolecular interactions within the Rx1 protein may play a role in explaining the nucleocytoplasmic distribution patterns. The relatively slow diffusion in the nucleus is only evident for the full CC domain (aa1-144) and fragment G (aa45-115). The effective diffusion coefficient of the CC-NB-ARC protein is much closer to the one observed for GFP (data not shown). This is reflected in the localization patterns as the CC has a stronger nuclear accumulation than the CC-NB-ARC protein (Fig. 2C & D). This suggests that there is an interaction surface on the CC exposed when the domain is expressed without the NB-ARC which directs its nuclear accumulation. In the presence of the NB-ARC this surface might not be available due to the known intramolecular interaction between the CC and NB-ARC domain (Rairdan et al., 2008). A similar phenomenon has been seen in interaction studies with other R proteins. Yeast two-hybrid experiments could only show the MLA-WRKY interactions for truncated MLA constructs missing the LRR (Shen et

al., 2007). The interaction between the N-terminal half of RPM1 and TIP49a, a nuclear factor involved in the transcriptional machinery and chromatin remodeling, could only be shown in the absence of the LRR in a yeast two-hybrid assay (Holt et al., 2002; Gallant, 2007). The other way around, the interaction between PP5 and the C-terminal part of the tomato I-2 LRR is decreased in the presence of the N-terminal domains of I-2 (de La Fuente van Bentem et al., 2005). Changes in the activation state of R proteins are characterized by conformational changes, as in a molecular switch (Takken et al., 2006). These conformational changes lead to different exposure of interacting surfaces, whereas the interactions observed in our R protein deletion constructs normally only occur in certain activated states of the R protein. Therefore, we suggest that the low diffusion coefficient of the CC domain in the nucleus, which is indicative for transient binding to large complexes like chromatin, is linked to a particular activation state of the Rx1 protein.

Our experiments show that the Rx1 protein is activated in the cytoplasm and that in the nucleus Rx1 is not able to initiate an HR. This functional differentiation between the cytoplasm and the nucleus is demonstrated by the loss of response once the PVX coat protein is targeted to the nucleus. The nuclear fraction of the cellular Rx1 pool seems unable to activate an HR in the presence of the coat protein. One explanation for this phenotype could be the requirement of a cytoplasmic complex with other protein partners, like HSP90 and SGT1, for Rx1 to be activated by the PVX coat protein. Moreover, the interaction between Rx1 and RanGAP2 could be interpreted as a guard-guardee complex following the guard hypothesis (de Wit, 2002). RanGAPs are mostly cytoplasmic proteins in correspondence with their RanGTPase activating function and plant RanGAPs contain a N-terminal WPP motif targeting them to the nuclear envelope (Jeong et al., 2005; Xu et al., 2007). The absence of the guardee in the nucleus would then make it impossible for Rx1 to sense the CP.

The threshold concentration of Rx1 required for the initiation of a hypersensitive reaction in the cytoplasm was shown to be relatively low when compared with its elicitor. Reducing the cytoplasmic pool of the Rx1 protein by adding an exogenous nuclear import signal did not have an observable effect, whereas forcing the PVX coat protein to the nucleus abrogated the HR. This shows that in our assay the Rx1 protein concentration is not rate limiting and that the Rx1 protein is effective at very low concentration levels. It is known that Rx1 confers a very efficient resistance response against PVX, better known as Extreme Resistance (Bendahmane et al., 1999), which results in inhibition of viral replication before a visible cell death response develops on the infected leaves. In natural situations the amount of Rx1 in the cell is relatively low and it is anticipated that Rx1 molecules will be largely outnumbered by PVX coat protein in infected cells. The excess of coat protein molecules in natural conditions may have resulted in only a weak affinity of Rx1 for the coat protein, leading to a relatively high threshold for the elicitor to initiate an HR.

The observation that the Rx1 protein is activated in the cytoplasm, while adding a nuclear import signal local to Rx1 itself has no effect on its functioning, illustrates limitations of adding exogenous targeting signals to redirect the localization of proteins. Localization signals do not completely exclude a protein from a compartment, but only decrease the protein concentration by increasing the rate of transport in one direction. Especially nuclear

export sequences are known to be relatively weak in comparison with nuclear localization signals (Kutay et al., 2005). As a result a small residual pool of Rx1 can still be able to mediate a cell death response. This is the most likely explanation why, in contrast to the nuclear R proteins MLA, N and RPS4, reducing the nuclear pool of Rx1 has no effect on its functioning.

Following a simple functional model the LRR recognizes the elicitor, the NB-ARC switches the protein from an inactive to an activated state and the CC or TIR domain relays the signal to downstream components. Studies on the R protein complexes (resistosomes) like N/NRIP1, RPS5/PBS1, Pto/Prf, RPM1/RPS2/RIN4, make clear that in reality the function of N-terminal domains is more complex. In these examples the N-terminal domains are thought to be adapters holding the virulence target (guardee or decoy) in the resistosome, enabling the complex to sense modifications of this target (Axtell and Staskawicz, 2003; Marathe and Dinesh-Kumar, 2003; Mucyn et al., 2006; Ade et al., 2007; Caplan et al., 2008b; van der Hoorn and Kamoun, 2008). Two studies have shown independently that the Rx1 CC domain specifically interacts with RanGTPase-Activating Protein 2 (RanGAP2) (Sacco et al., 2007; Tameling and Baulcombe, 2007). Silencing RanGAP2 in *N. benthamiana* compromised the Rx1 mediated resistance and overexpression of RanGAP2 caused the elicitor-independent activation of Rx1 CC-NB fragments. The Rx1 CC domain is also involved in intramolecular interactions with the NB-ARC-LRR domains. These interactions of the CC domain with the NB-ARC-LRR domains and with RanGAP2 are dependent on different, but overlapping, amino acid sequences (Rairdan et al., 2008). Interestingly, both interacting CC sequence stretches are located in fragment G (aa 45-116) showing nuclear accumulation in this study.

RanGAPs have a well-described role in the main nucleocytoplasmic shuttling machinery (Merkle, 2003; Stewart, 2006; Cook et al., 2007). The GTPase Ran travels to the nucleus in its GDP bound form and from the nucleus to the cytoplasm in its GTP-bound form (Gorlich et al., 1996). The interaction with RanGTPase Activating Protein in the cytoplasm strongly stimulates its GTPase activity, leading to a swift transfer from the GTP-bound to the GDP-bound form. In the nucleus the GDP is exchanged for a GTP by RanGEF, a nucleotide exchange factor. By the nature of its function, no functional RanGAP should be present in the nucleus. *Arabidopsis* RanGAP has been shown to have a cytoplasmic localization with a concentration on the nuclear envelope (Rose and Meier, 2001; Pay et al., 2002; Jeong et al., 2005). If RanGAP2 is involved in the Rx1 nucleocytoplasmic distribution is yet unknown. No evidence exists that RanGAPs themselves can shuttle proteins from the cytoplasm to the nucleus, and therefore it is not likely that it has a direct role in the transport of Rx1 to the nucleus. On the other hand, evidence exists that in yeast (*Schizosaccharomyces pombe*) RanGAP does enter the nucleus and functions there in heterochromatin assembly (Nishijima et al., 2006). The possibility that the role of RanGAP2 in Rx1 signaling is connected to its Ran activating function remains to be tested. Sacco et al. showed that targeting of RanGAP2 to the nuclear envelope by its WPP domain is not essential for the Rx1 activating overexpression phenotype (Sacco et al., 2007). Both options for RanGAP2, an active role in nucleocytoplasmic shuttling of Rx1 or a role as guardee and putative target for pathogen effectors, are still open.

By silencing SGT1 and to a lesser extent RAR1, the subcellular distribution of Rx1 shifts to the cytoplasm. Similar levels of fluorescence were measured in SGT1-silenced plants in comparison with non-silenced plants under identical microscope settings, indicating that the protein levels probably did not deviate much. In other studies both weak (Lu et al., 2003) and strong effects (Azevedo et al., 2006; Boter et al., 2007) of the absence of SGT1 on Rx1 stability have been observed. It is possible that in our experiments Rx1 is stabilized by the fusion with GFP. We observed the shift in localization towards the cytoplasm only for the full length Rx1. The CC and the CC-NB-ARC constructs were not affected. Therefore we conclude that SGT1 silencing affects Rx1 localization via the LRR. This concurs with earlier studies that connected SGT1 with the LRR of R proteins. A link between MLA mediated resistance and SGT1 and RAR1 was shown to be determined by a sequence in the LRR (Shen et al., 2003; Halterman and Wise, 2004). The LRR of MLA1 and the LRR of Bs2 have been shown to interact physically with SGT1 (Bieri et al., 2004; Leister et al., 2005). And in the formation of mammalian inflammasomes, SGT1 interacts with NOD-like receptor proteins (NLRs) via their LRR (Mayor et al., 2007).

SGT1 is thought to be a co-chaperone functioning in several pathways as a complex with RAR1, HSP90 and Hsc70 (Austin et al., 2002; Azevedo et al., 2002; Boter et al., 2007; Noel et al., 2007). SGT1 and HSP90 are the only proteins known to be absolutely essential for the Rx1 induced hypersensitive response (Peart et al., 2002b; Lu et al., 2003). RAR1 has a less pronounced function in Rx1 signaling as silencing showed in our experiments only a reduction in HR, not a complete inhibition (data not shown). The significant decrease of Rx1 in the nucleus upon SGT1 and RAR1 silencing may be explained by their role in correctly folding the Rx1 protein and allow in this way its transport through the nuclear pore. SGT1 has been shown to affect the conformation of R proteins. For example, the intramolecular interaction between the LRR and the CC-NB-ARC in the pepper R protein Bs2 can only be shown in the presence of SGT1 (Leister et al., 2005). In Rx1, however, the corresponding interaction between the N- and C-terminal domains was stable in SGT1 silenced plants (Moffett et al., 2002). Another scenario, not mutually exclusive with a co-chaperone function, is that SGT1 is involved in protein degradation processes. SGT1 has been linked to proteasomal regulation via its interaction with SKP1 in SCF-complexes and the COP9 signalosome (Azevedo et al., 2002; Liu et al., 2002). Therefore, it can not be excluded that SGT1 has a nuclear or cytoplasmic specific role in protecting Rx1 against breakdown or targeting for breakdown, leading to an Rx1 depletion from the nucleus in the absence of SGT1.

We have shown that a mutation in the P-loop of the nucleotide binding domain has a similar effect on the Rx1 localization pattern as the silencing of SGT1 and RAR1. The full-length protein is nearly excluded from the nucleus, but an Rx1 construct without the LRR is not affected. The NB-ARC module consists of three distinguishable subdomains, the NB, ARC1 and ARC2 (van der Biezen and Jones, 1998; Albrecht and Takken, 2006). Crystal structures of the homologous domains in the mammalian Apaf-1 and the nematode CED4 show that the NB and ARC domains are organized around the bound nucleotide with a conformation dependent on the type of nucleotide. The Apaf-1 structure containing ADP shows a compact conformation (Riedl et al., 2005), whereas the ATP-containing CED-4 structure shows a more open conformation with the ARC1 and ARC2 domains folded away from the NB

domain (Yan et al., 2005). Mutations in the R protein I-2 that disrupt ATPase activity, but not ATP-binding, result in constitutive activity, suggesting that the ATP-bound conformation represents the active state of the protein and the ADP-bound form the resting state (Tameling et al., 2006). The conserved lysine in the P-loop interacts with the beta- and gamma-phosphate of the nucleotide and the mutation (K176R) used in this study is known to stop ATP/ADP-binding efficiently, leading to an empty state of the protein. It was shown earlier that this P-loop mutation disrupts the intramolecular interaction between the CC and the NB-ARC-LRR, but not the interaction between the LRR and the CC-NB-ARC (Moffett et al., 2002; Leister et al., 2005). These changes in intramolecular interactions and thereby conformation may have a dramatic influence on nucleocytoplasmic distribution.

An interesting example of such a conformational dependent localization is given by the Class II Transactivator (CIITA) protein. CIITA has a role in the mammalian adaptive immune responses by regulating the activation of the Major Histocompatibility Complex Class I and Class II genes expression. The overall domain structure of CIITA resembles the structure of the NB-LRR R proteins and is composed of a central GTP-binding domain and a C-terminal LRR. The domains involved in transactivation are positioned on the N-terminal half of the protein. CIITA does not bind DNA by itself, but it serves as a scaffold, bringing together factors that together activate transcription (Sisk et al., 2001). The subcellular localization of the protein is determined by a rather complex interplay of multiple nuclear import and export sequences that are exposed or buried depending on the conformation and dimerization state of the protein (Kretsovali et al., 2001; Raval et al., 2003). Similar to what we find for Rx1, CIITA's nucleotide-binding is necessary for its nuclear localization. A mutation disabling GTP-binding causes CIITA to become predominantly cytoplasmic via enhanced export from the nucleus (Harton et al., 1999; Bewry et al., 2007). The fact that GTP-binding is essential for self-association of CIITA indicates involvement of the self-association in the regulation of its nucleocytoplasmic trafficking. Binding of the nuclear export receptor CRM1 has been mapped to the N-terminal 114 amino acids of the CIITA protein (Kretsovali et al., 2001). Raval et al. propose a model in which the NES is buried and multiple NLSs are exposed in the dimerized active form (Raval et al., 2003). No GTPase activity has been shown for wild-type CIITA. However, GTPase activity can be introduced by a single point mutation (L465Q) which results in lower levels of activity. The GTPase activity keeps the protein in its inactive, monomeric and cytoplasmic state (Harton et al., 1999; Bewry et al., 2007).

Similar to CIITA, our data support a model in which both folding states and domain interactions play a role in controlling the nucleocytoplasmic distribution of Rx1. Our findings may have also implications for various other NB-LRR proteins. Many NB-LRR proteins do not possess classical targeting signals and require also SGT1, Rar1 and a functional P-loop for mediating resistance. Also many Avr proteins have, like the PVX coat protein, a relatively small molecular mass and may diffuse freely between nucleus and cytoplasm. Here we show that although both Rx1 and its elicitor have a dual location, Rx1 can only be activated by its cognate elicitor in the cytoplasm. The nucleus lacks essential components to initiate the activation cycle of Rx1. Future research will focus on the role of RanGAP2 in the cytoplasm and the function of Rx1 in the nucleus by identifying the nuclear

components responsible for the transient binding of the CC domain to large nuclear complexes.

Acknowledgements

We like to express our gratitude to David Baulcombe (Sainsbury Laboratory, John Innes Centre, Norwich) for providing the PVX amplicons pgR105 en pgR106, Matthieu Joosten (Laboratory of Phytopathology, Wageningen University, the Netherlands) for the VIGS constructs for SGT1 and RAR1, Peter Moffett (Boyce Thompson Institute, Ithaca, NY) for the CC deletion constructs, and the Dutch Technology Foundation (STW), the EU project Insight inside and Bioexploit for funding. LNS and AJP also acknowledge CNCSIS grant PN2-ID-249 168/2007.

Materials and methods

Sequence characterization

Similarity searches were performed with BLAST using Blosum62. Patterns, profiles and domain recognition were scanned with InterPro (Quevillon et al., 2005) and CDART (Geer et al., 2002) which integrates Pfam, Prints, Prodom, SMART, TIGR and Prosite databases. Secondary structure predictions were performed with methods best ranked in CASP4: SOPMA (Geourjon and Deleage, 1995), GOR IV (Garnier et al., 1996), PsiPred (Jones, 1999), Jpred (Cole et al., 2008), HNN (Guermeur, 1997), PROF (Ouali and King, 2000). Specialized β -turn prediction was performed with BETATURNS (Chou, 2000). Domain linker prediction was performed using DLP (Miyazaki et al., 2002). Intrinsic disorder prediction was performed with DisEMBL (Linding et al., 2003).

Expression cassette construction

35S promoter constructs: The Rx1 encoding sequences, including introns, were inserted in the plant expression vector pRAP (Schouten et al., 1997). Rx1 was amplified using the primers 5GpRxbn (5'-TTT TTT GGA TCC ATG GCT TAT GCT GCT GTT ACT TCC C) and Rxrev (5'-GAT AGC GTC GAC CAC CTT AAC TAC TCG CTG CA) and transferred to NcoI-SalI digested pRAP, resulting in pRAP:Rx1. *Endogenous promoter constructs:* Comparison of the genomic sequence of Rx1 (Genbank AJ011801) and the homologous Gpa2 (Genbank AF195939) showed extensive similarity in the promoter and transcription terminator regions. These regions of similarity were taken as the basis for creation of vectors carrying endogenous transcription cassettes in a pUC19 background. The Rx1 3'-UTR (transcription termination) region (274bp) was amplified from pBIN:BAC-Rx1 (pBIN+:RGC4, (van der Vossen et al., 2000)) template using the primers 5UTRkp (5'-TGG TAC CTT CTG CAG CGA GTA GTT AAG GTG TTC TGA GGA C-3') and 3UTRrev (5'-CTT AAT TAA CCC GGG AGA TTG AGG ACT CCC AAG AAA GG-3') and cloned as KpnI-PacI fragment into pRAP:YFP, replacing the Tnos. The Rx1 promoter region (2804 bp between XbaI and ATG start codon) was constructed in two steps. First the region between the DraIII site (nt 1429) and the start codon was amplified from pBIN:BAC-Rx1 (pBIN+:RGC4) using the primers bRxAdelf (5'-GAG ATT CAC TAT GTG CAT CAC CCA C-3') and RxbnREV (5'-AGC ATA AGC CAT GGA TCC AAA AAA TAG AAA TAT CTC T-3'). This 1396 bp DraIII-

NcoI fragment was cloned alongside the 1431 bp AscI-DraIII fragment of pBIN:BAC-Rx1 into pRAP:YFP with Rx1 3'-UTR, from which the 35S region was removed by AscI-NcoI digestion. The resulting vector pRXI:YFP drives expression of the reporter gene. R-gene constructs were exchanged between the pRAP and pRXI using the unique NcoI and PstI sites. *Cloning R-gene fusion partners:* The YFP (enhanced YFP, Clontech), CFP (enhanced CFP, Clontech) and GFP (enhanced GFP, Clontech)(Yang et al., 1996) reporter genes were amplified by PCR using the primers 5CFPsbN (5'-GTC GAC GGA TCC ATG GTG AGC AAG GGC GAG GAG CTG TTC-3') and 3CFPsrk (5'-AGG TAC CTT AGC TCA TGA CTG ACT TGT AGA GCT CGT CCA TGC CGA GAG-3'). PCR products were cloned as NcoI-KpnI fragment in the vector pRAP between the 35S promoter and Tnos termination regions. *C-terminal fusions:* In the plant expression vector pRAP:cbp (cbp is a stuffer fragment) the SphI-PacI segment is replaced by the annealed oligos CBPY1+2 with SphI-NcoI overhang (3CBPY1 5'-CAC ACC GTA TGC GGC CGC TGC AGT CGA CGG TGA TGT GGT 3CBPY2 5'-CAT GAC CAC ATC ACC GTC GAC TGC AGC GGC CGC ATA CGG TGT GCA TG) and the YFP-Tnos NcoI-PacI fragment. In the resulting pRAP:cbp-YFP, the fluorescent protein is preceded by NotI and Sall. The Sall-SstI GFP fragment, derived from pCR2.1:GFP is cloned into pRAP:cbp-YFP. In the resulting vector an extra BamHI site is available for fusions. To create multimeric GFP molecules in the vector pRAP (AscI-PacI) the 35S-GFP AscI-BspHI and the GFP-Tnos NcoI-PacI fragment were joined, resulting in pRAP:GFP-GFP. By repeating the procedure with the AscI-BspHI 35S-GFP-GFP fragment and the NcoI-PacI GFP-Tnos a triple GFP fusion was constructed. In pRAP:YFP, opened between the SstI and KpnI site the annealed 3YN1+2 adapter (3YN1 5'-CTA CAA GTC AGC GGC CGC ATA AGG TAC, 3YN2 5'-CTT ATG CGG CCG CTG ACT TGT AGA GCT) was inserted. Then the NcoI-NotI YFP fragment is introduced into the pRAP:scFv3myc (Schouten et al., 1996). The resulting pRAP:YFPmyc carries an XbaI site after the myc-tag encoding sequence. In the vector pRAP:YFPmyc the GFP sequence is introduced via the NcoI-SstI sites. *Targeting signals:* The SV40 NLS (Fig. 7)(Lanford and Butel, 1984; Haasen et al., 1999), was generated as an annealed oligo (SV1:5'-CAT GGG CCC TAA AAA GAA GCG TAA GGT TGA GGA CCC TGG ATC CGT GAA TTC TG; SV2:5'-CTA GCA GAA TTC ACG GAT CCA GGG TCC TCA ACC TTA CGC TTC TTT TTA GGG CC) with NcoI-NheI overhang. It was cloned into pRAP:YFP-HA8 (NcoI-NheI), resulting in pRAP:SV-HA8. In the mutated control version pRAP:sv-HA8 the oligo pair SVmut1 (5' CAT GGG CCC TAA AAA CAA GCG TAA GGT TGA GGA CCC TGG ATC CGT GAA TTC TG) and SVmut2 (5' CTA GCA GAA TTC ACG GAT CCA GGG TCC TCA ACC TTA CGC TTG TTT TTA GGG CC) are used. The vectors pRAP:GFP-SV-HA8(NLS) and pRAP:GFP-sv-HA8(nls) were created by introduction into pRAP:SV-HA8 and pRAP:sv-HA8 (AscI-NcoI) of the 35S:GFP (AscI-BspHI fragment. As a NES the sequence from PKI (Wen et al., 1995) was used (Fig. 7). The annealed oligo pair PK1+PK2 (PK1: 5'-CTA CAA GGC CAT GGG TAA CGA GCT TGC ATT AAA GCT CGC TGG TCT TGA TAT TAA CAA GGG ATC CGG TG; PK2: 5'-CTA GCA CCG GAT CCC TTG TTA ATA TCA AGA CCA GCG AGC TTT AAT GCA AGC TCG TTA CCC ATG GCC TTG TAG AGC T) with SstI-NheI overhang was cloned in between the same sites in pRAP:GFP-HA8. The mutated version was inserted in a similar way. The resulting plasmids are pRAP:GFP-PK-HA8 (NES) and pRAP:GFP-pk-HA8 (nes). Versions of multimerized GFP with NLS/nls were created by introducing in pRAP:cbp-GFP-GFP (AscI-BamHI), the AscI-BamHI 35S:GFP-NLS or GFP-nls fragment. From these pRAP:GFP-NLS/nls-GFP-GFP constructs the AscI-BspHI segment was introduced into pRAP:NES/nes-HA8 resulting in the vectors

pRAP:GFP-NLS/nls-GFP-GFP-NES/nls-HA8. *R-gene derived segments:* In pRAP:Rx1, digested with AscI-NcoI, the 35S:YFP (AscI-SstI) fragment is ligated alongside the annealed oligo "112" (linker 12for: AGC TCT ACA AGG GCG GCG GAA GTG GAG GCG GAT CCG GGG GAG GCA GCA TG, linker12rev: CTG CCT CCC CCG GAT CCG CCT CCA CTT CCG CCG CCC TTG TAG (amino acid sequence "GGSG GSGG GS"). From the resulting pRAP:YFP-Rx1 an N-terminal GFP version is prepared by exchange of YFP for GFP via NcoI-SstI. Via NcoI-PstI the pRXI:GFP-Rx1 was prepared. To allow C-terminal fusions the Rx1 cDNA was prepared. The primer combination 5RxexFor (5'- CAA AGA GAT TGA TTT CGG GGG) and 3Rxnot (5'- GCT TCT TGC CGC AAT AAT GTC GAG GGT GCG GCC GCT TAA GGT ACC AG) amplified the Rx1 C-terminal end, which was then cloned into pCR2.1Topo and recloned as BspEI-KpnI fragment in pRAP NcoI-KpnI and the Rx1 NcoI-BspEI fragment. Finally in the cDNA version of pRAP:Rx1, digested with NotI-PacI, the GFP-Tnos from pRAP:cbp-GFP was inserted, leading to pRAP:Rx1-GFP. To create in the GFP-NLS/nls-Rx1 constructs, pRAP:GFP-Rx1 was opened with NcoI-BamHI and the GFP-SV(+) or GFP-sv(-) segments from pRAP:GFP-SV(+) or -sv(-) (NcoI-BamHI) were ligated. For introduction of the NES the AscI-BamHI fragment from pRAP:GFP-PK(+) or -PK(-) were introduced in AscI-BamHI digested pRAP:GFP-Rx1. The GFP-NLS/nls-Rx1 insert from the pRAP versions were introduced into pRXI using NcoI-PstI. For the GFP-NES/nls-Rx1 in pRXI:GFP-Rx1 the SstI-ClaI fragment was exchanged for the SstI-ClaI fragments from pRAP:GFP-PK(+)-Rx1 and GFP-PK(-)-Rx1. The P-loop mutant Rx1 K176R was generated by splicing by overlap extension of CC-NB-ARC fragments amplified with the primer combinations 5GpRxbn + 3LysRrev (5'-AGT TGT TCT CCC GAT GCC TCC CAT CCC) and 5LysRfor (5'-GGA GGC ATC GGG AGA ACA ACT TTG GCT ACA) and 3NBSeRev (5'-TGG TAC CTT AAG AAT TCA TGT TTC GAG CTT CCC TCA AAC AG) on Rx1 template. This R1-3K176R segment was introduced into pCR4-Topo (Invitrogen). Then the K176R mutation was cloned into pBAD:Rx1 (see LRR section) as fusion of BglII-StuI and StuI-ClaI fragments. The full sized pRAP:GFP-Rx1 K176R was generated by introducing into pRAP:GFP-Rx1 the HpaI-ClaI fragment from pBAD:Rx1K176R. *LRR constructs:* The sequence encoding the Rx1 LRR domain from amino acid 473 was amplified from cDNA using the primers For-LRRrx-1 (5'-ATG AAT TTT GTG AAT GTT ATC AGA GG) and Rev-LRRrx-1 (5'-CTC GAC ATT ATT GCG GCA AGA AGC) and then cloned into pBAD-topo (Invitrogen). In this vector the LRR is preceded by the extra amino acids "MGSGS GDDDD KLAL-". A full size pBAD:Rx1 was obtained by insertion of the NcoI-ClaI fragment from pRAP:Rx1. From pBAD:LRR the NcoI-ClaI fragment was cloned into pRAP:Rx1 (NcoI-ClaI) resulting in pRAP:LRR. In this vector the GFP sequence was introduced as NcoI-BamHI fragment taken from pRAP:GFP-Rx1. GFP and the LRR are connected by the linker "GGSG GSGD DDDKL AL". *CC-NB-ARC constructs:* The Rx1 CC-NB-ARC (aa 1-474) encoding sequence was amplified from Rx1 template using the primer set 5GpRxbn and NBSeRev (5'-TGG TAC CTT AAG AAT TCA TGT TTC GAG CTT CCC TCA AAC AG) and then cloned into pRAP via NcoI-KpnI restriction sites. A C-terminal fusion of YFP was accomplished by inserting into the AscI-NcoI cut pRAP:YFP the AscI-EcoRI fragment of pRAP:CC-NB-ARC and a linker peptide encoding EcoRI-NcoI fragment annealed from Ctyfp1 (5'-AAT TCT GGA GGT TCT GGT GGC GGA GGC TCA GGC GGT GGT GGA AG) and Ctyfp2 (5'-CAT GCT TCC ACC ACC GCC TGA GCC TCC GCC ACC AGA ACC TCC AG). In the resulting plasmid pRAP:CC-NB-ARC-YFP the linker "SGSG GSGG GGS" connects the CC-NB-ARC and YFP. pRAP:CC-NB-ARC-GFP was created by exchanging YFP for GFP from pRAP:NB-ARC-GFP

via ApaLI-PacI. The vector pRAP:YFP-CC-NB-ARC is created by introducing the StuI-PacI fragment from pRAP:CC-NB-ARC into pRAP:YFP-Rx1. The GFP-CC-NB-ARC K176R AscI-ApaLI fragment and the ApaLI-PacI fragment from pRAP:NB-ARC-GFP were inserted into AscI-PacI digested pRAP resulting in pRAP:GFP-CC-NB-ARC-GFP K176R. Then the AscI-HpaI fragment was replaced by AscI-HpaI from pRAP:Rx1 leading to pRAP:CC-NB-ARC-GFP K176R. To amplify the NB-ARC region of Rx1 the primer pair 5NBSf (5'- GAC CAT GGT TGG CCG TGA AAA TGA ATT TGA G) and ApaLRev (5'- GGT ACC TTA CTG CAT GGA TTG TGC ACA TGA AT) were used. The amplified product was cloned as NcoI-KpnI fragment into pRAP. To enable C-terminal fusions the annealed oligo AN1 +AN2 (AN1: 5'- TGC ACA ATC CAT GCA GGC GGC CGC TTA AGG TAC ; AN2: 5'-CTT AAG CGG CCG CCT GCA TGG ATT G) was introduced via ApaLI-KpnI. The resulting pRAP:NB-ARC-an carries a NotI site and allowed cloning of the GFP-Tnos NotI-PacI fragment to yield pRAP:NB-ARC-GFP. In this construct the Rx1 segment extends from amino acid 142 to 489.

CC deletion construct: To create the CC-deletion of Rx1-GFP, the AscI-ApaLI fragment from pRAP:NB-ARC-an was joined alongside the ApaLI-PacI fragment from pRAP:Rx1 in AscI-PacI digested pRAP resulting in pRAP:NB-ARC-LRR-GFP. ***CC-GFP and CC fragments:*** The CC domain of Rx1 was amplified from Rx1 template using the primer combination 5GpRxsbn and 3CCnot (5'- GTG GTA CCT TAA GCG GCC GCA CCA ACC ATT ATA TTC TCG GGC TGC). The fragment was cloned via NcoI-KpnI and into pRAP. Then after NotI-PacI digestion the GFP-Tnos fragment from pRAP:cbp-GFP was introduced leading to pRAP:CC-GFP. To create pRAP:CC-3GFP the AscI-BspHI CC-GFP fragment was fused to the NcoI-PacI GFP2-Tnos segment from pRAP:cbp-GFP2 in the vector pRAP (AscI-PacI). CC fragment constructs were prepared by amplification of Rx1 template with the following primer combinations. A: 5GpRxsbn and Rev-BamHI-AC (5'- AGG ATC CCA TTA TAT TGC AG). B: 5GpRxsbn and Rev-BamHI-B (5'- AGG ATC CAG TAA GTT CCA TTG). C: For-nco-CD (5'-TAC CAT GGA ACT TAC TGG ATG TG) and Rev-BamHI-AC. D2: For-nco-CD and Rev-BamHI-ED2 (5'-AGG ATC CAT TTC TTG ATT CCG AG). E1n: For-nco-EG (5'- TAC CAT GGG CGA TCA TGA GG) and Rev-BamHI-ED1 (5'- AGG ATC CCC TGC TTC TTT CCT C). F: For-nco-F (5'-TAC CAT GGC ACA GAA TTT GGA GG) and Rev-BamHI-GF (5'-AGG ATC CGC TGT CCG ATG TTG C). G: For-nco-EG and Rev-BamHI-GF. All these PCR fragments were ligated as NcoI-BamHI fragment with the pRAP AscI-NcoI fragment into the AscI-BamHI digested pRAP:cbp-GFP resulting in CC fragments with a C-terminal GFP fusion. The pRAP:R1-G-GFP3 was constructed by inserting into AscI-NcoI digested pRAPcbp-GFP2 the AscI-BspHI fragment from pRAP:R1-G-GFP.

PVX coat protein constructs: The PVX coat protein (CP) sequence was amplified from the PVX amplicons pGR106 (Jones et al., 1999) containing cDNA of the Rx1-avirulent PVX strain UK3 and pGR105 containing cDNA of the Rx1-resistance breaking strain HB (Goulden et al., 1993) using the primers 5UK3cp (5'-TCC ATG GGC GGT GGA GTC ATG AGC GCA CCA GCT AGC ACA ACA CAG CC) and 3UK3cp (5'-AGG TAC CTG CGG TTA TGG TGG TGG TAG AGT GAC AAC AGC) for cp106 and 5HBcp (5'-TCC ATG GGC GGT GGA GTC ATG ACT ACG CCA GCC AAC ACC ACT C) and 3HBcp (5'-AGG TAC CTG CGG TTA TGG TGG GGG TAG TGA GAT AAC AGC) for cp105. The products were cloned as BspHI-KpnI fragments into the NcoI-KpnI sites of pRAP. The N-terminal CFP fusions were constructed by inserting into pRAP AscI-KpnI the AscI-BspHI 35S:CFP and the NcoI-KpnI CP105 or CP106, resulting in pRAP:CFP-CP105 and pRAP:CFP-CP106. To provide the CP105 and CP106 with nuclear localization signals in a first step the BspHI-KpnI cp fragments were

cloned into pRAP:N-GFP yielding pRAP:N-cp105 and pRAP:N-cp106. Then these vectors were opened with *AscI*-*SpeI* followed by introduction of the *AscI*-*NheI* fragment from pRAP:GFP-SV(+) / -SV(-) / -PK(+) and -PK(-) thereby creating pRAP:GFP-NLS-CP105 and -CP106; pRAP:GFP-nls-CP105 and -CP106; pRAP:GFP-NES-cp105 and -cp106 and pRAP:GFP-nes-cp105 and -cp106.

Transient expression

For transient expression the expression cassettes from the pRAP constructs were cloned via *AscI*-*PacI* into the binary vector pBIN+ (van Engelen et al., 1995). *Agrobacterium tumefaciens* MOG101 (Hood et al., 1993) harboring the individual binary vectors was grown at 28 °C in YEB medium (per liter: 5 g. beef extract, 1 g. yeast extract, 5 g. peptone, 5 g. sucrose, 2 mL 1 M MgSO₄) with 50 mg/L kanamycin and 20 mg/L rifampicin. The bacteria were spun down and resuspended in MMAi infiltration medium (per liter: 5 g. MS salts, 1.95 g. MES, 20 g. sucrose, 1 mL 200 mM acetosyringone (Sigma-Aldrich)). The bacterial solution was diluted to an OD₆₀₀ of 0.5 in MMAi and incubated at room temperature for two hours before infiltration in *N. benthamiana* leaves. Infiltrated leaves were harvested two or three days after infiltration (depending on the expressed construct) for microscopy or protein extraction.

Confocal microscopy and FRAP

Images of the fluorescent protein constructs in *N. benthamiana* epidermal cells were obtained using a Zeiss LSM 510 confocal microscope (Carl-Zeiss, Germany) with a 40x 1.2 NA water corrected objective. For GFP imaging the 488nm line from an Argon laser was used for excitation and a 505-550 nm band pass filter for the emission. For CFP excitation the 458 nm line of a HeNe laser and for emission a 470-500 nm band pass filter were used. Chlorophyll emission was detected through a 650 nm long pass filter. Fluorescence intensities were quantified using the Java application ImageJ (Abramoff et al., 2004). FRAP experiments were performed with the Zeiss LSM 510 confocal microscope. A square 12 μm² region in the nucleus was bleached with the 488 nm laser line at 75% power and fluorescence recovery monitored with 60 ms intervals for 5 seconds. 20 to 35 recovery curves were acquired per construct with LSM 510 software. Half times ($\tau_{1/2}$) were determined by fitting bleaching corrected and normalized curves against a model for 2D diffusion (ref.). Effective diffusion coefficients (D_{eff} , μm²s⁻¹) were derived from $\tau_{1/2}$ via $D_{\text{eff}} = A/(4 \tau_{1/2})$ in which A represents the bleached surface. Statistical analyses were performed with the software package SPSS (SPSS for windows 12.01).

Protein extraction and immunodetection

In order to extract protein from leaf samples 0.5 g. leaf material was ground in extraction buffer (150 mM NaCl, 50 mM Tris-HCl pH 7.5, 1 mM EDTA, 10% glycerol, 10 mM DTT, 2% PVPP, and 0.5 mg/mL pefabloc SC protease inhibitor (Roche)). After spinning down the cell debris (5 min, 16000 rpm) the supernatans was combined with 4x laemmli buffer and the proteins were loaded on SDS-PAGE Tris-glycine gels for electrophoretic separation. SDS-PAGE separated protein was either detected with Coomassie Brilliant Blue staining or dry-blotted on a nitrocellulose membrane (immobilon-p, Millipore). GFP fused proteins were detected on western blots by HRP conjugated polyclonal rabbit anti-GFP (Novus, NB100-1184) or polyclonal rabbit anti-GFP (Abcam, 290-50) as primary antibody and peroxidase

conjugated Goat anti-Rabbit IgG (Jackson, 111-035-045) as secondary antibody. The peroxidase activity was visualized with ECL western blotting substrate (GE healthcare).

Plant transformation and virus resistance assay

The susceptible potato Line V was used for *Agrobacterium*-mediated plant transformation (van Engelen et al., 1994). Genomic DNA was extracted using the Dneasy plant mini kit (Qiagen GmbH, Hilden, Germany) and used for PCR analysis to check the incorporation of the transgenes in the plant genome. Primary transformants were used in the virus resistance assay. To obtain infectious virus particles, leaves of *N. benthamiana* were agro-infiltrated with the PVX amplicons pGR106 and pGR105. Systemically infected leaf material was homogenized in 10 ml of 50 mM sodiumphosphate buffer pH7, including 1 mM Pefabloc. Twenty µl was used for sap inoculation of the four lower leaves of four week-old transgenic potato plants dusted with carborundum powder. Inoculations were done *in triplo*. Infected plants were grown in the greenhouse at 20°C and 16 hours of light. Three weeks after infection leaf discs sampled from compound leaves of the apex were homogenized as described above and the virus concentration was determined using DAS-ELISA (Maki-Valkama et al., 2000). Plates were coated with a 1:1000 dilution of a polyclonal antibody against PVX to bind the antigen and a second polyclonal antibody against PVX conjugated with alkaline phosphatase was used for detection.

Virus induced gene silencing

Three week old *N. benthamiana* plants were co-infiltrated with *A. tumefaciens* GV3101 containing TRV1 and TRV2 vectors (Hellens et al., 2000; Ratcliff et al., 2001). The TRV silencing vectors TRV2:SGT1 and TRV2:RAR1 were kindly provided by M. Joosten (Gabriels et al., 2006; Gabriels et al., 2007). The empty TRV2 vector was used as negative control and TRV2:PDS (phytoene desaturase, (Ratcliff et al., 2001)) was used to visualize the silencing progression. Three or four weeks after inoculation the upper leaves of the plants were used for transient expression of the fluorescent Rx1 proteins.

References

- Abramoff, M.D., Magalhaes, P.J. and Ram, S.J.** (2004). Image processing with imageJ. *Biophotonics International* **11**, 36-41.
- Ade, J., Deyoung, B.J., Golstein, C. and Innes, R.W.** (2007). Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease. *Proc Natl Acad Sci U S A* **104**, 2531-2536.
- Alber, F., Dokudovskaya, S., Veenhoff, L.M., Zhang, W., Kipper, J., Devos, D., Suprpto, A., Karni-Schmidt, O., Williams, R., Chait, B.T., Sali, A. and Rout, M.P.** (2007). The molecular architecture of the nuclear pore complex. *Nature* **450**, 695-701.
- Albrecht, M. and Takken, F.L.** (2006). Update on the domain architectures of NLRs and R proteins. *Biochem Biophys Res Commun* **339**, 459-462.
- Austin, M.J., Muskett, P., Kahn, K., Feys, B.J., Jones, J.D.G. and Parker, J.E.** (2002). Regulatory role of SGT1 in early R gene-mediated plant defenses. *Science* **295**, 2077-2080.
- Axtell, M.J. and Staskawicz, B.J.** (2003). Initiation of RPS2-specified disease resistance in Arabidopsis is coupled to the AvrRpt2-directed elimination of RIN4. *Cell* **112**, 369-377.
- Azevedo, C., Betsuyaku, S., Peart, J., Takahashi, A., Noel, L., Sadanandom, A., Casais, C., Parker, J. and Shirasu, K.** (2006). Role of SGT1 in resistance protein accumulation in plant immunity. *Embo J* **25**, 2007-2016.
- Azevedo, C., Sadanandom, A., Kitagawa, K., Freialdenhoven, A., Shirasu, K. and Schulze-Lefert, P.** (2002). The RAR1 interactor SGT1, an essential component of R gene-triggered disease resistance. *Science* **295**, 2073-2076.
- Batten, J.S., Yoshinari, S. and Hemenway, C.** (2003). Potato virus X: a model system for virus replication, movement and gene expression. *Molecular Plant Pathology* **4**, 125-131.
- Beaudouin, J., Mora-Bermudez, F., Klee, T., Daigle, N. and Ellenberg, J.** (2006). Dissecting the Contribution of Diffusion and Interactions to the Mobility of Nuclear Proteins. *Biophys. J.* **90**, 1878-1894.
- Bendahmane, A., Kanyuka, K. and Baulcombe, D.C.** (1999). The Rx gene from potato controls separate virus resistance and cell death responses. *Plant Cell* **11**, 781-791.
- Bewry, N.N., Bolick, S.C., Wright, K.L. and Harton, J.A.** (2007). GTP-dependent Recruitment of CIITA to the Class II Major Histocompatibility Complex Promoter. *J Biol Chem* **282**, 26178-26184.
- Bieri, S., Mauch, S., Shen, Q.H., Peart, J., Devoto, A., Casais, C., Ceron, F., Schulze, S., Steinbiss, H.H., Shirasu, K. and Schulze-Lefert, P.** (2004). RAR1 positively controls steady state levels of barley MLA resistance proteins and enables sufficient MLA6 accumulation for effective resistance. *Plant Cell* **16**, 3480-3495.
- Boter, M., Amigues, B., Peart, J., Breuer, C., Kadota, Y., Casais, C., Moore, G., Kleanthous, C., Ochsenbein, F., Shirasu, K. and Guerois, R.** (2007). Structural and Functional Analysis of SGT1 Reveals That Its Interaction with HSP90 Is Required for the Accumulation of Rx, an R Protein Involved in Plant Immunity. *Plant Cell* **19**, 3791-3804.
- Burch-Smith, T.M., Schiff, M., Caplan, J.L., Tsao, J., Czymmek, K. and Dinesh-Kumar, S.P.** (2007). A Novel Role for the TIR Domain in Association with Pathogen-Derived Elicitors. *PLoS Biol* **5**, e68.
- Caplan, J., Padmanabhan, M. and Dinesh-Kumar, S.P.** (2008a). Plant NB-LRR immune receptors: from recognition to transcriptional reprogramming. *Cell Host Microbe* **3**, 126-135.
- Caplan, J.L., Mamillapalli, P., Burch-Smith, T.M., Czymmek, K. and Dinesh-Kumar, S.P.** (2008b). Chloroplastic Protein NRIP1 Mediates Innate Immune Receptor Recognition of a Viral Effector. *Cell* **132**, 449-462.
- Chou, K.C.** (2000). Prediction of tight turns and their types in proteins. *Anal Biochem* **286**, 1-16.
- Christophe, D., Christophe-Hobertus, C. and Pichon, B.** (2000). Nuclear targeting of proteins: how many different signals? *Cell Signal* **12**, 337-341.
- Chuderland, D., Konson, A. and Seger, R.** (2008). Identification and Characterization of a General Nuclear Translocation Signal in Signaling Proteins. *Mol. Cell* **31**, 850-861.
- Cingolani, G., Bednenko, J., Gillespie, M.T. and Gerace, L.** (2002). Molecular basis for the recognition of a nonclassical nuclear localization signal by importin beta. *Mol Cell* **10**, 1345-1353.
- Cokol, M., Nair, R. and Rost, B.** (2000). Finding nuclear localization signals. *EMBO Rep* **1**, 411-415.
- Cole, C., Barber, J.D. and Barton, G.J.** (2008). The Jpred 3 secondary structure prediction server. *Nucleic Acids Res* **36**, W197-201.
- Cook, A., Bono, F., Jinek, M. and Conti, E.** (2007). Structural biology of nucleocytoplasmic transport. *Annu Rev Biochem* **76**, 647-671.
- Cruz, S.S., Chapman, S., Roberts, A.G., Roberts, I.M., Prior, D.A.M. and Oparka, K.J.** (1996). Assembly and movement of a plant virus carrying a green fluorescent protein overcoat. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 6286-6290.
- de La Fuente van Bentem, S., Vossen, J.H., de Vries, K.J., van Wees, S., Tameling, W.I., Dekker, H.L., de Koster, C.G., Haring, M.A., Takken, F.L. and Cornelissen, B.J.** (2005). Heat shock protein 90 and its co-chaperone protein phosphatase 5 interact with distinct regions of the tomato I-2 disease resistance protein. *Plant J* **43**, 284-298.
- de Wit, P.** (2002). Plant biology - On guard. *Nature* **416**, 801-803.
- Deslandes, L., Olivier, J., Peeters, N., Feng, D.X., Khounlotham, M., Boucher, C., Somssich, I., Genin, S. and Marco, Y.** (2003). Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted

to the plant nucleus. *Proc Natl Acad Sci U S A* **100**, 8024-8029.

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

Dostie, J., Ferraiuolo, M., Pause, A., Adam, S.A. and Sonenberg, N. (2000). A novel shuttling protein, 4E-T, mediates the nuclear import of the mRNA 5' cap-binding protein, eIF4E. *Embo J* **19**, 3142-3156.

Eisenberg, D., Marcotte, E.M., Xenarios, I. and Yeates, T.O. (2000). Protein function in the post-genomic era. *Nature* **405**, 823-826.

Enright, A.J. and Ouzounis, C.A. (2001). Functional associations of proteins in entire genomes by means of exhaustive detection of gene fusions. *Genome Biol* **2**, RESEARCH0034.

Fagerlund, R., Melen, K., Kinnunen, L. and Julkunen, I. (2002). Arginine/lysine-rich nuclear localization signals mediate interactions between dimeric STATs and importin alpha 5. *J Biol Chem* **277**, 30072-30078.

Gabriels, S.H., Takken, F.L., Vossen, J.H., de Jong, C.F., Liu, Q., Turk, S.C., Wachowski, L.K., Peters, J., Witsenboer, H.M., de Wit, P.J. and Joosten, M.H. (2006). CDNA-AFLP combined with functional analysis reveals novel genes involved in the hypersensitive response. *Mol Plant Microbe Interact* **19**, 567-576.

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

Gallant, P. (2007). Control of transcription by Pontin and Reptin. *Trends Cell Biol.* **17**, 187-192.

Garnier, J., Gibrat, J.F. and Robson, B. (1996). GOR method for predicting protein secondary structure from amino acid sequence. *Methods Enzymol* **266**, 540-553.

Geer, L.Y., Domrachev, M., Lipman, D.J. and Bryant, S.H. (2002). CDART: protein homology by domain architecture. *Genome Res* **12**, 1619-1623.

Geourjon, C. and Deleage, G. (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Comput Appl Biosci* **11**, 681-684.

Görsch, S.M., Wachsmuth, M., Toth, K.F., Lichter, P. and Rippe, K. (2005). Histone acetylation increases chromatin accessibility. *J. Cell Sci.* **118**, 5825-5834.

Gorlich, D. and Kutay, U. (1999). Transport between the cell nucleus and the cytoplasm. *Annu Rev Cell Dev Biol* **15**, 607-660.

Gorlich, D., Pante, N., Kutay, U., Aebi, U. and Bischoff, F.R. (1996). Identification of different roles for RanGDP and RanGTP in nuclear protein import. *Embo J* **15**, 5584-5594.

Goulden, M.G., Kohm, B.A., Cruz, S.S., Kavanagh, T.A. and Baulcombe, D.C. (1993). A Feature Of The Coat Protein Of Potato Virus-X Affects Both Induced Virus-Resistance In Potato And Viral Fitness. *Virology* **197**, 293-302.

Haasen, D., Kohler, C., Neuhaus, G. and Merkle, T. (1999). Nuclear export of proteins in plants: ATXPO1 is the export receptor for leucine-rich nuclear export signals in *Arabidopsis thaliana*. *Plant J* **20**, 695-705.

Halterman, D.A. and Wise, R.P. (2004). A single-amino acid substitution in the sixth leucine-rich repeat of barley MLA6 and MLA13 alleviates dependence on RAR1 for disease resistance signaling. *Plant J* **38**, 215-226.

Harton, J.A., Cressman, D.E., Chin, K.C., Der, C.J. and Ting, J.P. (1999). GTP binding by class II transactivator: role in nuclear import. *Science* **285**, 1402-1405.

Heese, A., Hann, D.R., Gimenez-Ibanez, S., Jones, A.M., He, K., Li, J., Schroeder, J.I., Peck, S.C. and Rathjen, J.P. (2007). The receptor-like kinase SERK3/BAK1 is a central regulator of innate immunity in plants. *Proc Natl Acad Sci U S A* **104**, 12217-12222.

Hellens, R.P., Edwards, E.A., Leyland, N.R., Bean, S. and Mullineaux, P.M. (2000). pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. *Plant Mol.Biol.* **42**, 819-832.

Holt, B.F., Boyes, D.C., Ellerstrom, M., Siefers, N., Wiig, A., Kauffman, S., Grant, M.R. and Dangl, J.L. (2002). An evolutionarily conserved mediator of plant disease resistance gene function is required for normal *Arabidopsis* development. *Dev. Cell* **2**, 807-817.

Hood, E.E., Gelvin, S.B., Melchers, L., S. and Hoekema, A. (1993). New *Agrobacterium* helper plasmids for gene transfer to plants. *Transgenic Research* **2**, 208-218.

Houtsmuller, A.B. (2005). Fluorescence Recovery after Photobleaching: Application to Nuclear Proteins. In *Microscopy Techniques*, pp. 177-199.

Hutten, S. and Kehlenbach, R.H. (2007). CRM1-mediated nuclear export: to the pore and beyond. *Trends Cell Biol* **17**, 193-201.

Jeong, S.Y., Rose, A., Joseph, J., Dasso, M. and Meier, I. (2005). Plant-specific mitotic targeting of RanGAP requires a functional WPP domain. *Plant J* **42**, 270-282.

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

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

Jones, D.T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* **292**, 195-202.

- Jones, J.D. and Dangl, J.L.** (2006). The plant immune system. *Nature* **444**, 323-329.
- Jones, L., Hamilton, A.J., Voinnet, O., Thomas, C.L., Maule, A.J. and Baulcombe, D.C.** (1999). RNA-DNA Interactions and DNA Methylation in Post-Transcriptional Gene Silencing. *Plant Cell* **11**, 2291-2302.
- Kohler, A., Rinaldi, C., Duplessis, S., Baucher, M., Geelen, D., Duchaussoy, F., Meyers, B.C., Boerjan, W. and Martin, F.** (2008). Genome-wide identification of NBS resistance genes in *Populus trichocarpa*. *Plant Mol Biol* **66**, 619-636.
- Kretsovali, A., Spilianakis, C., Dimakopoulos, A., Makatounakis, T. and Papamatheakis, J.** (2001). Self-association of Class II Transactivator Correlates with Its Intracellular Localization and Transactivation. *J. Biol. Chem.* **276**, 32191-32197.
- Kutay, U., Guttinger, S., Gorlich, D. and Kutay, U.** (2005). Leucine-rich nuclear-export signals: born to be weak
Transport between the cell nucleus and the cytoplasm. *Trends Cell Biol* **15**, 121-124.
- Lahaye, T.** (2002). The Arabidopsis RRS1-R disease resistance gene--uncovering the plant's nucleus as the new battlefield of plant defense? *Trends Plant Sci* **7**, 425-427.
- Lanford, R.E. and Butel, J.S.** (1984). Construction and characterization of an SV40 mutant defective in nuclear transport of T antigen. *Cell* **37**, 801-813.
- Lange, A., Mills, R.E., Devine, S.E. and Corbett, A.H.** (2008). A PY-NLS nuclear targeting signal is required for nuclear localization and function of the *Saccharomyces cerevisiae* mRNA-binding protein Hrp1. *J Biol Chem* **283**, 12926-12934.
- Lange, A., Mills, R.E., Lange, C.J., Stewart, M., Devine, S.E. and Corbett, A.H.** (2007). Classical nuclear localization signals: definition, function, and interaction with importin alpha. *J Biol Chem* **282**, 5101-5105.
- Launholt, D., Merkle, T., Houben, A., Schulz, A. and Grasser, K.D.** (2006). Arabidopsis Chromatin-Associated HMGA and HMGB Use Different Nuclear Targeting Signals and Display Highly Dynamic Localization within the Nucleus. *Plant Cell* **18**, 2904-2918.
- Lee, B.J., Cansizoglu, A.E., Suel, K.E., Louis, T.H., Zhang, Z. and Chook, Y.M.** (2006). Rules for Nuclear Localization Sequence Recognition by Karyopherin[β]₂. *Cell* **126**, 543-558.
- Leister, R.T., Dahlbeck, D., Day, B., Li, Y., Chesnokova, O. and Staskawicz, B.J.** (2005). Molecular genetic evidence for the role of SGT1 in the intramolecular complementation of Bs2 protein activity in *Nicotiana benthamiana*. *Plant Cell* **17**, 1268-1278.
- Linding, R., Jensen, L.J., Diella, F., Bork, P., Gibson, T.J. and Russell, R.B.** (2003). Protein disorder prediction: implications for structural proteomics. *Structure* **11**, 1453-1459.
- Liu, Y., Burch-Smith, T., Schiff, M., Feng, S. and Dinesh-Kumar, S.P.** (2004). Molecular Chaperone Hsp90 Associates with Resistance Protein N and Its Signaling Proteins SGT1 and Rar1 to Modulate an Innate Immune Response in Plants. *J. Biol. Chem.* **279**, 2101-2108.
- Liu, Y., Schiff, M., Serino, G., Deng, X.W. and Dinesh-Kumar, S.P.** (2002). Role of SCF ubiquitin-ligase and the COP9 signalosome in the N gene-mediated resistance response to Tobacco mosaic virus. *Plant Cell* **14**, 1483-1496.
- Lu, R., Malcuit, I., Moffett, P., Ruiz, M.T., Peart, J., Wu, A.J., Rathjen, J.P., Bendahmane, A., Day, L. and Baulcombe, D.C.** (2003). High throughput virus-induced gene silencing implicates heat shock protein 90 in plant disease resistance. *Embo J* **22**, 5690-5699.
- Maki-Valkama, T., Valkonen, J.P., Kreuze, J.F. and Pehu, E.** (2000). Transgenic resistance to PVY(O) associated with post-transcriptional silencing of P1 transgene is overcome by PVY(N) strains that carry highly homologous P1 sequences and recover transgene expression at infection. *Mol Plant Microbe Interact* **13**, 366-373.
- Marathe, R. and Dinesh-Kumar, S.P.** (2003). Plant defense: one post, multiple guards?! *Mol Cell* **11**, 284-286.
- Martin, G.B., Bogdanove, A.J. and Sessa, G.** (2003). Understanding the functions of plant disease resistance proteins. *Annual Review Of Plant Biology* **54**, 23-61.
- Mayor, A., Martinon, F., De Smedt, T., Petrilli, V. and Tschopp, J.** (2007). A crucial function of SGT1 and HSP90 in inflammasome activity links mammalian and plant innate immune responses. *Nat Immunol* **8**, 497-503.
- Merkle, T.** (2003). Nucleo-cytoplasmic partitioning of proteins in plants: implications for the regulation of environmental and developmental signalling. *Curr Genet* **44**, 231-260.
- Mestre, P. and Baulcombe, D.C.** (2006). Elicitor-mediated oligomerization of the tobacco N disease resistance protein. *Plant Cell* **18**, 491-501.
- Miyazaki, S., Kuroda, Y. and Yokoyama, S.** (2002). Characterization and prediction of linker sequences of multi-domain proteins by a neural network. *J Struct Funct Genomics* **2**, 37-51.
- Moffett, P., Farnham, G., Peart, J. and Baulcombe, D.C.** (2002). Interaction between domains of a plant NBS-LRR protein in disease resistance-related cell death. *Embo J.* **21**, 4511-4519.
- Mucyn, T.S., Clemente, A., Andriotis, V.M., Balmuth, A.L., Oldroyd, G.E., Staskawicz, B.J. and Rathjen, J.P.** (2006). The tomato NBARC-LRR protein Prf interacts with Pto kinase in vivo to regulate specific plant immunity. *Plant Cell* **18**, 2792-2806.
- Navarro, L., Zipfel, C., Rowland, O., Keller, I., Robatzek, S., Boller, T. and Jones, J.D.** (2004). The transcriptional innate immune response to flg22. Interplay and overlap with Avr gene-dependent defense responses and bacterial pathogenesis. *Plant Physiol* **135**, 1113-1128.
- Nishijima, H., Nakayama, J., Yoshioka, T., Kusano, A., Nishitani, H., Shibahara, K. and Nishimoto, T.** (2006). Nuclear RanGAP is required for the heterochromatin assembly and is reciprocally regulated by histone H3 and Clr4 histone

methyltransferase in *Schizosaccharomyces pombe*. *Mol Biol Cell* **17**, 2524-2536.

Noel, L.D., Cagna, G., Stuttmann, J., Wirthmuller, L., Betsuyaku, S., Witte, C.P., Bhat, R., Pochon, N., Colby, T. and Parker, J.E. (2007). Interaction between SGT1 and cytosolic/nuclear HSC70 chaperones regulates *Arabidopsis* immune responses. *Plant Cell* **19**, 4061-4076.

Ouali, M. and King, R.D. (2000). Cascaded multiple classifiers for secondary structure prediction. *Protein Sci* **9**, 1162-1176.

Pay, A., Resch, K., Frohnmeyer, H., Fejes, E., Nagy, F. and Nick, P. (2002). Plant RanGAPs are localized at the nuclear envelope in interphase and associated with microtubules in mitotic cells. *Plant J* **30**, 699-709.

Peart, J.R., Cook, G., Feys, B.J., Parker, J.E. and Baulcombe, D.C. (2002a). An EDS1 orthologue is required for N-mediated resistance against tobacco mosaic virus. *Plant J* **29**, 569-579.

Peart, J.R., Lu, R., Sadanandom, A., Malcuit, I., Moffett, P., Brice, D.C., Schausser, L., Jaggard, D.A.W., Xiao, S.Y., Coleman, M.J., Dow, M., Jones, J.D.G., Shirasu, K. and Baulcombe, D.C. (2002b). Ubiquitin ligase-associated protein SGT1 is required for host and nonhost disease resistance in plants. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 10865-10869.

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R. and Lopez, R. (2005). InterProScan: protein domains identifier. *Nucleic Acids Res* **33**, W116-120.

Rairdan, G.J., Collier, S.M., Sacco, M.A., Baldwin, T.T., Boettrich, T. and Moffett, P. (2008). The Coiled-Coil and Nucleotide Binding Domains of the Potato Rx Disease Resistance Protein Function in Pathogen Recognition and Signaling. *Plant Cell* **20**, 739-751.

Rairdan, G.J. and Moffett, P. (2006). Distinct Domains in the ARC Region of the Potato Resistance Protein Rx Mediate LRR Binding and Inhibition of Activation. *Plant Cell* **18**, 2082-2093.

Ratcliff, F., Martin-Hernandez, A.M. and Baulcombe, D.C. (2001). Tobacco rattle virus as a vector for analysis of gene function by silencing. *Plant J* **25**, 237-245.

Raval, A., Weissman, J.D., Howcroft, T.K. and Singer, D.S. (2003). The GTP-Binding Domain of Class II Transactivator Regulates Its Nuclear Export. *J Immunol* **170**, 922-930.

Reits, E.A. and Neefjes, J.J. (2001). From fixed to FRAP: measuring protein mobility and activity in living cells. *Nat Cell Biol* **3**, E145-147.

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

Rose, A. and Meier, I. (2001). A domain unique to plant RanGAP is responsible for its targeting to the plant nuclear rim. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 15377-15382.

Sacco, M.A., Mansoor, S. and Moffett, P. (2007). A RanGAP protein physically interacts with the NB-LRR protein Rx, and is required for Rx-mediated viral resistance. *Plant J* **52**, 82-93.

Schouten, A., Roosien, J., deBoer, J.M., Wilimink, A., Rosso, M.N., Bosch, D., Stiekema, W.J., Gommers, F.J., Bakker, J. and Schots, A. (1997). Improving scFv antibody expression levels in the plant cytosol. *FEBS Lett.* **415**, 235-241.

Schouten, A., Roosien, J., van Engelen, F.A., de Jong, G.A., Borst-Vrensens, A.W., Zilverentant, J.F., Bosch, D., Stiekema, W.J., Gommers, F.J., Schots, A. and Bakker, J. (1996). The C-terminal KDEL sequence increases the expression level of a single-chain antibody designed to be targeted to both the cytosol and the secretory pathway in transgenic tobacco. *Plant Mol Biol* **30**, 781-793.

Seksek, O., Biwersi, J. and Verkman, A.S. (1997). Translational diffusion of macromolecule-sized solutes in cytoplasm and nucleus. *J Cell Biol* **138**, 131-142.

Shen, Q.H., Saijo, Y., Mauch, S., Biskup, C., Bieri, S., Keller, B., Seki, H., Ulker, B., Somssich, I.E. and Schulze-Lefert, P. (2007). Nuclear activity of MLA immune receptors links isolate-specific and basal disease-resistance responses. *Science* **315**, 1098-1103.

Shen, Q.H. and Schulze-Lefert, P. (2007). Rumble in the nuclear jungle: compartmentalization, trafficking, and nuclear action of plant immune receptors. *Embo J* **26**, 4293-4301.

Shen, Q.H., Zhou, F., Bieri, S., Haizel, T., Shirasu, K. and Schulze-Lefert, P. (2003). Recognition specificity and RAR1/SGT1 dependence in barley Mla disease resistance genes to the powdery mildew fungus. *Plant Cell* **15**, 732-744.

Sisk, T.J., Roys, S. and Chang, C.H. (2001). Self-association of CIITA and its transactivation potential. *Mol. Cell. Biol.* **21**, 4919-4928.

Sprague, B.L. and McNally, J.G. (2005). FRAP analysis of binding: proper and fitting. *Trends Cell Biol* **15**, 84-91.

Sprague, B.L., Pego, R.L., Stavreva, D.A. and McNally, J.G. (2004). Analysis of binding reactions by fluorescence recovery after photobleaching. *Biophys J* **86**, 3473-3495.

Stewart, M. (2006). Structural basis for the nuclear protein import cycle. *Biochem Soc Trans* **34**, 701-704.

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

Tameling, W.I. and Baulcombe, D.C. (2007). Physical association of the NB-LRR resistance protein Rx with a Ran GTPase-activating protein is required for extreme resistance to Potato virus X. *Plant Cell* **19**, 1682-1694.

Tameling, W.I.L., Elzinga, S.D.J., Darmin, P.S., Vossen, J.H., Takken, F.L.W., Haring, M.A. and Cornelissen, B.J.C. (2002). The tomato R gene products I-2 and MI-1 are functional ATP binding proteins with ATPase activity. *Plant Cell* **14**, 2929-2939.

Tameling, W.I.L., Vossen, J.H., Albrecht, M., Lengauer, T., Berden, J.A., Haring, M.A., Cornelissen, B.J.C. and Takken, F.L.W. (2006). Mutations in the NB-ARC Domain of I-2 That Impair ATP Hydrolysis Cause Autoactivation. *Plant Physiol.* **140**, 1233-1245.

Thompson, N.L., Lieto, A.M. and Allen, N.W. (2002). Recent advances in fluorescence correlation spectroscopy. *Curr Opin Struct Biol* **12**, 634-641.

Tuskan, G.A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R.R., Bhalerao, R.P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G.L., Cooper, D., Coutinho, P.M., Couturier, J., Covert, S., Cronk, Q., Cunningham, R., Davis, J., Degroove, S., Dejardin, A., Depamphilis, C., Detter, J., Dirks, B., Dubchak, I., Duplessis, S., Ehlting, J., Ellis, B., Gendler, K., Goodstein, D., Gribskov, M., Grimwood, J., Groover, A., Gunter, L., Hamberger, B., Heinze, B., Helariutta, Y., Henrissat, B., Holligan, D., Holt, R., Huang, W., Islam-Faridi, N., Jones, S., Jones-Rhoades, M., Jorgensen, R., Joshi, C., Kangasjarvi, J., Karlsson, J., Kelleher, C., Kirkpatrick, R., Kirst, M., Kohler, A., Kalluri, U., Larimer, F., Leebens-Mack, J., Leple, J.C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D.R., Nelson, C., Nieminen, K., Nilsson, O., Pereda, V., Peter, G., Philippe, R., Pilate, G., Poliakov, A., Razumovskaya, J., Richardson, P., Rinaldi, C., Ritland, K., Rouze, P., Ryaboy, D., Schmutz, J., Schrader, J., Segerman, B., Shin, H., Siddiqui, A., Sterky, F., Terry, A., Tsai, C.J., Uberbacher, E., Unneberg, P., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**, 1596-1604.

Ueda, H., Yamaguchi, Y. and Sano, H. (2006). Direct Interaction between the Tobacco Mosaic Virus Helicase Domain and the ATP-bound Resistance Protein, N Factor during the Hypersensitive Response in Tobacco Plants. *Plant Mol Biol* **61**, 31-45.

Ursula Stochaj, K.L.R. (1999). Nucleocytoplasmic trafficking of proteins: With or without Ran? *Bioessays* **21**, 579-589.

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

van der Hoorn, R.A. and Kamoun, S. (2008). From Guard to Decoy: a new model for perception of plant pathogen effectors. *Plant Cell* **20**, 2009-2017.

van der Vossen, E.A.G., van der Voort, J., Kanyuka, K., Bendahmane, A., Sandbrink, H., Baulcombe, D.C., Bakker, J., Stiekema, W.J. and Klein-Lankhorst, R.M. (2000). Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: a virus and a nematode. *Plant J.* **23**, 567-576.

van Engelen, F.A., Molthoff, J.W., Conner, A.J., Nap, J.P., Pereira, A. and Stiekema, W.J. (1995). pBINPLUS: an improved plant transformation vector based on pBIN19. *Transgenic Res* **4**, 288-290.

van Engelen, F.A., Schouten, A., Molthoff, J.W., Roosien, J., Salinas, J.s., Dirkse, W.G., Schots, A., Bakker, J., Gommers, F.J., Jongma, M.A., Bosch, D. and Stiekema, W.J. (1994). Coordinate expression of antibody subunit genes yields high levels of functional antibodies in roots of transgenic tobacco. *Plant Molecular Biology (Historical Archive)* **26**, 1701-1710.

Wen, W., Meinkoth, J.L., Tsien, R.Y. and Taylor, S.S. (1995). Identification of a Signal for Rapid Export of Proteins from the Nucleus. *Cell* **82**, 463-473.

Wirthmueller, L., Zhang, Y., Jones, J.D. and Parker, J.E. (2007). Nuclear Accumulation of the Arabidopsis Immune Receptor RPS4 Is Necessary for Triggering EDS1-Dependent Defense. *Curr Biol* **17**, 2023-2029.

Xu, L. and Massague, J. (2004). Nucleocytoplasmic shuttling of signal transducers. *Nat Rev Mol Cell Biol* **5**, 209-219.

Xu, X.M., Meulia, T. and Meier, I. (2007). Anchorage of plant RanGAP to the nuclear envelope involves novel nuclear-pore-associated proteins. *Curr Biol* **17**, 1157-1163.

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

Yang, T.T., Cheng, L. and Kain, S.R. (1996). Optimized codon usage and chromophore mutations provide enhanced sensitivity with the green fluorescent protein. *Nucl. Acids Res.* **24**, 4592-4593.

Zipfel, C., Robatzek, S., Navarro, L., Oakeley, E.J., Jones, J.D., Felix, G. and Bolter, T. (2004). Bacterial disease resistance in Arabidopsis through flagellin perception. *Nature* **428**, 764-767.

Chapter 7

General Discussion

This chapter aims to shine a light on the mechanistical diversity and parallels that can be gleaned from various well-studied plant NB-LRR R proteins and structural or functional analogs in other organisms. A conformational activation model and the relation between the R protein's localization and functioning are discussed in the context of recent scientific literature. This chapter finishes with perspectives for future research based on various remaining questions about the functional mechanisms of R proteins.

The role of intramolecular interactions in controlling R protein activity

Our understanding of how R proteins specifically recognize pathogens, are activated, and initiate the signalling pathways leading to a resistance response, has grown tremendously since the cloning of the first NB-LRR encoding R genes (Bent et al., 1994; Mindrinos et al., 1994; Whitham et al., 1994; Grant et al., 1995). Even in those early studies the modular nature of NB-LRR proteins and the potential functions of the structural domains in pathogen recognition and signalling were recognized. Structural homologs of the TIR (Toll and Interleukin-1 Receptor), loosely defined CC, NB-ARC and LRR domains were known in other, often metazoan, proteins and assuming that function follows structure, a simple functional model could be drawn. The most straightforward mechanistical interpretation of the genetic concept of gene-for-gene interaction was the direct recognition of the pathogens avirulence product by a receptor-like R protein in the plant (Hammond-Kosack and Jones, 1997). Leucine-rich repeats directly interact with ligands in many receptor proteins and were predicted to have such a function in R proteins as well. In analogy with several known receptors the N-terminal domains of R proteins were assumed to relay the LRR mediated recognition to downstream signalling components. Although the basic idea of an R protein as a sensor for pathogen derived products still stands, the underlying mechanisms turned out to be much more complex and varied than originally anticipated and several basic steps in the working of R proteins remain to be elucidated.

The activation cycle of NB-LRR R proteins can be depicted as a series of more-or-less discrete events taking it from an inactive autoinhibited state to its active signalling state after pathogen recognition and back to inactivation. The NB-ARC domain plays a central role in this cycle as its conformational state integrates the functioning of the surrounding domains (Takken et al., 2006). Inadvertent activation of R proteins is extremely costly for plants, at the worst the hypersensitive response cell death destroys the cells, but even an upregulation of resistance pathways can stunt plant growth (Grant et al., 2003; Weaver et al., 2006). It is not surprising then that regulatory mechanisms have been found at the levels of gene transcription, translation, protein stability, and protein interactions (Jones et al., 2004; Holt et al., 2005; Halterman and Wise, 2006; Tan et al., 2007; Hubert et al., 2009). Similar layers of security can be found for metazoan apoptotic proteins as well (Borner, 2003; Danial and Korsmeyer, 2004; Holcik and Sonenberg, 2005; Yan and Shi, 2005; Faustin et al., 2009).

At the level of the single protein, activation is suppressed in the absence of an elicitor by intramolecular interactions between the domains. The inhibitory function of certain interdomain interaction becomes apparent if deletion of a domain results in constitutive activation. However, the fact that these domains are often also involved in signalling functions complicates disentangling the network of internal inhibitory mechanisms. Simply truncating the protein will not be informative in such cases.

The LRR domain is involved in keeping the R protein in an inactive state

CC-NB-LRR proteins

For Rx1, it was found that the autoactivation caused by overexpression of the protein in *Nicotiana tabacum* developed more rapidly when the LRR and different lengths of the ARC domain were deleted (Bendahmane et al., 2002). The additional deletion of a small part of the NB (from AA 282 onwards) reduced the autoactivity. Interestingly the autoactive response of the deletion mutants was not observed in *N. benthamiana* in an identical assay. Only when stabilised by GFP this deletion construct would induce an HR in *N. benthamiana* (Rairdan et al., 2008). Further deletion of the CC domain or introducing a P-loop mutation did not stop this constitutive activation (Rairdan et al., 2008), which establishes the NB as the signalling module of Rx1. The Rx1 CC domain only binds the NB-ARC when the LRR is present, and does not bind the LRR alone (Rairdan and Moffett, 2006; Rairdan et al., 2008).

Another CC-NB-LRR for which the interdomain interactions and their relevance in signalling have been studied is RPS5 from *Arabidopsis thaliana* (Warren et al., 1998) which confers resistance to *Pseudomonas* strains carrying *AvrPphB*. RPS5 interdomain interactions vary from the interdomain interactions of Rx1. The RPS5 CC binds the NB-ARC, even when the LRR is not present. Its LRR binds the NB-ARC as well. Furthermore the CC, the NB-ARC and the LRR can each form homodimers, like the full-length protein (Ade et al., 2007). If the LRR is deleted, the truncated construct becomes constitutively active, but this response is weaker than the elicitor induced response (Ade et al., 2007). This suggests that the release of the LRR mediated inhibition does not fully activate the signalling and the LRR is needed in the active state, for example in stabilising a conformation or aiding in the binding of downstream components.

RPS2 is also a CC-NB-LRR from *Arabidopsis* that confers resistance against *Pseudomonas syringae* strains, but expressing *AvrRpt2*. Overexpression of RPS2 gives an autoactivation response (Tao et al., 2000). However, when the C-terminal end of the LRR was deleted ($\Delta 890$), this autoactivation disappeared. Only after deleting the complete LRR ($\Delta 492$) the same overexpression activation appeared again. Partial deletion of the LRR seems to strengthen its inhibitory role in RPS2. Further deletions of the CC-NB-ARC lost the overexpression induced autoactivation, showing that the complete NB-ARC was required for activity.

TIR-NB-LRR proteins

Similar examples of autoactivation after the deletion of the LRR have also been found for Toll- Interleukin receptor like (TIR)-NB-LRR proteins. The TIR-NB-LRR RPS4 from *Arabidopsis* gives an elicitor-independent activation when expressed in *N. benthamiana* (Zhang et al., 2004). Deletion of the LRR strongly reduces the protein concentration, but not the HR. The fact that similar expression levels in an *Arabidopsis* background do not lead to autoactivity could mean that in *Arabidopsis* external RPS4 regulators are present. A similar difference between expression in *Arabidopsis* and heterologous expression in *N. tabacum* could be seen for a RPP1A deletion missing the LRR (Weaver et al., 2006). Deleting the NB-ARC domain, leaving only the TIR, does abolish the HR for both RPS4 and RPP1A. However, if the small region in between the TIR and the P-loop of the NB is retained in the deletion construct a strong HR occurs (Swiderski et al., 2009) as was earlier seen for L10 from flax (Frost et al., 2004). For these TIR-NB-LRR proteins clearly not the NB, but the TIR is the effector domain sufficient for downstream signalling. The presence of the short sequence also stabilises the TIR domain; without it the RPS4 TIR is no longer detectable (Weaver et al., 2006; Swiderski et al., 2009). In the same study a similar deletion construct was created for several TIR-NB-LRRs, but surprisingly not all of them could be activated by this specific truncation, showing that within the class of TIR-NB-LRR proteins multiple regulatory mechanisms might exist (Swiderski et al., 2009).

Interestingly, for both RPS4 and the TIR-NB-LRR protein N from tobacco alternative splicing forms exist that give natural truncated versions of the proteins missing varying lengths of the LRR domain. These alternative transcripts are needed for full resistance (Dinesh-Kumar and Baker, 2000; Zhang and Gassmann, 2003; Takabatake et al., 2006). One of the alternative splice forms of RPS4 (RPS4^{TN4L}) that is upregulated in the presence of the elicitor, is highly unstable, but able to induce an HR and is proposed to function as an amplifier of the response (Zhang and Gassmann, 2007). Alternative transcripts have been observed for 11 TIR-NB-LRR R genes in *Arabidopsis* (Tan et al., 2007), but whether they play a similar role in CC-NB-LRR functioning is not resolved yet.

Summarising, the LRR domain has an inhibitory function in both CC-NB-LRR and TIR-NB-LRR. However, the mechanism seems to differ. Intradomain interactions like reported for Rx1 and RPS5 have not been reported for TIR-NB-LRR proteins.

Modification of the interaction between ARC2 and LRR results in autoactivation

The LRR domain has not only an inhibitory function that would be lifted upon elicitor recognition. Several of the LRR deletion constructs give a response that is weaker than the elicitor induced response or the response caused by for example a mutation in the MHD motif. When the D460V mutation is introduced in the CC-NB-ARC of Rx1, no autoactivation can be seen in the absence of the LRR. Coexpression of the LRR enables the constitutive activity *in trans* (Moffett et al., 2002). Apparently, an interaction with the LRR is needed to bring the CC-NB-ARC D460V in its activated state. If the NB of Rx1 is sufficient for signalling, then the ARC domains too have an inhibitory effect on signalling. The LRR might help the ARC domains to release this inhibition from the NB.

Remarkably, coexpression with an ARC-LRR construct does not enable the autoactivity of CC-NB-ARC D460V *in trans*, unless the Rx1 elicitor is present (Moffett et al., 2002). In the ARC-LRR construct the LRR is probably bound intramolecularly to the ARC as described in Chapter 5 and not available for intermolecular interaction with the CC-NB-ARC D460V. The presence of the elicitor might modify the ARC-LRR interaction in such a manner that the LRR becomes available for an interaction with the CC-NB-ARC *in trans*.

For both Rx1 and Mi-1.2, a root knot nematode resistance gene from tomato, it was shown that the physical interaction between the NB-ARC and LRR is not easily broken by either loss-of-function or gain-of-function mutations (Rairdan and Moffett, 2006; van Ooijen et al., 2008a). Several alanine substitutions in motifs interacting with the nucleotide in the Rx1 NB-ARC, DD244AA, GLP330ALA and MHDV458AAAA, did reduce the interaction with the LRR compared with the wild-type, but did not completely abolish it. Similarly, substituting the RNBS-D motif for alanines (CFly389AAAA), which is positioned in the interface with the LRR in our docking model (Chapter 5), results in a loss-of-function of Rx1, but not in a loss of the interaction with the LRR in pull down experiments. It is interesting that the interaction between the LRR and the NB-ARC is aspecific enough to occur between NB-ARC and LRR domains of more distantly related proteins like HRT, Bs2 and Rx1. At least between Rx1 and Bs2 the basic residues in the LRR and the acidic residues in the ARC2 domain predicted to be involved in electrostatic interactions are conserved. These heterologous interactions do not result in a functional protein, however.

In Mi-1.2 a series of autoactivating mutations in the NB and ARC domains was tested for trans-complementation with the LRR and for effect on the interaction between the NB-ARC and the LRR. Interestingly, NB-ARC constructs with autoactivating mutations in the MHD motif in ARC2 (H840A, D841V) could not be complemented *in trans* with the Mi LRR, but autoactivating mutations in the NB (T557S, D630E) were signalling competent when coexpressed with the LRR. However, none of these mutations affected the physical interaction with the LRR (van Ooijen et al., 2008a). It would be worthwhile to investigate if distinct areas in the interaction surface are responsible for the functional interaction, the communication, between the NB-ARC and LRR and for the physical interaction between the domains. Based on the docking model presented in Chapter 5, the electrostatic interactions between the basic patch in the LRR and the acidic loop in the ARC2 domain could contribute strongly to the physical interaction. The mutagenesis of these target regions resulted in a loss-of-function *in trans*, but not to autoactivation or loss-of-function *in cis*, which means that the functional interaction between the LRR and NB-ARC domains is not disturbed. However, this hypothesis still has to be tested by showing that mutating the charged patches affects the binding of the domains.

The close cooperation between the ARC and LRR domain is evident in the Gpa2 NB-ARC-LRR docking model presented in Chapter 5. The N-terminal repeats of the LRR are in close contact with the ARC2 at the position of the groove between the ARC2 and NB domains. The ARC2 region and the first repeats of the LRR were shown to co-evolve, in contrast to more distant regions with similar variability. If sequences were exchanged between Rx1 and Gpa2 in these regions either concentration dependent constitutive activity or loss-of-function phenotypes were observed (Chapter 4 and Chapter 5). Normal functioning could

be regained by combining the matching ARC2 and N-terminal LRR fragments, showing that these two subdomains should be compatible to operate as a functional unit.

Several autoactivating or activation sensitizing mutations have been reported in literature that map on the proposed docking interface between the ARC2 and LRR in Rx1 (Bendahmane et al., 2002; Farnham and Baulcombe, 2006). It is tempting to speculate that the conservation of several motifs in the opposing interaction surfaces of the LRR and ARC2 domains has its origin in a conserved interaction and functional cooperation between the ARC2 and LRR. The sequence of the second and third LRR repeats in Rx1 and Gpa2 are conserved among many CC-NB-LRR R proteins (described as LRR motif 1 in (Meyers et al., 2003).) Mutation E572K in the RPS5 VLDL repeat (VLDLSE) was shown to have a dominant negative effect on multiple R proteins (Warren et al., 1998), whereas mutation D543E in the Rx1 VLDL repeat (VLDLGLN) resulted in constitutive activity (Bendahmane et al., 2002).

The role of the N-terminal domain in controlling R protein activity

Analogous to metazoan immune receptors or apoptosis regulators like the Toll receptor or Nod1 and Apaf1, the role of the N-terminal domain of R proteins could be an effector or adaptor domain. The cytoplasmic TIR domain of Toll-like receptors for example interacts with the TIR domain of MyD88, which in turn interacts with its Death domain (DD) with the DD domain of the Interleukin-1 receptor associated kinase (IRAK). From there the signalling pathway follows several protein interactions and eventually ends in transcriptional reprogramming via transcription factor activation (Akira and Takeda, 2004). The N-terminal CARD domains of Apaf-1, CED-4 and NOD recruit caspases via CARD-CARD interactions and oligomerization leads these caspases activate each other via induced proximity (Inohara et al., 1999; Shiozaki et al., 2002; Wagner et al., 2009). For the N-terminal CC, TIR or extended Solanaceae specific domains (SD) of R proteins, no clear role in downstream signal transduction has been shown yet. There is on the other hand evidence that the N-terminal domains can have other roles in R protein functioning, for example in the regulation of the R protein's activation or in the formation of complexes with virulence targets.

Signalling

The kinase Pto interacts with the N-terminal (NT) domain upstream of the CC and the Solanaceae domain (SD) in Prf (Mucyn et al., 2006). Both the kinase Pto and the N-terminally extended CC-NB-LRR protein Prf are needed to mediate resistance against *Pseudomonas* strains carrying AvrPto or AvrPtoB (Salmeron et al., 1996). In a mechanism probably comparable with the RanGAP2 and Rx1 interaction (Sacco et al., 2007; Tameling and Baulcombe, 2007), the overexpression of both Pto and Prf in *N. benthamiana* induces an elicitor-independent HR. This could mean that the binding of Pto brings Prf in a signalling competent conformation, needed for elicitor dependent activation, but in case of overexpression resulting in autoactivation. Pto kinase activity (autophosphorylation) and N-myristoylation are required for this response. Prf and Pto enhance each others protein accumulation. The instability and breakdown of each non-complexed protein could be a feed-back mechanism that ensures the presence of only signalling competent complexes. Interestingly, RPM1 that is not bound to RIN4 is also actively degraded (Mackey et al.,

2002). Apparently, R protein functioning is also regulated by modulation of the amount of protein present in the cell. It will be interesting to see if similar mechanisms are involved in the regulation of Rx1 and Gpa2 activity, which was shown to be concentration dependent (Chapter 4).

The CC and NB-ARC-LRR of Rx1 can function in an elicitor-dependent activation *in trans*. Autoactive versions of Rx1, like the D460V or the Y712H mutation, which are active if the CC-NB-ARC and LRR are coexpressed, do not function when the CC and NB-ARC-LRR are coexpressed (Rairdan et al., 2008). Rx1 NB-ARC-LRR D460V does not interact with the CC, in contrast to the wild type NB-ARC-LRR. The conformation of the NB-ARC-LRR might have changed by the D460V mutation, and thereby its surface characteristics. Interestingly, the Y712H autoactivating mutation in the LRR does not impair the CC interaction with the NB-ARC-LRR construct. The differences in *in trans* functionality between the CC-NB-ARC and the CC seem to indicate that interaction between the CC and NB-ARC-LRR plays a role at a different moment in the activation cycle of the R protein than the interaction between the NB-ARC and the LRR. If the NB is the minimal signalling domain and the internal inhibition of the ARC domain is lifted by D460V *in cis*, or in a LRR deletion construct only with the cooperation of the LRR domain *in trans*, then the function of the CC seems to be needed prior to the activation state mimicked by the D460V mutation.

In Mi-1.1 and Mi-1.2 a specific interaction between the LRR and the extended N-terminal domain prevents autoactivation. Sequence exchange between the proteins lead to constitutive activation, if the N-terminal domain and a small region of the LRR did not match (Hwang et al., 2000). Coexpression of the N-terminal domain matching the LRR in the autoactive construct could suppress the activation (Hwang and Williamson, 2003). For Rx1 and Gpa2, the coexpression of the free CC or CC-NB-ARC domain with the autoactive chimeric constructs could suppress the HR phenotype (unpublished data). However no specificity in the suppression for either Gpa2 or Rx1 derived CC domains was seen. The CC domains of both Rx1 and Gpa2 specifically bind to the RanGAP2 (Ran GTPase Activating Protein 2) N-terminal domain (Sacco et al., 2007; Tameling and Baulcombe, 2007). Decreasing the available RanGAP2 protein leads to a partial loss of Rx1 resistance (Tameling and Baulcombe, 2007), whereas overexpression increases the autoactivity of CC-NB or CC-NB-ARC constructs (Sacco et al., 2007). The CC-RanGAP2 interaction could force the NB-ARC in a conformation that is easier to activate than the unbound version. The mechanism by which this suppression works and if RanGAP2 is involved remains to be identified. Mutations made in the CC domain of RPS2 resulted in a dominant negative suppression of coexpressed wild type RPS2. This might be explained by the sequestering of downstream components, or other RPS2 proteins forming complex unable to signal (Tao et al., 2000).

Guarding

The signalling capacity of TIR domains expressed in the absence of the NB-ARC-LRR as discussed previously, is one of the few examples where a N-terminal R protein domain appears to function as a signalling domain. But even in these examples the TIR domain's role is more complex. The TIR domain of N from tobacco for example, interacts with a chloroplast enzyme, a rhodanese sulfurtransferase named NRIP1, which is translocated

from the chloroplast to the cytoplasm by the TMV helicase p50 (Caplan et al., 2008). P50 is the elicitor of N and binds NRIP1 too. This means that only in the presence of p50 NRIP1 and N can meet in the same subcellular compartment and only when NRIP1 interacts with the N TIR domain p50 can be in one complex with N and induce N's activation. So, next to a role in signal transduction, the N TIR domain has a role in the interaction of the R protein with a guarded virulence target. A similar interaction with a guarded virulence target is seen for several CC domains. The CC domains of RPM1 and RPS2 for example bind RIN4, the target of various *Pseudomonas* effectors (Mackey et al., 2003; Kim et al., 2005a). The RPS5 CC interacts with the kinase PBS1, which is the target of the *Pseudomonas* effector protease AvrPphB (Shao et al., 2003; Ade et al., 2007).

Recognition

A role for a N-terminal domain in recognition was found for flax resistance genes in the L locus (Luck et al., 2000). Exchanging just the TIR domain between L6 and L2 shifted the specificity of the chimeric protein. The interaction with the flax rust derived elicitors still depends on the LRR region (Dodds et al., 2006). Maybe the specificity is determined by a cooperation of N- and C-terminal domains

The constitutive activity in the chimeric constructs created by sequence exchanges between Mi-1.1 and Mi-1.2 was coupled to a functional interaction between the extended N-terminal domain and the LRR. In an extensive targeted mutagenesis study all 40 positions in which the Mi-1.1 and Mi-1.2 LRRs varied were mutated individually. In this study the mutation of one position from the Mi-1.2 identity to the Mi-1.1 identity (R961D) was found to stop the constitutive activity in the chimeric construct and introduce constitutive activity in Mi-1.2 (Hwang and Williamson, 2003). Most of the other mutations that caused a loss of constitutive activation in the chimeric construct also showed a loss of elicitor-dependent activation in the Mi-1.2 background, showing that activational control and recognition are closely linked.

In our studies we found that exchanging the CC domain of Rx1, the CC and NB domain, or the ARC1 domain by the corresponding Gpa2 sequences resulted in chimeric proteins with a more pronounced response to the virulent PVX CP105, without showing autoactivation (chapter 5). Either the activation of these constructs is sensitized and therefore able to respond to the virulent CP, or the recognition specificity is slightly broadened. The latter phenotype has earlier been associated with mutations in the C-terminal half of the LRR (Farnham and Baulcombe, 2006). Both N- and C-terminal fusions of fluorescent proteins to Rx1 brought about a similar response to the virulent PVX CP105. In the Gpa2 docking model the N- and C-terminal parts of the NB-ARC-LRR are relatively close in space. One could envision that the CC domain in this structure contacts the LRR and that this interaction has a role in regulating the sensitivity of the recognition.

In conclusion, the N-terminal domains of R proteins have been shown to function in internal inhibitory interactions, resistance response signalling, the interaction with guarded virulence targets, and can even contribute to specificity. For Rx1 and Gpa2 an interaction of the CC with RanGAP2 has been identified for which the exact function in signalling is still

unknown (Sacco et al., 2007; Tameling and Baulcombe, 2007), but which could be a virulence target. Furthermore, specific sequences in the CC domain have been identified that are required for either the NB-ARC-LRR binding or the interaction with RanGAP2 (Rairdan et al., 2008). Both these sequences in the CC lie within the CC fragment that we show to be the minimal domain that shows a strong nuclear accumulation and slow diffusion (Chapter 6).

The role of the central NB-ARC domain in R protein activation

The conserved NB-ARC domain is the central focus of the inhibitory and activating actions of the R protein's N- and C-terminal domains described above. It functions as molecular switch that integrates the state of the surrounding domains into signalling or inactive conformations. The interaction of the NB and ARC with the bound nucleotide determines the conformational on- or off-state of the switch. NB-LRR R proteins can be grouped together with the metazoan apoptosis regulators CED-4 and Apaf-1 into a structural family based on the similarity in the NB and ARC domains (van der Biezen and Jones, 1998). Together with the NACHT family (Koonin and Aravind, 2000), the NB-ARC proteins are part of an ancient group of signal transduction ATPases with numerous domains (STAND) found in a wide variety of organisms (Leipe et al., 2004). Several studies have shown that the structural similarities between the NB domains in these protein families underlie functional similarities.

ATP binding

Several lines of evidence point at the ATP-bound state as the active conformation of the NB-ARC structure, as opposed to an inactive ADP-bound state. The R protein I-2 and Mi-1.2 have been shown to bind ADP and ATP, and to have ATPase activity (Tameling et al., 2002). Based on the fact that mutations disrupting I-2's ATPase activity, but not its ATP binding capacity, resulted in constitutively active signalling, it is likely that the ATP bound conformation of the NB-ARC is the active form (Tameling et al., 2006). The ATPase activity can switch the protein from the active to the resting state. So far, no ATPase activity has been shown for Rx1 despite several attempts (Takken *et al.*, pers. comm.).

In the bacterial transcriptional regulator MalT a similar mutation inhibiting ATP hydrolysis resulted in constitutive activation (Marquet and Richet, 2007). MalT belongs to the NACHT subclass of STAND proteins, instead of the NB-ARC class of R proteins and Apaf-1. MalT is an example of how STAND proteins can act as integrators of multiple input signals by combining intermolecular interactions and different nucleotide-dependent activation states. The monomeric ADP-bound resting form is stabilized through interaction with different negative effectors (MalK, MalY, Aes) binding to the NB and helical domain (which is similar to the ARC1 domain in NB-ARC proteins). The presence of its inducer maltotriose, sensed via the C-terminal SUPR-type domain, overrides the negative regulators and makes nucleotide exchange and ATPase activity possible. Interestingly, the kinase-2 mutation that inhibits ATP hydrolysis also reduces the sensitivity of MalT for two out of three activation repressors (MalK and MalY). ATPase activity in itself is not needed for active signalling, it seems just to regulate the protein's activity by returning it to its resting state (Danot et al., 2009).

An idea of how the ADP-bound conformation and the ATP-bound conformation of the NB-ARC module differ from each other can be gained from comparing the crystal structures of the ADP-bound Apaf-1 and the ATP-bound CED-4 from *C. elegans* (Riedl et al., 2005; Yan et al., 2005). In Apaf-1 the NB, ARC1 and ARC2 have a compact conformation, with amino acids from each subdomain interacting with the bound ADP. The conformation of the ATP-bound CED-4 is strikingly different; the ARC2 domain is folded away from the NB domain, around the ARC1 domain, creating an elongated, open structure. CED-4 was crystallized as homodimer complexed with the negative regulator CED-9. The two CED-4 molecules in the structure bind each other via an NB-NB interaction and via a separate interaction of the ARC2 domain in one protomer and the ARC1 domain in the other protomer. No ATPase activity has been detected for CED-4 and no ADP bound forms have been found through mass spectrometry (Seiffert et al., 2002; Yan et al., 2005), which could mean that the resting state of this CED-4 is also an ATP bound form. Maybe the inhibitory binding of CED-9 replaces intramolecular interactions found in the NB-ARC of Apaf-1. The suppression by CED-9 is released via a forced conformation change of CED-9 by the positive regulator EGL-1 (del Peso et al., 2000; Yan et al., 2004).

The difference between the closed ADP-bound Apaf-1 conformation and the open ATP-bound CED-4 conformation suggests that the binding surface of an NB-ARC module changes dramatically upon nucleotide exchange or hydrolysis, providing a structural basis to the domains molecular switch function.

Even if ATPase activity is not a prerequisite for activation, nucleotide binding in itself is necessary for functionality of NB-ARC proteins. Mutations of the phosphate-binding loop (P-loop) of the NB consistently resulted in a loss-of-function phenotype in numerous studies. Apparently, both the open and closed state, as described above, need a nucleotide as organizing centre for the NB-ARC subdomains.

The GFP stabilized NB domain of Rx1 is an interesting exception in that respect (Rairdan et al., 2008). The elicitor-independent activation of this truncated protein was not hindered by a mutation in the P-loop, a structural motif in the NB domain of R proteins involved in ATP/ADP binding. This suggests that the Rx1 NB domain might be in an uninhibited activation state in which ATP binding is no longer required. The nucleotide-binding of STAND proteins can be regulated by interacting proteins. It was recently shown that Bcl-2 and Bcl-XL inhibit the functioning of the NACHT protein NALP1 by inhibiting its ATP-binding (Danot et al., 2009; Faustin et al., 2009). A short (20 AA) conserved loop from these proteins binds NLRP1 with high affinity. However, the constitutively active NLRP1 Δ LRR cannot be inhibited by Bcl-2 or this inhibitory loop if added as peptide. Like in the Rx1 NB domain it seems that in NALP1 in the absence of certain internal inhibitory interactions the nucleotide-binding state is no longer regulating activation.

Nucleotide-dependent conformation and oligomerization

Many proteins from the AAA+ structural superfamily, to which the STAND proteins (NB-ARC and NACHT) belong, have been shown to oligomerize as part of their activational

mechanism (Diemand and Lupas, 2006). The bacterial transcription regulator MalT, for example, has been shown to form multimeres through homotypic interactions of its ATP-bound NB domain. Mutations that abolish ATP hydrolysis increase the MalT self-association (Steebhorn et al., 2001; Marquenet and Richet, 2007). The heptameric state of ATP-bound Apaf-1 (Kim et al., 2005b) forms large wheel-like structures, named apoptosomes (Cain et al., 2000; Acehan et al., 2002).

Two theories have been proposed on how the ATP-bound conformation of Apaf-1 monomers forms oligomers, based on the 12.8 Å resolution electron cryomicroscopy based structure of the Apaf-1 apoptosome. In the original paper describing the 12.8 Å structure, a model was presented in which the central ring of oligomerizing domains consists of alternating NB and ARC2 domains (winged-helix domain) (Yu et al., 2005). Diemand and Lupas on the other hand propose an oligomerization based on the binding via homotypic interactions of the NB-ARC1 regions of the proteins (Diemand and Lupas, 2006). This latter model is more in line with the association seen for other proteins of the AAA+ superfamily, to which the STAND proteins belong. Many of the proteins in this family form ring-like multimers, most often hexamers. In the multimeric state of these ATPases the NB and the helical domain (analogous to ARC1) mediate oligomerization. The nucleotide is bound in the interfaces between the proteins. In one protomer it interacts with the P-loop, kinase 2, sensor I and sensor II motifs and in the second it interacts with a conserved Arginine-finger (Danot et al., 2009). Being buried in the interface would protect the nucleotide-triphosphate from hydrolyzation and exchange, and is therefore a factor stabilizing the ATP-bound state.

The common propensity of STAND proteins to form multimers raises the question whether R proteins also need to form oligomers in order to initiate defence signalling. The tobacco TIR-NB-LRR N (conferring TMV resistance) was shown to oligomerize in the presence of its elicitor, the viral helicase p50 (Mestre and Baulcombe, 2006). This oligomerization was not influenced by silencing the downstream signalling components NRG1 and EDS1, which means that oligomerization is an early event in N's signalling. Inhibiting ATP binding by mutating the P-loop or silencing the chaperone-like SGT1 on the other hand did abolish the oligomerization. A mutation in the RNBS-A motif, one of the conserved NB-ARC motifs varying between CC-NB-LRR and TIR-NB-LRR proteins, does not affect the oligomerization, but does abrogate the elicitor induced stabilization of N.

For the CC-NB-LRR RPS5 di- or oligomerization of the full-length proteins has been shown (Ade et al., 2007). Even the separate CC, NB-ARC or LRR domains could form homodimers. Inhibition of ATP hydrolysis by mutating the catalytic aspartic acid residue in the kinase 2 motif resulted constitutive activity, like shown for I-2 (Tameling et al., 2006). The dimerization of the LRR or CC domains in the absence of the NB-ARC region indicate that multimerization of RPS5 could be regulated differently from proteins in which only the NB-ARC or analogous domain is involved in multimerization. It would be interesting to know how the RPS5 nucleotide-bound state influences the formation of oligomers.

ADP/ATP exchange

The ADP-bound Apaf-1 Δ WD40, used as template for the Gpa2 NB-ARC modelling in Chapter 5, forms a compact structure in which ADP is buried deeply between the NB and ARC2 Domain (Riedl et al., 2005). Because there is limited contact between the pocket containing the nucleotide and the solvent, only a shift to a more open structure would allow an ADP/ATP exchange. In Apaf-1 the closed conformation is further stabilized by the binding of the N-terminal CARD to the NB and ARC2 interface, and probably also by the interactions of the WD40 domain with the NB (Hu et al., 1998; Acehan et al., 2002). On the other hand the intramolecular interactions that inhibit ADP/ATP exchange, are also competing with intermolecular interactions that play a role in signalling. When the CARD domain binds to the NB and ARC2 it cannot recruit Caspase-9 (Qin et al., 1999). Similarly the NB and ARC1 domains in this closed state are not available for oligomerizing interactions with NB and ARC1 domains from other Apaf-1 molecules. One layer of inhibition is released when cytochrome c binds to the C-terminal WD40 repeat. This cytochrome c interaction thereby enhances the nucleotide exchange rate (Jiang and Wang, 2000). However, even the closed form of Apaf-1 can exchange nucleotides (e.g. ADP for dADP), suggesting that the conformation does not remain completely fixed but might open up from time to time. The direct binding of calcium to Apaf-1 has been shown to fixate the closed conformation, thereby strongly inhibiting nucleotide exchange and the possibility to form apoptosomes (Bao et al., 2007). The nucleotide exchange (ATP for ADP) that is essential for NALP1's oligomerization only occurs in the presence of the NALP1 specific inducer muramyl dipeptide (Faustin et al., 2007). Deletion of the NALP1 LRR made the ATP binding inducer-independent.

The conserved MHD motif is suggested to play an important role in the nucleotide exchange. The conserved histidine (H459 in Gpa2 and Rx1, H438 in Apaf-1) coordinates the β -phosphate of ADP via a hydrogen bond (Riedl et al., 2005). These contacts are likely stabilize the NB-ARC conformation and the bound ADP (van Ooijen et al., 2008b). Mutations of the MHD motif in NB-LRR R proteins often result in constitutive activity, which could be caused by a reduced inhibition of nucleotide exchange (Bendahmane et al., 2002; Howles et al., 2005). Inhibiting ATP-hydrolysis does not further enhance the autoactivity caused by this mutation (van Ooijen et al., 2008b). This could mean that with mutations in the MHD motif, the rate of nucleotide exchange stabilizing the activated conformation is much higher than the rate of hydrolysis.

Activation model for the NB-ARC of Rx1 and Gpa2

The presence of multiple inhibitory mechanisms in R proteins, some epistatic to others, predicts the existence of intermediate activation stages. From experimental data from comparable protein structures a series of events can be proposed through which the proteins go before being fully activated. The LRR domain has been found in several cases to be involved in inhibitory interactions with the NB-ARC and N-terminal domains, in the absence of the specific elicitor. Similarly, N-terminal domains have been found to function in intramolecular inhibition in the resting state of the protein. In some cases a guarded virulence target binds the N-terminal domain, and this interaction has been shown to bring the protein into a conformation, which is easier to activate. One could envision that the

recruitment of the guardee competes with intramolecular interactions. Within the core NB-ARC further activation is regulated by interactions between the NB-ARC subdomains and the nucleotide. These determine if nucleotide exchange can take place and if an open conformation can be formed. Other intermolecular interactions will determine how stable the open conformation is. Nucleotide hydrolysis could return the structure to an ADP-bound resting state. Oligomerization on the other hand has been shown to block hydrolysis activity as shown for several STAND protein oligomers in which the ATP is locked, resistant to exchange in the interface of the individual protomers. The LRR is not only an inhibitory domain released after specific recognition, but is also required for activation. Its cooperation with the ARC2 domain, as observed in the sequence exchange experiments in Chapter 5 could be a clue to understanding its activating function. It is tempting to speculate that through the interaction between the N-terminal half of the LRR and the ARC2 domain, the LRR can regulate nucleotide exchange.

In the Rx1 / Gpa2 sequence exchange constructs maybe two or three steps in the activation process are influenced. The CC and CC-NB exchange, which broadens the specific recognition to include the virulent PVX CP is one in which the match between the CC and the NB-ARC-LRR seems central. The exchanges disrupting the match between the ARC2 and the N-terminal half of the LRR would be another, in which the tuned communication between the LRR and the ARC2 changes the LRR's control over nucleotide exchange. The two mutations based on the minimal Gpa2 ARC2 sequence incompatible with the Rx1 LRR indicate that a further division can be made. One of the residues influences specifically the elicitor-dependent activation and is predicted, based on the structural model, to reside in the interface between the NB and the ARC2 domains.

The scenario as sketched above for the steps in R protein activation is of course a generalization, combining mechanisms seen in various proteins containing NB-ARC or NACHT domains. Although there clearly are common themes as seen in most NB-LRR proteins, a remarkable variation is observed in their mode of action at a more detailed level.

Perspectives

The potato R proteins Rx1 and Gpa2 have proven to be a rich model system for the study of R protein functioning and evolution. Like many R proteins their analysis is seriously hindered by low expression levels, limited protein stability, the need for laborious resistance assays when studying their functioning in potato, and last but not least, the hypersensitive phenotype which signifies functionality but greatly hampers the study of this functionality in the cell. However, several other properties compensate for all these disadvantages. For example, they can be studied in simple transient expression assays, for both Gpa2 and Rx1 the cognate elicitors are known. Fluorescent fusion constructs of these proteins are (just) stable enough to allow observations by confocal microscopy. And especially Rx1 can be split in many small fragments, which still assemble in the cell and function as an intact protein, a reductionist's dream. In addition, more than 70 homologous genes have been obtained from distantly related wild *Solanum* species (Butterbach, thesis), gaining insights in the evolutionary dynamics of the genes encoding Rx and Gpa2. These protein characteristics and the increasing knowledge about their function provide a solid basis to address several remaining questions about the structure, function and cellular dynamics of CC-NB-LRR R proteins.

What other components exist in the Rx1/Gpa2 signalling pathway?

How the activation of the Rx1 or Gpa2 (or any other R protein) eventually results in cell death is still a mystery. Now that it has been shown that a truncated construct consisting of only the NB domain is signalling competent (Rairdan et al., 2008), it is likely that this subdomain is responsible for downstream signalling and therefore likely to interact with downstream signalling components. Only a few proteins have been shown to be absolutely required for Rx1 and Gpa2 functioning, namely SGT1 and HSP90. Both are probably involved in R protein stability or assisting in conformational changes, for example by stabilizing intermediate stages in the activation process. RanGAP2, another interactor binding the CC domain of both Rx1 and Gpa2, is not likely to be a signalling component. Even truncated constructs only leaving the binding domain, seem to be able to bring Rx1 in a more activation competent state. Further investigation of the association of the CC to a large nuclear component (Chapter 6) will maybe bring another interactor to light, which is involved in the regulation of Rx signalling.

What is the role of the dual localization of Rx1 in its signalling?

Several R proteins are inactivated if they are forced out of the nucleus. For Rx1 several lines of evidence point to the opposite; if its elicitor or the interactor RanGAP2 (pers. comm. Wladimir Tameling) are forced into the nucleus, Rx1 is no longer activated. We know this type of R proteins go through conformational changes in their activation cycle. However, nothing is known about how these interactions take place in real time and how subcellular localization could influence these conformations or how the conformation contribute to the specific subcellular targeting of R proteins as demonstrated for the P-loop mutation (Chapter 6). Physiological conditions, including the ratio of ATP and ADP, differ greatly

between the cytoplasm and the nucleus and it will be interesting to monitor how this affects R protein trafficking in plant cells.

What are generic and specific structural properties of R proteins?

Several loss-of-function and gain-of-function mutations are known for Rx1. Their relative position in the protein structure and the strength of their phenotype, suggest they affect various different intramolecular regular mechanisms and stages in the activation process. More information could be gained by combining mutations to study if they are epistatic or if they affect the functionality of them in trans-expressed domains in the same way as they affect the full length protein. More detailed mutagenesis studies could be conducted trying to answer the structure-function relationships in R proteins. Additional targeted mutations can be designed based on the structural models obtained in Chapter 5 and thereby refining it. Then, comparing distinct R proteins will enable to resolve the specific and generic structural features important for their function that determine their mode of action. It is anticipated that for example by modifying specific residues in the NB-ARC domain can increase R protein sensitivity and thereby, enhance or broaden its disease resistance response in some cases.

What is the role of the LRR in pathogen recognition?

The LRR can be divided in several distinct functional surfaces, based on selection patterns, surface charge and mutant phenotypes. According to the NB-ARC-LRR docking model, both the NB and the ARC2 interact with the LRR. It will be interesting to investigate whether these predicted interactions exist and if they have different functions. Furthermore, structural results obtained in Chapter 5 show that the LRR domain can be split in two distinct parts, the C-terminal recognition part of the LRR and the N-terminal part involved in intramolecular communication. If the LRR can be accurately split in activation and recognition then it should be possible to exchange recognition specificities between unrelated R proteins as demonstrated for Gpa2 and Rx1 in Chapter 4. This knowledge could be used in creating artificial specificities, which can teach us a lot about R protein mediated recognition and activation and eventually lead to the application of engineered resistance in crop protection.

References

- Acehan, D., Jiang, X., Morgan, D.G., Heuser, J.E., Wang, X. and Akey, C.W.** (2002). Three-dimensional structure of the apoptosome: implications for assembly, procaspase-9 binding, and activation. *Mol Cell* **9**, 423-432.
- Ade, J., Deyoung, B.J., Golstein, C. and Innes, R.W.** (2007). Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease. *Proc Natl Acad Sci U S A* **104**, 2531-2536.
- Akira, S. and Takeda, K.** (2004). Toll-like receptor signalling. *Nature Reviews Immunology* **4**, 499-511.
- Bao, Q., Lu, W., Rabinowitz, J.D. and Shi, Y.** (2007). Calcium blocks formation of apoptosome by preventing nucleotide exchange in Apaf-1. *Mol Cell* **25**, 181-192.
- Bendahmane, A., Farnham, G., Moffett, P. and Baulcombe, D.C.** (2002). Constitutive gain-of-function mutants in a nucleotide binding site-leucine rich repeat protein encoded at the Rx locus of potato. *Plant J.* **32**, 195-204.
- Bent, A.F., Kunkel, B.N., Dahlbeck, D., Brown, K.L., Schmidt, R., Giraudat, J., Leung, J. and Staskawicz, B.J.** (1994). Rps2 of *Arabidopsis thaliana* - a Leucine-Rich Repeat Class of Plant-Disease Resistance Genes. *Science* **265**, 1856-1860.
- Borner, C.** (2003). The Bcl-2 protein family: sensors and checkpoints for life-or-death decisions. *Mol. Immunol.* **39**, 615-647.
- Cain, K., Bratton, S.B., Langlais, C., Walker, G., Brown, D.G., Sun, X.M. and Cohen, G.M.** (2000). Apaf-1 oligomerizes into biologically active similar to 700-kDa and inactive similar to 1.4-MDa apoptosome complexes. *J. Biol. Chem.* **275**, 6067-6070.
- Caplan, J.L., Mamillapalli, P., Burch-Smith, T.M., Czymbek, K. and Dinesh-Kumar, S.P.** (2008). Chloroplastic Protein NRIP1 Mediates Innate Immune Receptor Recognition of a Viral Effector. *Cell* **132**, 449-462.
- Danial, N.N. and Korsmeyer, S.J.** (2004). Cell death: critical control points. *Cell* **116**, 205-219.
- Danot, O., Marquet, E., Vidal-Ingigliardi, D. and Richet, E.** (2009). Wheel of Life, Wheel of Death: A Mechanistic Insight into Signaling by STAND Proteins. *Structure* **17**, 172-182.
- del Peso, L., Gonzalez, V.M., Inohara, N., Ellis, R.E. and Nunez, G.** (2000). Disruption of the CED-9 - CED-4 complex by EGL-1 is a critical step for programmed cell death in *Caenorhabditis elegans*. *J. Biol. Chem.* **275**, 27205-27211.
- Diemand, A.V. and Lupas, A.N.** (2006). Modeling AAA+ ring complexes from monomeric structures. *J Struct Biol* **156**, 230-243.
- Dinesh-Kumar, S.P. and Baker, B.J.** (2000). Alternatively spliced N resistance gene transcripts: their possible role in tobacco mosaic virus resistance. *Proc Natl Acad Sci U S A* **97**, 1908-1913.
- Dodds, P.N., Lawrence, G.J., Catanzariti, A.M., Teh, T., Wang, C.I., Ayliffe, M.A., Kobe, B. and Ellis, J.G.** (2006). Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proc Natl Acad Sci U S A* **103**, 8888-8893.
- Farnham, G. and Baulcombe, D.C.** (2006). Artificial evolution extends the spectrum of viruses that are targeted by a disease-resistance gene from potato. *Proc Natl Acad Sci U S A* **103**, 18828-18833.
- Faustin, B., Chen, Y., Zhai, D., Le Negrate, G., Lartigue, L., Satterthwait, A. and Reed, J.C.** (2009). Mechanism of Bcl-2 and Bcl-X(L) inhibition of NLRP1 inflammasome: loop domain-dependent suppression of ATP binding and oligomerization. *Proc Natl Acad Sci U S A* **106**, 3935-3940.
- Faustin, B., Lartigue, L., Bruey, J.M., Luciano, F., Sergienko, E., Bailly-Maitre, B., Volkman, N., Hanein, D., Rouiller, I. and Reed, J.C.** (2007). Reconstituted NALP1 inflammasome reveals two-step mechanism of caspase-1 activation. *Mol Cell* **25**, 713-724.
- Frost, D., Way, H., Howles, P., Luck, J., Manners, J., Hardham, A., Finnegan, J. and Ellis, J.** (2004). Tobacco transgenic for the flax rust resistance gene L expresses allele-specific activation of defense responses. *Mol Plant Microbe Interact* **17**, 224-232.
- Grant, J.J., Chini, A., Basu, D. and Loake, G.J.** (2003). Targeted activation tagging of the Arabidopsis NBS-LRR gene, ADR1, conveys resistance to virulent pathogens. *Mol Plant Microbe Interact* **16**, 669-680.
- Grant, M.R., Godiard, L., Straube, E., Ashfield, T., Lewald, J., Sattler, A., Innes, R.W. and Dangl, J.L.** (1995). Structure of the Arabidopsis Rpm1 Gene Enabling Dual-Specificity Disease Resistance. *Science* **269**, 843-846.
- Halterman, D.A. and Wise, R.P.** (2006). Upstream open reading frames of the barley Mla13 powdery mildew resistance gene function co-operatively to down-regulate translation. *Molecular Plant Pathology* **7**, 167-176.
- Hammond-Kosack, K.E. and Jones, J.D.G.** (1997). Plant disease resistance genes. In *Annu. Rev. Plant Physiol. Plant Molec. Biol.*, pp. 575-607.
- Holcik, M. and Sonenberg, N.** (2005). Translational control in stress and apoptosis. *Nat Rev Mol Cell Biol* **6**, 318-327.
- Holt, B.F., 3rd, Belkadir, Y. and Dangl, J.L.** (2005). Antagonistic control of disease resistance protein stability in the plant immune system. *Science* **309**, 929-932. Epub 2005 Jun 2003.
- Howles, P., Lawrence, G., Finnegan, J., McFadden, H., Ayliffe, M., Dodds, P. and Ellis, J.** (2005). Autoactive alleles of the flax L6 rust resistance gene induce non-race-specific rust resistance associated with the hypersensitive response. *Mol. Plant-Microbe Interact.* **18**, 570-582.
- Hu, Y., Ding, L., Spencer, D.M. and Nunez, G.** (1998). WD-40 Repeat Region Regulates Apaf-1 Self-association and Procaspase-9 Activation. *J. Biol. Chem.* **273**, 33489-33494.

- Hubert, D.A., He, Y., McNulty, B.C., Tornero, P. and Dangl, J.L.** (2009). Inaugural Article: Specific Arabidopsis HSP90.2 alleles recapitulate RAR1 cochaperone function in plant NB-LRR disease resistance protein regulation. *Proc Natl Acad Sci U S A* **106**, 9556-9563.
- Hwang, C.F., Bhakta, A.V., Truesdell, G.M., Pudlo, W.M. and Williamson, V.M.** (2000). Evidence for a role of the N terminus and leucine-rich repeat region of the Mi gene product in regulation of localized cell death. *Plant Cell* **12**, 1319-1329.
- Hwang, C.F. and Williamson, V.M.** (2003). Leucine-rich repeat-mediated intramolecular interactions in nematode recognition and cell death signaling by the tomato resistance protein Mi. *Plant J.* **34**, 585-593.
- Inohara, N., Koseki, T., del Peso, L., Hu, Y.M., Yee, C., Chen, S., Carrio, R., Merino, J., Liu, D., Ni, J. and Nunez, G.** (1999). Nod1, an Apaf-1-like activator of caspase-9 and nuclear factor- κ B. *J. Biol. Chem.* **274**, 14560-14567.
- Jiang, X.J. and Wang, X.D.** (2000). Cytochrome c promotes caspase-9 activation by inducing nucleotide binding to Apaf-1. *J. Biol. Chem.* **275**, 31199-31203.
- Jones, A.M., Thomas, V., Truman, B., Lilley, K., Mansfield, J. and Grant, M.** (2004). Specific changes in the Arabidopsis proteome in response to bacterial challenge: differentiating basal and R-gene mediated resistance. *Phytochemistry* **65**, 1805-1816.
- Kim, H.-S., Desveaux, D., Singer, A.U., Patel, P., Sondek, J. and Dangl, J.L.** (2005a). The *Pseudomonas syringae* effector AvrRpt2 cleaves its C-terminally acylated target, RIN4, from Arabidopsis membranes to block RPM1 activation. *PNAS* **102**, 6496-6501.
- Kim, H.E., Du, F.H., Fang, M. and Wang, X.D.** (2005b). Formation of apoptosome is initiated by cytochrome c-induced dATP hydrolysis and subsequent nucleotide exchange on Apaf-1. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 17545-17550.
- Koonin, E.V. and Aravind, L.** (2000). The NACHT family - a new group of predicted NTPases implicated in apoptosis and MHC transcription activation. *Trends Biochem.Sci.* **25**, 223-224.
- Leipe, D.D., Koonin, E.V. and Aravind, L.** (2004). STAND, a class of P-loop NTPases including animal and plant regulators of programmed cell death: Multiple, complex domain architectures, unusual phyletic patterns, and evolution by horizontal gene transfer. *J. Mol. Biol.* **343**, 1-28.
- Luck, J.E., Lawrence, G.J., Dodds, P.N., Shepherd, K.W. and Ellis, J.G.** (2000). Regions outside of the leucine-rich repeats of flax rust resistance proteins play a role in specificity determination. *Plant Cell* **12**, 1367-1377.
- Mackey, D., Belkhadir, Y., Alonso, J.M., Ecker, J.R. and Dangl, J.L.** (2003). Arabidopsis RIN4 is a target of the type III virulence effector AvrRpt2 and modulates RPS2-mediated resistance. *Cell* **112**, 379-389.
- Mackey, D., Holt, B.F., 3rd, Wiig, A. and Dangl, J.L.** (2002). RIN4 interacts with *Pseudomonas syringae* type III effector molecules and is required for RPM1-mediated resistance in Arabidopsis. *Cell* **108**, 743-754.
- Marquenet, E. and Richet, E.** (2007). How integration of positive and negative regulatory signals by a STAND signaling protein depends on ATP hydrolysis. *Mol Cell* **28**, 187-199.
- Mestre, P. and Baulcombe, D.C.** (2006). Elicitor-mediated oligomerization of the tobacco N disease resistance protein. *Plant Cell* **18**, 491-501.
- Meyers, B.C., Kozik, A., Griego, A., Kuang, H. and Michelmore, R.W.** (2003). Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. *Plant Cell* **15**, 809-834.
- Mindrinis, M., Katagiri, F., Yu, G.L. and Ausubel, F.M.** (1994). The *A. thaliana* disease resistance gene RPS2 encodes a protein containing a nucleotide-binding site and leucine-rich repeats. *Cell* **78**, 1089-1099.
- Moffett, P., Farnham, G., Peart, J. and Baulcombe, D.C.** (2002). Interaction between domains of a plant NBS-LRR protein in disease resistance-related cell death. *Embo J.* **21**, 4511-4519.
- Mucyn, T.S., Clemente, A., Andriotis, V.M., Balmuth, A.L., Oldroyd, G.E., Staskawicz, B.J. and Rathjen, J.P.** (2006). The tomato NBARC-LRR protein Prf interacts with Pto kinase in vivo to regulate specific plant immunity. *Plant Cell* **18**, 2792-2806.
- Qin, H., Srinivasula, S.M., Wu, G., Fernandes-Alnemri, T., Alnemri, E.S. and Shi, Y.** (1999). Structural basis of procaspase-9 recruitment by the apoptotic protease-activating factor 1. *Nature* **399**, 549-557.
- Rairdan, G.J., Collier, S.M., Sacco, M.A., Baldwin, T.T., Boettrich, T. and Moffett, P.** (2008). The Coiled-Coil and Nucleotide Binding Domains of the Potato Rx Disease Resistance Protein Function in Pathogen Recognition and Signaling. *Plant Cell* **20**, 739-751.
- Rairdan, G.J. and Moffett, P.** (2006). Distinct Domains in the ARC Region of the Potato Resistance Protein Rx Mediate LRR Binding and Inhibition of Activation. *Plant Cell* **18**, 2082-2093.
- Riedl, S.J., Li, W.Y., Chao, Y., Schwarzenbacher, R. and Shi, Y.G.** (2005). Structure of the apoptotic protease-activating factor 1 bound to ADP. *Nature* **434**, 926-933.
- Sacco, M.A., Mansoor, S. and Moffett, P.** (2007). A RanGAP protein physically interacts with the NB-LRR protein Rx, and is required for Rx-mediated viral resistance. *Plant J* **52**, 82-93.
- Salmeron, J.M., Oldroyd, G.E., Rommens, C.M., Scofield, S.R., Kim, H.S., Lavelle, D.T., Dahlbeck, D. and Staskawicz, B.J.** (1996). Tomato Prf is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the Pto Kinase gene cluster. *Cell* **86**, 123-133.
- Seiffert, B.M., Vier, J. and Hacker, G.** (2002). Subcellular localization, oligomerization, and ATP-binding of *Caenorhabditis elegans* CED-4. *Biochem Biophys Res Commun* **290**, 359-365.
- Shao, F., Golstein, C., Ade, J., Stoutemyer, M., Dixon, J.E. and Innes, R.W.** (2003). Cleavage of Arabidopsis PBS1 by a bacterial type III effector. *Science* **301**, 1230-1233.

- Shiozaki, E.N., Chai, J.J. and Shi, Y.** (2002). Oligomerization and activation of caspase-9, induced by Apaf-1 CARD. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 4197-4202.
- Steebhorn, C., Danot, O., Huber, R. and Clausen, T.** (2001). Crystal structure of transcription factor MalT domain III: a novel helix repeat fold implicated in regulated oligomerization. *Structure* **9**, 1051-1060.
- Swiderski, M.R., Birker, D. and Jones, J.D.** (2009). The TIR domain of TIR-NB-LRR resistance proteins is a signalling domain involved in cell death induction. *Mol Plant Microbe Interact* **22**, 157-165.
- Takabatake, R., Seo, S., Mitsuhashi, I., Tsuda, S. and Ohashi, Y.** (2006). Accumulation of the Two Transcripts of the N gene, Conferring Resistance to Tobacco Mosaic Virus, is Probably Important for N Gene-dependent Hypersensitive Cell Death. *Plant Cell Physiol* **47**, 254-261.
- Takken, F.L., Albrecht, M. and Tameling, W.I.** (2006). Resistance proteins: molecular switches of plant defence. *Curr Opin Plant Biol* **9**, 383-390.
- Tameling, W.I. and Baulcombe, D.C.** (2007). Physical association of the NB-LRR resistance protein Rx with a Ran GTPase-activating protein is required for extreme resistance to Potato virus X. *Plant Cell* **19**, 1682-1694.
- Tameling, W.I.L., Elzinga, S.D.J., Darmin, P.S., Vossen, J.H., Takken, F.L.W., Haring, M.A. and Cornelissen, B.J.C.** (2002). The tomato R gene products I-2 and Mi-1 are functional ATP binding proteins with ATPase activity. *Plant Cell* **14**, 2929-2939.
- Tameling, W.I.L., Vossen, J.H., Albrecht, M., Lengauer, T., Berden, J.A., Haring, M.A., Cornelissen, B.J.C. and Takken, F.L.W.** (2006). Mutations in the NB-ARC Domain of I-2 That Impair ATP Hydrolysis Cause Autoactivation. *Plant Physiol.* **140**, 1233-1245.
- Tan, X., Meyers, B.C., Kozik, A., West, M.A., Morgante, M., St Clair, D.A., Bent, A.F. and Michelmore, R.W.** (2007). Global expression analysis of nucleotide binding site-leucine rich repeat-encoding and related genes in Arabidopsis. *BMC Plant Biol* **7**, 56.
- Tao, Y., Yuan, F.H., Leister, R.T., Ausubel, F.M. and Katagiri, F.** (2000). Mutational analysis of the Arabidopsis nucleotide binding site- leucine-rich repeat resistance gene RPS2. *Plant Cell* **12**, 2541-2554.
- van der Biezen, E.A. and Jones, J.D.G.** (1998). The NB-ARC domain: A novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. *Curr. Biol.* **8**, R226-R227.
- van Ooijen, G., Mayr, G., Albrecht, M., Cornelissen, B.J.C. and Takken, F.L.W.** (2008a). Transcomplementation, but not Physical Association of the CC-NB-ARC and LRR Domains of Tomato R Protein Mi-1.2 is Altered by Mutations in the ARC2 Subdomain. *Mol Plant* **1**, 401-410.
- van Ooijen, G., Mayr, G., Kasiem, M.M.A., Albrecht, M., Cornelissen, B.J.C. and Takken, F.L.W.** (2008b). Structure-function analysis of the NB-ARC domain of plant disease resistance proteins. *J. Exp. Bot.* **59**, 1383-1397.
- Wagner, R.N., Proell, M., Kufer, T.A. and Schwarzenbacher, R.** (2009). Evaluation of Nod-like receptor (NLR) effector domain interactions. *PLoS ONE* **4**, e4931.
- Warren, R.F., Henk, A., Mowery, P., Holub, E. and Innes, R.W.** (1998). A mutation within the leucine-rich repeat domain of the arabidopsis disease resistance gene RPS5 partially suppresses multiple bacterial and downy mildew resistance genes. *Plant Cell* **10**, 1439-1452.
- Weaver, M.L., Swiderski, M.R., Li, Y. and Jones, J.D.G.** (2006). The Arabidopsis thaliana TIR-NB-LRR R-protein, RPP1A: protein localization and constitutive activation of defence by truncated alleles in tobacco and Arabidopsis. *The Plant Journal* **47**, 829-840.
- Whitham, S., Dinesh-Kumar, S.P., Choi, D., Hehl, R., Corr, C. and Baker, B.** (1994). The product of the tobacco mosaic virus resistance gene N: similarity to toll and the interleukin-1 receptor. *Cell* **78**, 1101-1115.
- Yan, N., Chai, J.J., Lee, E.S., Gu, L.C., Liu, Q., He, J.Q., Wu, J.W., Kokel, D., Li, H.L., Hao, Q., Xue, D. and Shi, Y.G.** (2005). Structure of the CED-4-CED-9 complex provides insights into programmed cell death in *Caenorhabditis elegans*. *Nature* **437**, 831-837.
- Yan, N. and Shi, Y.** (2005). Mechanisms of apoptosis through structural biology. *Annu Rev Cell Dev Biol* **21**, 35-56.
- Yan, N.G., Gu, L.C., Kokel, D., Chai, J.J., Li, W.Y., Han, A.D., Chen, L., Xue, D.N. and Shi, Y.G.** (2004). Structural, biochemical, and functional analyses of CED-9 recognition by the proapoptotic proteins EGL-1 and CED-4. *Mol. Cell.* **15**, 999-1006.
- Yu, X., Acehan, D., Menetret, J.F., Booth, C.R., Ludtke, S.J., Riedl, S.J., Shi, Y., Wang, X. and Akey, C.W.** (2005). A structure of the human apoptosome at 12.8 Å resolution provides insights into this cell death platform. *Structure (Camb)* **13**, 1725-1735.
- Zhang, X.C. and Gassmann, W.** (2003). RPS4-mediated disease resistance requires the combined presence of RPS4 transcripts with full-length and truncated open reading frames. *Plant Cell* **15**, 2333-2342.
- Zhang, X.C. and Gassmann, W.** (2007). Alternative splicing and mRNA levels of the disease resistance gene RPS4 are induced during defense responses. *Plant Physiol* **145**, 1577-1587.
- Zhang, Y.** (2009). Protein structure prediction: when is it useful? *Curr. Opin. Struct. Biol.* **19**, 145-155.
- Zhang, Y., Dorey, S., Swiderski, M. and Jones, J.D.G.** (2004). Expression of RPS4 in tobacco induces an AvrRps4-independent HR that requires EDS1, SGT1 and HSP90. *Plant J* **40**, 213-224.

Summary

Resistance proteins are part of the plant's immune system and mediate a defence response upon recognizing their cognate pathogens. They are thought to be present in the cell as part of a larger protein complex. The modular architecture of R proteins suggests that they form a scaffold for various interacting proteins, involved in pathogen recognition, downstream signalling or protein stabilization. However, few common interactors have been found for the CC-NB-ARC domains despite extensive screenings for downstream interactors. The objective of this thesis was to get new insights in the structure, function and localization of R proteins by using the potato resistance genes Rx1 and Gpa2 as a model system. Initially, a novel T7 phage display method was developed to facilitate high throughput selection of interacting molecules (Chapter 2). However, the use of a T7 cDNA phage library to identify interactors of the CC-NB-ARC domains of Rx1 resulted in the discovery of a large set of highly basic peptides (Chapter 3). In Chapter 4, the functional role of the CC-NB-ARC domain in mediating disease resistance was explored by creating chimeric proteins between Rx1 and Gpa2. This resulted in the observation that the CC-NB-ARC is able to confer both virus and nematode resistance in potato. Furthermore, it was shown that the CC-NB-ARC of Rx1 and the LRR of Gpa2 are incompatible and *vice versa*. This phenomenon was studied in more detail in Chapter 5, in which a docking model for the interacting surface of these domains was constructed based on the individual structural domains. Finally, the subcellular localisation was investigated to get a better understanding about the R proteins function in the cell (Chapter 6).

The lytic T7 phages form a powerful platform for the display of large cDNA libraries and offer the possibility to screen for strong interactions with a variety of substrates. To visualise these interactions directly by fluorescence microscopy, we constructed fluorescent T7 phages by exploiting the flexibility of phages to incorporate modified versions of its capsid protein (**Chapter 2**). By applying translational frameshift sequences, helper plasmids were constructed that expressed a fixed ratio of both wild-type capsid protein (gp10) and capsid protein fused to enhanced yellow fluorescent protein (EYFP). The frameshift sequences were inserted between the 3'-end of the capsid gene and the sequence encoding EYFP. Fluorescent fusion proteins are only formed when the ribosome makes a -1 shift in reading frame during translation. As far as we know this is the first report of using a translational frameshift for a biotechnological purpose. The phages formed in this way have capsids composed of three different variants of their capsid protein; EYFP-fused versions derived by frameshift translation, non-fused versions derived by regular translation from the helper plasmid, and versions that display peptides encoded in the library ligated in the phage genome. Using standard fluorescence microscopy, we could sensitively monitor the enrichment of specific binders in a cDNA library displayed on fluorescent T7 phages. Closely monitoring the effect of the selection procedure enables fine tuning, and obviates the need for more laborious ELISA or plaque lift assays. Furthermore, with the fast pace of developments in single molecule detection technologies and sorting systems, these fluorescent phages open the way to high throughput platforms for the direct selection of binding molecules.

In **Chapter 3**, cDNA phage display was applied as an alternative method to identify additional downstream Rx1 interactors, which could further resolve the Rx1 signalling pathway. In a pilot experiment the value of T7 phage display to identify specific interactors was demonstrated by using an antibody raised against the PVY coat protein. Screening of a PVY-infected *N. benthamiana* cDNA phage display library resulted in the selection of peptides harbouring the known PRIKAI epitope. Next, phage display was explored as technique to discover proteins interacting with the potato R protein Rx1. The system turned out to be prone to pick up interactors binding to matrices like Ni-NTA or to fusion proteins like thioredoxine. A possible way to circumvent this weakness was to design the selection procedure in such a way that it alternates between different matrices and to limit the number of selection round. This adapted approach resulted in the identification of a series of highly basic protein fragments and random peptides, for which a specific interaction could be shown. Two cDNA sequences encoded the ribosomal proteins L19 and L36a, which showed a stunted growth phenotype upon gene silencing in *N. benthamiana* using VIGS and a slightly reduced Rx-mediated HR.

The nematode resistance protein Gpa2 and the virus resistance protein Rx1 provide an excellent test system to investigate the exchangeability of recognition and signalling domains and explore the evolutionary flexibility of R proteins, for they confer resistance to completely unrelated pathogens (*Globodera pallida* and potato virus X, respectively). In Chapter 4, we provide evidence for the hypothesis, that, via intergenic sequence exchanges and various types of mutations, NB-LRR proteins have the potential to alter resistance specificities towards taxonomically unrelated pathogens in relatively short evolutionary time periods. Both the regulatory sequences and CC-NB domains of the paralogs Gpa2 and Rx1 are non-pathogen specific and exchangeable. Remarkably, the genetic fusions of the CC-NB of *Rx1* with the LRR of *Gpa2* (Rx1_{CN}/Gpa2_L) and the reciprocal domain swap (Gpa2_{CN}/Rx1_L) were not functional when driven by the endogenous promoters or 35S promoter. Gain of wild type resistance was obtained by re-introducing the first five LRRs of Rx1 in Rx1_{CN}/Gpa2, restoring the compatibility between the N-terminal part of the LRR and the ARC2 domain. Decreasing the expression levels for Gpa2_{CN}/Rx1_L resulted in extreme resistance against PVX, indistinguishable from wild type plants. Our results indicate that not only coding sequences, but that also optimizing the expression levels may play a role in generating novel resistances.

The CC, NB-ARC, and LRR domains of the Rx1 and Gpa2 proteins interact with each other and recognition of the elicitor mediated by the LRR is translated in an activation of the NB-ARC. The available functional and evolutionary data make Gpa2 a suitable candidate for structural modelling of the individual domains and their interaction (Chapter 5). A structural model of the NB-ARC / LRR interaction could function as a framework for the interpretation of known empirical data and the design of new experiments to test R protein operational mechanisms (Zhang, 2009). Therefore, computer aided modelling of the 3D structure domain models for the NB-ARC and the LRR domains were obtained and used as basis for a domain docking study. The functional interaction between the domains was studied via a detailed analysis of their incompatibility in chimeric Gpa2 and Rx1 proteins. A large set of sequence exchanges between the two proteins was created for that purpose. Both in the LRR and in the ARC2 domain small regions could be identified in which the amino

acids differing between Gpa2 and Rx1 led to domain incompatibility. Five of the ARC2 positions required for LRR compatibility and three known autoactivating positions from the RX1 LRR were used as constraints in domain docking computation to limit the potential search space. The resulting docking model indicated an important role in the NB-ARC-LRR interaction for electrostatic and hydrophobic interactions. A loop region rich in acidic residues in the ARC2 domain was found close in space to a patch of basic residues grouped together in the LRR. Hydrophobic residues on both the NB and the ARC2 contacted hydrophobic residues on the surface of the LRR. A correlation analysis of the NB-ARC and LRR subdomains detected coevolution between the interacting surfaces, which supports a direct interaction between these two domains. Site-directed mutagenesis and pull-down experiments were used to test the role of surface features that might play an important role in the interdomain docking interface.

In Chapter 6, we have made use of the characteristic of the Rx1 protein that it remains functional when its domains are co-expressed as separate polypeptides. This allowed us to create fluorescent constructs, not only of the full length protein, but also of the separate subdomains. Most of these tagged constructs still form functional proteins. C- and N-terminal fusions of Green Fluorescent Protein (GFP) variants to Rx, made it possible to study its subcellular localization in *Nicotiana benthamiana* cells. Contrary to our expectations we observed the presence of Rx1 in both the cytoplasm and the nucleus. Rx1 does not contain known nuclear localization signals and the size of the protein (140 kDa including GFP) exceeds the limit for passive diffusion through the nuclear pore. Fluorescent fusions of a series of deletion constructs, CC-NB-ARC, NB-ARC, NB-ARC-LRR, CC and LRR showed three distinct patterns of subcellular localization. The NB-ARC-LRR and LRR constructs have a cytoplasmic localization and are mostly absent in the nucleus. The NB-ARC and CC-NB-ARC constructs showed equal fluorescence intensities in both the nucleus and the cytoplasm. The CC alone fused to GFP, however, seems to preferentially accumulate in the nucleus resulting in a three to four times higher fluorescence intensity in the nucleus compared to the cytoplasm. The diffusion behaviour inside the nucleus for both the complete CC and a CC fragment containing the two predicted helices downstream of the central turn, showed that their nuclear accumulation coincides with a significantly reduced nuclear diffusion as compared to unfused GFP and the other CC fragments. This difference might point to a potential interaction between the CC and an unknown nuclear component. Furthermore, SGT1 and Rar1 are thought to function as chaperones involved in stabilizing R proteins. Both the silencing experiments with these two proteins and the P-loop mutation show that the nuclear localisation of Rx1 is probably conformation dependent. Two approaches were followed to see if CP recognition or Rx1 signalling pathway were linked to a certain cellular compartment. At one hand the Rx1 protein itself or its subdomains were directed to either the nucleus or the cytoplasm by fusion to exogenous targeting signals (Nuclear Export Signals or Nuclear Localization Signals). On the other hand the elicitor, the PVX coat protein, was directed to the nucleus or cytoplasm. The PVX coat protein is a much smaller protein and can under normal circumstances diffuse freely between the cytoplasm and the nucleus. The surprising result was that no effect was found for retargeting Rx1, but when the elicitor was targeted to the nucleus, it could not activate Rx1 anymore, indicating that recognition might to take place in the cytoplasm.

In the final chapter, the results obtained in this thesis are put into perspective by studying parallels in scientific literature on NB-LRR proteins with similar functions in other organisms.

Samenvatting

Resistentie-eiwitten (R eiwitten) vormen een belangrijk onderdeel van het immuunsysteem van planten. Ze kunnen een verdedigingsreactie in gang zetten na directe of indirecte herkenning van ziekteverwekkers die de plant binnendringen. Er zijn aanwijzingen dat R eiwitten functioneren als onderdeel van een groter eiwitcomplex. De modulaire opbouw van R eiwitten uit structureel en functioneel van elkaar te onderscheiden domeinen suggereert dat ze een platform vormen waaraan verschillende andere eiwitten kunnen binden die betrokken zijn bij pathogeen herkenning, eiwitvouwing of het doorgeven van signalen. Ondanks uitgebreide studies zijn er nog relatief weinig eiwitten gevonden die een gemeenschappelijk onderdeel zijn van diverse R eiwit complexen. Het belangrijkste doel van het onderzoek, beschreven in dit proefschrift, was het verkrijgen van inzicht in de structuur, de functionele mechanismen en de rol van de subcellulaire lokalisatie van R eiwitten. Als model systeem zijn hiervoor de Rx1 en Gpa2 bestudeerd, twee nauw verwante R eiwitten afkomstig uit aardappel. In eerste instantie is een op fluorescente T7 bacteriofagen gebaseerd systeem opgezet dat de selectie van bindende eiwitten via *phage display* technologie moest vereenvoudigen (Hoofdstuk 2). Het gebruik van T7 cDNA *phage display* om eiwitten te selecteren die binden aan de N-terminale CC-NB-ARC domeinen van Rx1, leverde een grote verzameling sterk positief geladen peptiden op (Hoofdstuk 3). In Hoofdstuk 4 is vervolgens de rol van de CC-NB-ARC domeinen bij het in gang zetten van de resistentie reactie verkend door middel van het uitwisselen van de domeinen tussen Gpa2 en Rx1. Hieruit kon worden geconcludeerd dat de CC-NB-ARC domeinen zowel een zeer efficiënte resistentie tegen virussen als een resistentie tegen parasitaire aaltjes kunnen aanschakelen. De CC-NB-ARC en LRR domeinen van Rx1 en Gpa2 waren echter niet zomaar met elkaar verenigbaar in een eiwit, wat eiwit varianten opleverde die of continu aangeschakeld waren, of juist niet meer aangeschakeld konden worden. Dit fenomeen is in meer detail bestudeerd in Hoofdstuk 5, waarin de ruimtelijke structuur van de NB-ARC en LRR domeinen beschreven wordt. De exacte regio's binnen de domeinen die bepalend zijn voor een goede samenwerking tussen de domeinen, zijn als uitgangspunt genomen voor een ruimtelijk model van de onderlinge binding van de NB-ARC en LRR domeinen. De lokalisatie van het R eiwit binnen de cel en de rol die de verschillende domeinen daarbij spelen is nader onderzocht in Hoofdstuk 6.

De lytische T7 bacterievirussen zijn bijzonder geschikt als platform voor het selecteren van eiwitten, die aan allerhande substraten binden. Eiwitten die gecodeerd worden door cDNA dat geplaatst is in het genoom van een faag worden aangeboden op het oppervlak van het deze faag. Om de bindingen tussen de eiwitten op de fagen en een substraat zichtbaar te maken door middel van fluorescentie microscopie, hebben we fluorescente T7 fagen geconstrueerd (**Hoofdstuk 2**). Hierbij is gebruik gemaakt van de flexibiliteit van fagen in het opnemen van aangepaste versies van de T7 manteleiwitten. Met behulp van translationele *frameshift* sequenties zijn helper plasmiden gemaakt die vanaf één gen zowel het wild type manteleiwit als een fusieproduct met een geel fluorescent eiwit (*enhanced yellow fluorescent protein*, EYFP) tot expressie brachten in verschillende vastliggende ratio's. De *frameshift* sequenties waren geplaatst tussen het 3'-einde van het gen dat codeert voor het T7 manteleiwit en de EYFP coderende DNA sequentie. Alleen wanneer het ribosoom zich in

leesraam (*reading frame*) vergist op de *frameshift* sequentie en een nucleotide terug schuift, wordt niet een stopcodon gelezen, maar leest het ribosoom door in het EYFP gen. Zover wij weten is dit de eerste maal dat een dergelijke frameshift gebruikt is met een biotechnologisch doel. De bacteriofagen die gevormd worden bij gebruik van dit helperplasmide hebben eiwitmantels die bestaan uit drie verschillende versies van het manteleiwit; EYFP gefuseerde versies die ontstaan zijn door de verschuiving in leesraam, ongefuseerde versies die gevormd worden wanneer het leesraam niet verschuift en versies die gefuseerd zijn aan eiwitten die gecodeerd zijn in het genoom van de faag. De beschreven toevoeging van EYFP aan het manteleiwit maakte het mogelijk met standaard fluorescentie microscopische technieken de verrijking in bindende fagen in een *phage display* cDNA bibliotheek te volgen tijdens een selectieprocedure. Het direct gedetailleerd volgen van verrijking in binders in reactie op een selectieprocedure maakt het mogelijk om de condities precies af te stellen zonder de noodzaak om meer bewerkelijke methoden als ELISA of *plaque lift assays* toe te passen. Verder zouden fluorescente fagen, met het oog op de snelle ontwikkelingen in moleculaire sorteersystemen en de toenemende detectiegevoeligheid van microscopen, de deur kunnen openen naar zogenaamde *high throughput* systemen die directe binding van moleculen kunnen selecteren.

In **Hoofdstuk 3** is T7 cDNA *phage display* gebruikt als alternatieve methode om eiwitten te identificeren die het resistentie eiwit Rx1 binden. Deze bindende eiwitten zouden inzicht kunnen geven in de manier waarop signalen worden doorgegeven via dit type R eiwitten. In een proefexperiment, waarbij gezocht is in een cDNA bibliotheek naar eiwitfragmenten die specifiek bonden aan een antilichaam, ooit ontwikkeld om het aardappelvirus PVY te herkennen, bewees deze techniek zijn waarde. Uit de cDNA bibliotheek, samengesteld uit cDNA van PVY geïnfecteerde *Nicotiana benthamiana* planten, kon een serie peptiden geselecteerd worden die allemaal het al bekende epitope (PRIKAI) van het antilichaam bevatten. De selectie van Rx1-bindende eiwitten bleek lastiger. De relatief zwakke eiwit-eiwit interacties ondergaan veel competitie van bindingen met onder andere de matrix waaraan het Rx1 eiwit bevestigd was. Dit probleem kon omzeild worden door tijdens de selectie procedure af te wisselen tussen verschillende bindingsmatrices en door de hoeveelheid selectie rondes te beperken. Gevolg hiervan was wel dat bewerkelijke controle stappen noodzakelijk waren om specifiek bindende (poly-)peptiden te identificeren. Uiteindelijk werden een serie eiwitfragmenten en peptiden gevonden die bonden aan de CC-NB-ARC domeinen van Rx1. *Pull-down* experimenten toonden aan dat deze interactoren specifiek waren voor Rx1. Een opvallende overeenkomst tussen al deze bindende eiwitfragmenten was hun hoge gehalte aan de basische aminozuren arginine en lysine. Twee van de cDNA fragmenten codeerden voor de ribosomale eiwitten L19 en L36a. Het uitschakelen van de expressie van deze eiwitten in planten veroorzaakte een vrijwel volledige blokkering van de groei van de planten en een vermindering van de door de activatie van Rx1 veroorzaakte celdood.

De R eiwitten Gpa2 en Rx1 vormen een perfect systeem om de uitwisselbaarheid van herkennings- en signaaldomeinen te testen en de evolutionaire flexibiliteit van R eiwitten te onderzoeken. Ze geven in de plant resistentie tegen ongerelateerde ziekteverwekkers, namelijk het aaltje *Globodera pallida* en het virus PVX, maar zijn wat betreft aminozuur

volgorde sterk aan elkaar gelijk. De resultaten die we presenteren in **Hoofdstuk 4** ondersteunen de hypothese dat R eiwitten via mutaties en DNA uitwisselingen tussen hun genen, in relatief korte tijd herkenningsspecificiteiten kunnen ontwikkelen tegen taxonomisch niet verwante ziekteverwekkers. Zowel de DNA sequenties die de expressie van beide genen controleren als de CC-NB-ARC domeinen van Gpa2 en Rx1 zijn uitwisselbaar en niet specifiek voor een bepaalde ziekteverwekker. Opvallend genoeg zijn de CC-NB-ARC en LRR domeinen van Gpa2 en Rx1 niet zomaar met elkaar verenigbaar in een eiwit. Wanneer de CC-NB-ARC van Gpa2 gecombineerd werd met de LRR van Rx1 resulteerde dit in een eiwit dat continu actief was en daardoor celdood veroorzaakte in de cellen waarin het tot expressie kwam. Het complementaire construct, met de CC-NB-ARC van Rx1 en de LRR van Gpa2, verloor juist alle functie. De autoactiviteit van het construct met de Gpa2 CC-NB-ARC kon voorkomen worden door de expressie niveaus van het eiwit sterk te verlagen. Daarbij herwon het eiwit weer de mogelijkheid specifiek op het virus PVX te reageren en resistentie te geven in transgene aardappelplanten. De inactiviteit van het construct dat de LRR van Gpa2 bevatte kon hersteld worden door het eerste deel van de LRR (de eerste vijf LRR *repeats*) te vervangen door de Rx1 sequentie. Blijkbaar vormt het NB-ARC domein een functioneel geheel met het eerste deel van de LRR en leiden verstoringen daarin tot een verlies van functie. Deze resultaten laten zien dat er functionele beperkingen bestaan voor het uitwisselen van R eiwit fragmenten, maar dat onder andere het afstemmen van de R eiwitniveaus in de cel een belangrijke rol kan spelen bij de het ontstaan van nieuwe resistentie specificiteiten.

De CC (*coiled coil*), NB-ARC (nucleotide bindend domein) en de LRR van Rx1 en Gpa2 binden elkaar en herkenning van de bijbehorende *elicitor*, die afkomstig is van het pathogeen, door de LRR, wordt binnen het eiwit vertaald naar een activatie van het NB-ARC domein. De beschikbaarheid van functionele en evolutionaire kennis maken Gpa2 een geschikte kandidaat voor de modellering van de ruimtelijke structuur van de losse domeinen, maar ook hun onderlinge binding (**Hoofdstuk 5**). Een structuurmodel van de binding tussen de NB-ARC en de LRR zou kunnen dienen als raamwerk voor de interpretatie van bekende empirische gegevens en voor het ontwikkelen van nieuwe hypothesen en experimenten om de werking van R eiwitten te achterhalen. Om die reden zijn met behulp van computerprogramma's structuurmodellen gemaakt van de R eiwit domeinen, die gebaseerd zijn op vergelijkbare, empirisch bepaalde structuren. Deze structuurmodellen vormden vervolgens weer de basis voor het modelleren van de binding tussen de NB-ARC en LRR domeinen.

De functionele interactie tussen de NB-ARC en LRR was nader bestudeerd in een gedetailleerde analyse van de onverenigbaarheid van de Gpa2 en Rx1 domeinen die al in hoofdstuk 4 beschreven was. Binnen zowel de NB-ARC als de LRR konden kleine regio's aangewezen worden die onderling afhankelijk waren. De regio in het ARC2 domein van de NB-ARC dat voor correct functioneren gecombineerd moet zijn met de bijbehorende LRR bleek in maar zeven aminozuur posities te verschillen tussen Gpa2 en Rx1. Vijf van deze zeven lagen aan de oppervlakte van de structuur en zijn gebruikt als uitgangspunt voor het modelleren van domeinbinding. Vanaf de kant van de LRR bleek een klein gebied bestaande uit de eerste drie LRR *repeats* gecombineerd te moeten worden met de bijpassende NB-ARC

voor een juiste werking. Verder zijn voor de modellering posities van drie activerende mutaties gebruikt die al eerder gepubliceerd waren.

Het resulterende model voor de NB-ARC/LRR binding liet zien dat deze waarschijnlijk voor een belangrijk deel bepaald wordt door zowel elektrostatische als hydrofobe interacties. Een lus-structuur in het ARC2 subdomein, rijk aan negatief geladen aminozuren, lag in het bindingsmodel dicht in de buurt van een grote groep basische aminozuren op het oppervlak van de LRR in het bindingsmodel. Hydrofobe aminozuren in het oppervlak van zowel het NB als het ARC2 subdomein bleken te binden aan hydrofobe aminozuren in het oppervlak van de binnenzijde van de hoefijzervormige LRR structuur. Correlaties tussen mutaties in de bindende oppervlakken wijzen op co-evolutie tussen de domeinen binnen deze eiwitten. Gerichte mutagenese, functionele testen en interactie proeven zijn gebruikt om het belang van deze oppervlakte eigenschappen in de binding aan te tonen en zo het interactie model te toetsen.

In **Hoofdstuk 6** is bepaald op welke plaatsen in de cel het R eiwit Rx1 voorkomt door het te fuseren met fluorescente eiwitten, die zijn afgeleid van het groen fluorescente eiwit GFP dat afkomstig is uit de kwal *Aequorea victoria*. Rx1 heeft de interessante eigenschap dat het zelfs wanneer het in gescheiden onderdelen tot expressie wordt gebracht toch door de onderlinge binding van de domeinen in de cel nog een functionerend eiwit kan vormen. Deze eigenschap konden we gebruiken om te bepalen hoe de verschillende functionele domeinen bijdroegen aan de subcellulaire lokalisatie van Rx1. De meeste fluorescente fusies van losse domeinen die we maakten behielden hun functionaliteit. Het volledige eiwit troffen we, ondanks zijn grootte aan in zowel het cytoplasma als de kern van de cel. Rx1 bevat geen bekende kern lokalisatie sequentie en is groter (met GFP fusie 140 kDa) dan de maximale doorgang voor passieve diffusie de kern in (40-60 kDa). De fluorescente fusies van losse domeinen of combinaties van domeinen lieten drie verschillende lokalisatie patronen zien. De CC-NB-ARC, NB-ARC en het volledig eiwit waren gelijkmatig (wat betreft fluorescentieniveau) verdeeld tussen de kern en het cytoplasma. De LRR en de NB-ARC-LRR constructen konden voornamelijk in het cytoplasma worden waargenomen en waren grotendeels afwezig in de kern. Het CC domein alleen echter, toonde een grote concentratie in de kern, hoger dan wat voor vrij GFP werd waargenomen. Zowel de CC als een kleiner fragment van de CC, dat alleen de twee helices bevatte die voorspeld worden voor de tweede helft van het domein, hadden in de kern een significant lagere diffusiesnelheid dan vrij GFP. Dit zou er op kunnen duiden dat de CC via deze twee helices bindt aan een nog onbekende component van de kern. Verder bleek dat zowel het uitschakelen van de mogelijke chaperone SGT1 via *virus-induced gene silencing* (VIGS) als het blokkeren van de binding van nucleotiden aan het Rx1 NB domein door een gerichte mutatie daarin ervoor zorgden dat het volledig Rx1 vrijwel niet meer in de kern voorkwam. Deze verschuiving in lokalisatie bleek afhankelijk van het LRR domein.

Twee strategieën zijn gevolgd om te bestuderen of PVX herkenning en Rx1 activatie en het doorgeven van een signaal verbonden waren met een bepaalde subcellulaire lokalisatie. Enerzijds zijn zowel Rx1 als zijn losse domeinen naar, ofwel de kern, ofwel het cytoplasma gedirigeerd via de fusie van signaalsequenties (Nuclear Export Signal; NES, of Nuclear Localization Signal; NLS). Anderzijds is de *elicitor*, het PVX manteleiwit, op vergelijkbare

wijze naar de kern of het cytoplasma gedwongen. Opvallend genoeg werd geen effect gevonden voor de afgedwongen lokalisaties van Rx1 zelf, hoewel fluorescentie microscopie liet zien dat de signaalsequenties functioneerden. Echter, het naar de kern sturen van het PVX manteleiwit zorgde ervoor dat het Rx1 niet meer werd geactiveerd. Dit laat naar ons idee zien dat de in ieder geval de activatie van Rx1 door het manteleiwit van PVX in het cytoplasma moet plaatsvinden. Mogelijk omdat bepaalde samenwerkende eiwitten zoals de Rx1 interactor RanGAP2 voornamelijk in het cytoplasma gevonden worden.

In het laatste hoofdstuk, **Hoofdstuk 7**, worden de resultaten gevonden in dit proefschrift besproken in de context van de wetenschappelijke literatuur en de rol die verschillende functionele domeinen van NB-LRR eiwitten spelen in hun activatiecyclus.

Acknowledgments

Finishing this thesis has been a long and interesting process. This is a good moment to look back to all those years and acknowledge the many people who have contributed to it along the way. I think I could not have written this thesis without their support. First of all I would like to thank the people who did a lot of the lab work. Jan Roosien is the mastermind behind most of the DNA constructs found in this thesis (to get an idea of the enormous amount of work involved, one should look at the methods section of chapters 5 and 6). Jan reads DNA sequences like most people read the newspaper and elevated cloning to a form of art. Rikus Pomp helped me a lot with the protein work. First with technical advice and by teaching me how to use the Äkta, and since last year with co-immunoprecipitation experiments that are becoming a crucial part of our research. And of course, both Jan and Rikus were also always there to answer difficult questions from our students in the lab. Casper van Schaik and Liesbeth Bouwman created the transgenic potato plants used to study our R proteins in their native background. Transforming potato plants is a laborious process, which can take up to half a year, but which is essential for really understanding the proteins we study. Jan Willem Borst and Mark Hink from the Microspectroscopy centre and Jan Vos from the Laboratory of Plant Cell Biology helped me with the microscopic and microspectroscopic techniques like FCS and FRAP. The colourful results from confocal microscopy fill most of chapter 6. Not every experiment we attempted worked eventually, or made it to this thesis, but the ones that did, made me feel that the many hours behind the microscopes were well spent. Already some new and promising projects involving microspectroscopic techniques are under way and I look forward to extending our cooperations. The presented structural models of the Gpa2 domains and of their interaction are the work of Laurentiu Spiridon and Andrei Petrescu from the Institute of Biochemistry of the Romanian Academy. This collaboration in which their structural models help us design new experiments and our experimental data helps them refine the models has been extremely fruitful. With the availability of this structural model we can start forming hypotheses on the actual mechanism of activation of these protein and explain various phenotypes we observed in previous experiments. I think this is also the place to thank for the wonderful time Aska, Jan and I had in Romania last summer when we came to visit your laboratory and you gave us a guided tour through the Carpathian Mountains and Bucharest. During the early years of my PhD project I worked in the group of Arjen Schots, the LMA (which stands for the unexpectedly long Laboratory for Molecular recognition and Antibody technology). The LMA group further consisted of Hans, Willemien, Diego and Rikus. I still appreciate the good atmosphere within our little group, the lunches, dinners, movies we went to, etc. I have good memories of the times Hans and I visited phage display congresses in Boston or EU meetings in Vienna, where we always stayed a few days extra to explore these cities. Our spontaneous brainstorm sessions (especially on Friday afternoons) could lead to interesting ideas, and chapter 1 shows that following up on such an idea can even result in good publications. At the LMA we had the tradition of ending the work week with a beer on Friday. Rikus and I are still keeping this tradition alive, so if you are in town, you know where to find us. Of course I should not forget to thank the undergraduate students who worked on my project. Matthijs and Lidwien were in the lab when our overambitious aim was still to unravel the cell death machinery via cDNA phage display. They often picked up

techniques faster than I did myself and were of great help to my research. A few years ago Wladimir started working here in Wageningen at the Laboratory of Phytopathology studying the interaction between Rx1 and RanGAP2. Almost immediately it was clear that there were interesting between his research on RanGAP2 and our work on the localization of Rx1. I am glad we used this great opportunity to work together and share ideas and materials. Already our cooperation resulted in some interesting insights and I am convinced more will follow. In 2006 Kamila and I had the opportunity to visit the lab of Peter Moffett at the Boyce Thompson Institute in Ithaca. Peter's work on Rx1 has contributed enormously to the understanding of R protein functioning and has always been a great inspiration, so it was an honour to spend that week in Ithaca. Erin and I shared an office for many years. It was good to have somebody around to listen to new ideas, to discuss the weather, politics, and of course the latest research. We share a love for classical music, and could therefore agree to enliven the afternoons in the office with Bach's Mattheus Passion or other classical pieces. I owe a lot to my supervisors during this project, Arjen and Aska. Arjen for his original ideas and ambition. Aska for structuring my often chaotic planning and work, for the feedback and reassurance at the many moments that experiments did not work out as planned or did not give the results we hoped for. I would like to thank both for the almost complete freedom I was given to set out my own course in research, a freedom which I learned to use better over the years. Jaap's enthusiasm for my results, his eye for detail when going meticulously through manuscripts and our many discussions also meant a great deal to me. Lisette is the organizing centre of Nematology. Many thanks for patiently reminding me of all the administrative necessities which I would otherwise forget and also for helping me filling out forms and expense accounts for which I harbour an irrational anxiety. My fellow PhD students and other colleagues on the lab, I cannot possibly name you all, but thanks for the good atmosphere, the 'labuitjes', the Friday afternoon drinks, the traditional birdwatching excursions, and even the workdiscussions. Jan, Kristof, Pjotr, you are some of the most intelligent and ambitious scientists I know and I liked our discussions. Maybe even more because our fields of research are so far apart. It helps to keep a broad view of the world and not lock oneself into a just one narrow specialization when there is so much more interesting and complex stuff out there. And finally I would like to thank all the people who maybe did not contribute directly to this thesis, but certainly made my life outside the lab more enjoyable. There is more to life than science. Erik, Nienke, Erwin, Judith, Mark, Christien, Henk-Jan, Suzan, Joren, Willemy, Dieuwertje, Stijn, Simon, Jos, Matthijs, and others, I have good memories of all the travels, cycling trips, RISK evenings, wine tasting, movies, gardening adventures, etc. we had together. And last but certainly not least, I would like to thank my family and especially my parents, who stimulated my interest in nature and gave me the freedom to choose the study I liked most.

Publications

Journal articles

Shah, K. Tung, C.H., Chang, C.H., Slootweg, E.J., O'Loughlin, T., Breakefield, X.O., Weissleder, R. (2004). In vivo imaging of HIV protease activity in amplicon vector-transduced gliomas. *Cancer Research* **64**, 273-278.

Slootweg, E.J., Keller, H.J., Hink, M.A., Borst, J.W., Bakker, J., Schots, A. (2006). Fluorescent T7 display phages obtained by translational frameshift. *Nucleic Acids Research* **34**, e137.

Slootweg, E., Tameling, W., Roosien, J., Lukasik, E., Joosten, M., Takken, F., Bakker, J., Govere, A. (2009). An outlook on the localisation and structure-function relationships of R proteins in *Solanum*. *Potato Research* **52**, 229-235.

Manuscripts in preparation

Slootweg, E.J., Roosien, J., Spiridon, L.N., Petrescu, A.-J., van Schaik, C., Dees, R., Borst, J.W., Smant, G., Schots, A., Bakker, J., Govere, A. Nuclear localization of the NB-LRR Resistance protein Rx1 is dependent on coiled-coil domain, ATP-binding, SGT1 and Rar1, and its activation is triggered in the cytoplasm.

Slootweg, E.J., Spiridon, L., Roosien, J., Butterbach, P., Pomp, R., Smant, G., Petrescu, A., Bakker, J., Govere, A. A docking model for the NB-ARC and LRR domains of the CC-NB-LRR Resistance protein Gpa2.

Slootweg, E.J., Koropacka, K., Roosien, J., Dees, R., van Schaik, C., Pomp, R., Bouwman, L., Schots, A., Smant, G., Bakker, J., Govere, A. Domain exchange between Rx1 and Gpa2 reveals flexibility of CC-NB-LRR genes to switch between virus and nematode resistance.

Congress abstracts

Slootweg, E., Keller, H., van Sprundel, L., Orzaez, D., Hink, M., Bakker, J., Schots, A. (2003). Protein-protein interactions during programmed cell death in plants. In the Keystone Symposium "Plant Biology: Functions and Control of Cell Death". Snowbird, Utah, USA. 121, p. 51.

van Bers, N.E.M., Overmars, H.A., Slootweg, E.J., Qin Ling, Govere, A., Bakker, J., Smant, G. (2007). The nematode effector peptide dgl-3 accumulates in the nucleolus of tobacco cells. In NEMAGENICS, Exploiting genomics to understand plant-nematode interactions. COST 872 Workshop, La Colle-sur-Loupe, France.

Koropacka, K., Sloatweg, E., Roosien, J., Dees, R., Sacco, M.A., Moffett, P., Smant, G., Govere, A. (2007). Structural and functional analysis of *Gpa2*-mediated resistance against the potato cyst nematode *Globodera pallida*. In the 13th International congress on molecular plant-microbe interactions, Sorrento, Italy, CS 6-5, p. 67.

Butterbach, P., Sloatweg, E., Koropacka, K., Spiridon, L., Dees, R., Roosien, J., Bakker, E., Arens, M., Petrescu, A., Smant, G., Bakker, J., Govere, A. (2007). Functional constraints and evolutionary dynamics of the *Rx1/Gpa2* cluster in potato. In the 13th International congress on molecular plant-microbe interactions, Sorrento, Italy, CS 15-2, p. 84 & Proceedings. MPMI XIII.

Sloatweg, E., Roosien, J., Hink, M., Schots, A., Govere, A., Bakker, J. (2007). Protein regions determining the nucleo-cytoplasmic distribution of the R protein Rx1. In the 13th International congress on molecular plant-microbe interactions, Sorrento, Italy, PS 1-76, p. 120.

Sloatweg, E., Roosien, J., Borst, J.W., Schots, A., Govere, A., Bakker, J. (2008). Nucleo-cytoplasmic distribution of the R protein Rx1 and its subdomains. In the Keystone Symposium "Plant Innate Immunity", Keystone, Colorado, USA. 306, p. 133.

Sloatweg, E., Spiridon, L., Roosien, J., Pomp, R., Smant, G., Petrescu, A., Bakker, J., Govere, A. (2009). Sequence exchanges and targeted mutagenesis based on structural modeling of Gpa2 and Rx1 provide novel insights in NB-LRR protein functioning. In The 14th International Congress on Molecular Plant-Microbe Interactions, Quebec, Canada. **PS 8-445.**

Sloatweg, E., Roosien, J., Spiridon, L., Borst, J.W., Petrescu, A., Schots, A., Bakker, J., Govere, A. (2009). Nucleo-cytoplasmic distribution of the R protein Rx1 and its domains. In the 14th International Congress on Molecular Plant-Microbe Interactions, Quebec, Canada. **PS 5-211.**

Tameling, W.I., Nooijen, C., Sloatweg, E.J., Govere, A., Ludwig, N.R., Joosten, M.H. (2009). Forced changes in nucleo-cytoplasmic partitioning of the NB-LRR Rx dictate initiation of defence signalling. In the 14th International Congress on Molecular Plant-Microbe Interactions, Quebec, Canada. **PS 7-430.**

Curriculum vitae

Erik Jeroen Slootweg was born in Breda, the Netherlands, on January 4, 1978. He graduated from grammar school at the Stedelijk Gymnasium Breda in 1996. He studied Biology (BSc and Msc) at Wageningen University, where he specialized in plant biology. He conducted the research for his undergraduate theses at the Laboratory of Nematology (Functional analysis of the homologous resistance genes Rx1 and Gpa2 by means of deletion studies and domain swapping), and in the group of Sacco de Vries at the Laboratory of Molecular Biology (Characterization of an AtSERK2, a receptor-like kinase protein, subcellular localization, phosphorylation studies and expression patterns). The final stage of his study was a 5 month internship in Xandra Breakefields group at the Laboratory of Molecular Neurogenetics, Harvard Medical School & Massachusetts General Hospital (Charlestown, MA, USA) under supervision of Khalid Shah (Protease regulation of the tumour specific apoptosis-inducing protein TRAIL combined with near-infrared fluorescence imaging in a Herpes simplex virus-based gene therapy system). In March 2002 he graduated with distinction and started his PhD research on the EU project "Insight Inside: Signal transducing protein machines revealed by subcellular single molecule detection" in the group of Arjen Schots at the Laboratory for Molecular Recognition and Antibody technology (MLA) of the Wageningen University. From 2006 onwards he worked in the group of Aska Goverse at the Laboratory of Nematology of the Wageningen University, first in an extension to his PhD program and now as a post-doctoral associate within the EU project BIO-EXPLOIT.

**Education Statement of the Graduate School
Experimental Plant Sciences**



Issued to: Erik Jeroen Slootweg
Date: 23 October 2009
Group: Laboratory of Nematology, Wageningen University

1) Start-up phase ► First presentation of your project Plant programmed cell death ► Writing or rewriting a project proposal ► Writing a review or book chapter ► MSc courses ► Laboratory use of isotopes	<p align="right"><i>date</i></p> <p align="center">20 Mar 2002</p>	
<i>Subtotal Start-up Phase</i>		1.5 credits*
2) Scientific Exposure ► EPS PhD student days PhD student day, Radboud University Nijmegen PhD student day, Wageningen University ► EPS theme symposia EPS theme 2 symposium 'Interactions between plants and biotic agents', Leiden University EPS theme 2 symposium 'Interactions between plants and biotic agents', University of Amsterdam ► NWO Lunteren days and other National Platforms NWO-ALW Lunteren days (2 days per year) ► Seminars (series), workshops and symposia EPS Workshop 'Programmed Cell Death', Leiden NWO-ALW System Biology meeting, Egmond aan Zee BIO-Exploit scientific meetings, Ede EPS Flying seminars EPS Frontiers in plant development seminars ► Seminar plus ► International symposia and congresses CHI symposium "Phage Display: the chemistry set for proteins" IV, Cambridge, MA, US Keystone symposium "Functions and control of Cell Death, Snowbird, UT, US CHI symposium "Phage Display: the chemistry set for proteins" VI, Cambridge, MA, US XIII International MPMI congress, Sorrento, Italy Keystone symposium "Plant Innate Immunity", Keystone, CO, US 17 th triennial conference of the EAPR, Brasov, Romania XIV International MPMI congress, Quebec, Canada ► Presentations Presentation at Keystone symposium on Plant biology: "Functions and control of cell death", Snowbird, UT, US Presentation at the conference of the EAPR, Brasov, Romania Presentation at the PhD Summerschool Plant-Pathogen Interactions, Wageningen Presentation at the NWO/ALW meeting in Lunteren 2009 Presentation at the BIO-EXPLOIT summit 2007, Ede Poster presentations at MPMI 2007, MPMI 2009 (2x), Keystone 2003, Keystone 2008 (total 5 posters). ► IAB interview ► Excursions	<p align="right"><i>date</i></p> <p align="center">03 Jun 2005 19 Sep 2006 23 Jun 2005 02 Feb 2007 2002-2009 Feb 2002 Mar 2003 2006-2009 2003-2009 2003-2009 Apr 2002 Apr 2003 Apr 2004 Jul 2007 Feb 2008 Jul 2008 Jul 2009 Apr 2003 Jul 2008 20 Jun 2009 07Apr 2009 Oct 2007 3003-2009 03 Jun 2004</p>	
<i>Subtotal Scientific Exposure</i>		23.5 credits*
3) In-Depth Studies ► EPS courses or other PhD courses FEBS Advanced Course Microspectroscopy: Monitoring interactions and reactions in living cells PhD Summerschool Plant-Pathogen Interactions, Wageningen ► Journal club Literature discussion group 'Plant Resistance' (Nematology department) ► Individual research training	<p align="right"><i>date</i></p> <p align="center">29 Sep-05 Oct 2002 18-20 Jun 2008 2006-2008</p>	
<i>Subtotal In-Depth Studies</i>		4.8 credits*
4) Personal development ► Skill training courses OWU course 'supervision of PGO groups IOW' (3 days) OWU Course 'hoorcollege geven' (3 days) ► Organisation of PhD students day, course or conference ► Membership of Board, Committee or PhD council	<p align="right"><i>date</i></p> <p align="center">Jul 2002 Jan 2004</p>	
<i>Subtotal Personal Development</i>		1.8 credits*
TOTAL NUMBER OF CREDIT POINTS*		31,6

Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS which comprises of a minimum total of 30 credits

* A credit represents a normative study load of 28 hours of study

Financial support

The research presented in this thesis was performed at the Laboratory of Nematology of the Wageningen University with financial support from the European Union 6th framework project BIOEXPLOIT (FOOD-CT-2005-513959) and at the Laboratory for Molecular recognition and Antibody technology (LMA) of the Wageningen University with financial support from the European Union 5th framework project Insight Inside (QLRT-2000-01428).

Cover image

The cover image depicts *Nicotiana benthamiana* leaf epidermis cells. The outer cell layer of the leaf is the first line of defence the plant has against pathogens. Their strange puzzle-like appearance has a functional origin. It allows the cells to have a larger connecting surface, thereby increasing the strength by which the cells are attached to each other.