

# An Integrated Approach to Comprehend MYMIV-Susceptibility of Blackgram Cv. T9 Possessing Allele of *CYR1*, the Cognate R-Gene

Anju Patel<sup>1\*</sup>, Soumitra Maiti<sup>1\*</sup>, Santosh Kumar<sup>1</sup>, Sayak Ganguli<sup>2</sup>, Amita Pal<sup>1#</sup>

<sup>1</sup>Division of Plant Biology, Bose Institute, Kolkata, India

<sup>2</sup>DBT Centre for Bioinformatics, Presidency University, Kolkata, India

Email: anjupatel2009@gmail.com, soumitragalaxy@gmail.com, skg.research@gmail.com, ganguli.sayak@gmail.com, #amita\_pal@yahoo.com, #amita@jcbse.ac.in

Received 7 October 2015; accepted 14 February 2016; published 17 February 2016

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution-NonCommercial International License (CC BY-NC).

<http://creativecommons.org/licenses/by-nc/4.0/>



Open Access

## Abstract

Blackgram, an important legume crop, faces the constraint of Mungbean yellow mosaic India virus (MYMIV)-stress resulting in severe crop penalty. MYMIV-resistant plants exhibit incompatible response via a cognate *CYR1* gene-mediated interaction with virus effector molecule. In this study, we searched for the susceptible allele of the “R” gene in Cv. T9. Southern hybridization study confirmed presence of an allele in Cv. T9. However, transcripts of the *CYR1* could not be detected either by RT-PCR or by Northern hybridization in Cv. T9 and also in other susceptible blackgram line. The allele was isolated, sequenced and referred as *cyr1*. *In silico* study revealed that *cyr1* also encodes a CC-NBS-LRR protein like *CYR1*. However the CC domain of *cyr1* is truncated by 128 amino acid residues indicating functional impairment with respect to the signal transduction after pathogen invasion. Comparative 3D structural modeling, hydrogen bonding and Van der Waals interaction studies revealed differences between *CYR1* and *cyr1*. Lys519 and Thr490 present in the largest pockets of the *CYR1* are the key interacting hotspots between *CYR1* and MYMIV coat protein (CP). The weak Van der Waals interactions and intermolecular hydrogen bonding between *cyr1* and CP confers less stability to the molecular recognition complex, unlike *CYR1*. Thus, the present investigation revealed Cv. T9 shows compatible interaction with MYMIV due to the truncation in the *cyr1* sequence and consequent structural difference in the N-terminal of CC-domain.

\*Both these authors are equal contributors.

#Corresponding author.

## Keywords

**Blackgram, CC-NBS-LRR, *CYR1*, *cyr1*, Incompatible Interaction, Mungbean Yellow Mosaic India Virus**

## 1. Introduction

Yellow mosaic disease of blackgram (*Vigna mungo*) is caused by Mungbean yellow mosaic India virus (MYMIV). Successful disease establishment resulted in irregular, chlorotic, yellow patches on the leaves, the characteristic phenotype of MYMIV-infected susceptible plants. The yield penalty reaches to the extent of cent percent when pathogenesis affects MYMIV-susceptible plants at the early stage of development. Interactions between blackgram and MYMIV can lead to establishment of the disease, *i.e.* compatible response; on the contrary, MYMIV-resistant plants exhibit complete resistance, the phenomenon referred to as incompatibility. In one of our previous studies we have isolated a candidate disease resistance (R) gene, *CYR1* that co-segregates with MYMIV-resistant populations of *V. mungo* [1]. *CYR1* that coded *in silico* translated protein sequence comprised of 1176 amino acids with coiled coil (CC) structure at the N-terminus, central nucleotide binding site (NBS) and C-terminal leucine-rich repeats (LRRs) that belongs to non-TIR (CC)-NBS-LRR subfamily of plant R genes. *CYR1* transcripts were unambiguously expressed during incompatible plant virus interactions. Additional molecular and *in silico* studies suggest involvement of *CYR1* in recognizing MYMIV-effector molecule thus contributing incompatible interaction with MYMIV. We have also introgressed this cognate R gene in the MYMIV-susceptible cultivar T9 through molecular breeding and raised number of agronomically useful recombinant inbred lines establishing the biological function of *CYR1* [2]. These class of genes encodes CC-NBS-LRR proteins that recognize specific pathogen-derived products and initiate a resistance response that often includes a type of cell death known as the hypersensitive response (HR).

Under this backdrop we questioned the role of allele (*cyr1*) of *CYR1*, if any, present in the susceptible Cv. T9. Hence, at first, we detected the presence of an allele of *CYR1* in Cv. T9 through Southern hybridization. However, we could neither detect any transcripts of *CYR1* by RT-PCR nor by Northern hybridization techniques in the MYMIV challenged Cv. T9 or in segregating susceptible lines (raised by crossing T9 with the MYMIV-resistant line). Subsequently, we have isolated the allele through PCR amplification followed by genome walking and sequencing. It was revealed that *cyr1* possess truncated N-terminal CC-region. It was suggested by Feys *et al.* [3] that N-terminus of NBS-LRR proteins influences downstream signaling components after recognizing pathogen effector proteins. The barley (*Hordeum vulgare*) Mla proteins bind to WRKY transcription factors through their CC domains [4], although it was not clear that whether such interaction was involved in recognition or signaling. Similarly, N-terminal TIR domain of RPP1A, one of the R proteins of *Arabidopsis thaliana* plays a crucial role in the cell death signaling pathway [5]. Later, Swiderski *et al.* [6] have confirmed through a comprehensive study that the TIR domain is indeed an effector domain for cell death signaling, like CC domain.

Nevertheless, the genomic evidences of host-pathogen interaction could not unambiguously explain the inability of expression of *cyr1* transcripts even after challenging with the virus, which led us to generate 3D models using comparative modelling of the *CYR1* protein from the resistant and *cyr1* from susceptible Cv. T9 along with the MYMIV-coat protein (CP), which is the most potent pathogen effector molecule. The models of the *CYR1* and *cyr1* proteins were then aligned and an alignment Z Score of 0.84 was obtained with an RMSD of 3.6 (Å); which indicated significant structural similarities. These were then docked separately with the MYMIV-CP and analyzed for intermolecular interactions which revealed strong associations involving hydrogen bonded and Van der Waals contact mediated interactions for the *CYR1* and MYMIV-CP; and only non-bonded weak Van der Waals interaction between *cyr1* and MYMIV-CP.

All these above evidences put together point towards the inability of *cyr1* allele of T9 cultivar to confer resistance against MYMIV and hence showed compatible interaction.

## 2. Materials and Methods

### 2.1. Plant Materials

The MYMIV-susceptible *Vigna mungo*, Cv. T9 and MYMIV-resistant line VM1 [7] were grown in the insectary

room with suitable growth conditions at the Madhyamgram Experimental Farm (MEF) of Bose Institute, Kolkata, India. In addition, one susceptible inbred line (B20) and two resistant inbred lines (VMR B33 and B38) [2] were used as experimental materials for studying transcript expression of targeted alleles of *CYRI* amongst virus challenged *V. mungo* plants.

## 2.2. Artificial, Forced Inoculation of *V. mungo* with MYMIV

Experimental plants were challenged with the viruliferous whiteflies, vector of MYMIV, using the forced feeding method as described by Kundu and Pal [2]. Briefly, white fly populations (*B. tabaci*) were reared on susceptible blackgram plants, grown in insect proof cages and maintained at 27°C in the insectary facility of the MEF. About 25 - 30 adult white flies were captured in transparent glass trappers and allowed for an acquisition access period of 24 h on symptomatic leaves of naturally infected blackgram plants. Subsequently, the trappers with viruliferous whiteflies were attached to the leaves of 3-week-old healthy plants of selected experimental lines and allowed to feed for 24 h inoculation access period, thereby transmitting the virus particles from the viruliferous flies to the plant. Likewise another set of experimental plants were forcefully fed by the starved, non-viruliferous whiteflies and referred to as mock inoculated control. Accumulation of MYMIV particles in the host plants after challenging with the virus was determined by qPCR as previously reported by Kundu *et al.* [8].

## 2.3. Isolation of Genomic DNA

DNA was isolated from the seedlings of *V. mungo* Cv. T9 and MYMIV-resistant line VM1 using modified CTAB method of DNA isolation following Basak *et al.* [7]. Isolated DNA was quantified in a NanoDrop™ 1000 Spectrophotometer (Thermo Scientific, USA) and diluted to a concentration of 0.1 µg/µl. The size and purity was checked by loading 1 µl of sample DNA and 1 µl of control genomic DNA (0.1 µg/µl; supplied with Genome Walking kit, Clontech) on a 0.6% agarose/EtBr gel along with 1 kb Plus DNA ladder (Fermentas, France). The purity of genomic DNA was checked by *EcoRV* digestion.

## 2.4. Southern Hybridization

Twenty-five µg genomic DNA was isolated separately from five experimental plants as stated above. These were subjected to restriction digestion with *Bam*HI, *Eco*RI, *Hind*III and *Sma*I, separated on agarose gel and blotted onto a nylon membrane. The objective of these hybridization experiments was to check the presence and copy number of *CYRI* and its allele within the whole genomes of VM1 and T9 Cv. respectively.

Following steps were performed sequentially to conduct the hybridization experiments. A PCR fragment amplified from VM1 genomic DNA, using 5'-GGGTGGATTGGGTAAGACCAC-3' Forward and 5'-TCTGAAAAGTGTGTGGCGCT-3' Reverse primer combinations. Amplified product was 1235bp length and labeled with  $\alpha$ P<sup>32</sup>dCTP using Prime-a-gene labeling system kit (Promega, USA) as per manufacturer's instructions. The membrane was incubated at 68°C in pre-hybridization solution for 3 hr under constant rotation in a hybridization oven (Stuart Scientific, UK). The radio-labeled probe was denatured and added to the membrane and incubated at 68°C. The membrane was washed on the following day and exposed to X-ray film.

## 2.5. Isolation of Total RNA

Leaf samples from control and MYMIV-infected plants were collected at 6, 9, 12, 18, 24 and 36 hr after forced inoculation. Experiments were performed three times on different dates to obtain three independent biological replications. Total RNA was isolated from mock inoculated and infected leaf tissues by using the Trizol reagent (Invitrogen, Carlsbad, CA) treated with DNase-I (Sigma-Aldrich, USA) to eliminate traces of genomic DNA and purified using the RNeasy Plant Mini Kit (Qiagen, USA) following manufacturer's instruction. Integrity of the isolated RNA samples were assessed by agarose gel electrophoresis and purity and quantity of individual samples were determined spectrophotometrically (NanoDrop™ 1000 Spectrophotometer, Thermo Scientific, USA).

## 2.6. Semi-Quantitative RT-PCR Analyses

Expression of *CYRI* and its allele, *cyr1* were analyzed by RT-PCR method from total RNA of control and MYMIV-infected MYMIV-resistant and -susceptible lines, respectively at 6, 9, 12, 18, 24 and 36 hr. Titan One

Tube RT-PCR System (Boehringer Mannheim, USA) was used to synthesize the first strand of cDNA using AMV reverse transcriptase at 50°C for 30 min followed by the pre-amplification of 10 cycles (denaturation at 94°C for 30 sec, primer annealing at 60°C for 30 sec, elongation at 68°C for 1 min), amplification of 25 cycles with a blend of Taq DNA polymerase and Pwo DNA polymerase (denaturation at 94°C for 30 sec, primer annealing at 60°C for 30 sec and primer extension at 68°C for 1 min with an increment of 5 sec in each cycle) and a final extension at 68°C for 7 min. At the same time, actin amplification was done with the primers VActin-F (5'-ATGACTCAGATCATGTTTGAG-3') and VActin-R (5'-AGCCTTCGCAATCCACATCTG-3') were used in RT-PCR, reaction. The amplified products were analyzed using 1.5% agarose gel electrophoresis and stained with EtBr.

## 2.7. RNA Blot Hybridization

Total RNA (20 µg) from each sample was separated on a 1.0% agarose-formaldehyde gel and transferred to Hybond-N<sup>+</sup> Nylon membrane (Amersham, Pharmacia, UK). *CYR1* gene specific dCTP-[ $\alpha$ -<sup>32</sup>P]-labelled probe was generated and RNA blot hybridization was performed using the method described above for Southern hybridization.

## 2.8. Isolation of *cyr1* Allele by PCR-Based Genome Walking

T9 *cyr1* sequence has been elucidated using genomic DNA from Cv. T9 as template, using Genome Walking kit (Clontech, USA) following manufacturer's instruction to obtain full length sequence of the allele. The digested and purified DNA was ligated to Genome Walker Adaptor with T4 DNA ligase at 16°C overnight in a thermal cycler. The reaction was stopped by adding TE buffer (pH 7.5) and incubated at 70°C for 5 min, then vortexed at slow speed for 10 sec. The adapter was ligated at both ends of the entire digested product to create Genome Walker libraries. To screen the libraries, primary PCR was set up with outer adapter primer (AP1) and outer gene specific primer (GSP1). Secondary PCR was then set up from primary PCR product with nested adapter primer (AP2) and nested gene specific primer (GSP2). Major PCR product was sequenced in ABI automated DNA sequencer (ABI Prism 3100) and subsequently used for further genome walking using different primers (Table 1), which were designed from conserved regions of *CYR1* (Table 1) and different primer combinations were used.

**Table 1.** Primers used in genome walking.

| Primer code           | Primer sequence                              | Annealing temperature (°C) |
|-----------------------|--|----------------------------|
| GW1CF                 | 5'-CAACCACGGAAAAGCGGTTTGTAAATGCATG-3'        | 74                         |
| GW1CFN                | 5'-TCTGTGGGGACATCTGTTTCAGGTTG-3'             | 76                         |
| GW2CF                 | 5'-TGTGGGGACATCTGTTTCAGGTTG-3'               | 70                         |
| GW2CFN                | 5'-AATACATAACGAAAACGACCCGTC-3'               | 68                         |
| GW3CF                 | 5'-AAGGCTGGAGGAATTGCCCTCAAAT-3'              | 72                         |
| GW3CFN                | 5'-ACTGACCAAATTGCGTTGCCTAGAA-3'              | 70                         |
| GW4CF                 | 5'-TCTTTTCCACGTCTTCAACATCTGTCT-3'            | 74                         |
| GW4CFN                | 5'-TATGGATCACTGTCCCGAGCTGAAAG-3'             | 76                         |
| GW5CF                 | 5'-TAAAACATGCGTCGCTTTCAAGTT-3'               | 64                         |
| GW5CFN                | 5'-TTAAGAGAGTCATTGGATGCAAAC-3'               | 64                         |
| CYRT-Fd               | 5'-TAGCAAAGCATGGGTCTGTG-3'                   | 55                         |
| CYRT-Rv               | 5'-GAGTTTGTACAGCTTCCCATTCGTCTC-3'            | 54                         |
| GW4CR                 | 5'-GAGTTCAAGCGATGATGTGTCCATGT-3'             | 71                         |
| <i>CYR1</i> -F-Bam-Tn | 5'-TTTTTTTGGATCCATGGCAGCAGCTTGTC-3'          | 78                         |
| <i>CYR1</i> -R-Kpn-Tn | 5'-TTTTTTTGGTACCCATTGAAAAATTGATTATTA GGTA-3' | 69                         |
| S5RcCyR3              | 5'-TGAAGTCTTCTGTTCTGCAT-3'                   | 55                         |

## 2.9. Structural Analysis of Targeted Sequence Using Tools of Bioinformatics

The nucleotide sequence thus obtained was translated using ExPASy translate tool and referred to as *cyr1*, and conserved domains were analysed using BLASTP and CDS (Conserved Domain Search) tool of NCBI, respectively. To compare *cyr1* with *CYR1* multiple sequence alignment was done in ClustalW server.

## 2.10. Secondary Structural Analyses, Comparative Modelling and Simulation

Secondary structural analysis was performed using PSIPRED which uses a simple and accurate secondary structure prediction method, incorporating two feed-forward neural networks which perform an analysis on output obtained from PSI-BLAST (Position Specific Iterated-BLAST). Once the secondary structural features were identified, the proteins were modelled using Modeller 1.12 using multiple comparative modelling. The multiple templates used for the comparative modelling procedure were 4UO8A, 3SFZA, 4TZHA, 4MN8A and 4F3GA for the *CYR1* protein and 1Z6TA, 4TZHA, 4FCGA and 4LXRA for the *cyr1* protein respectively. Generated structures were subsequently analyzed for their geometries using RAMPAGE server. Following the analyses of geometry the structures were subjected to simulation using the consensus view of protein near-native dynamics obtained from 10-nanosecond MD simulations of all-atom, explicit water, for all protein metafolds using the four most popular force-fields (AMBERff99sb, Consistent Force Field, CHARMM and GROMOS). The model obtained following simulation using AMBERff99sb [9] was found to be the most suitable for further calculations.

## 2.11. 3D Structure Analyses

Following this the CASTp server was used for the analyses of 3D structural motifs present in the structures. CASTp uses a combination of Delaunay Triangulation scoring scheme based on the basic voronoi diagram architecture which enables the proper detection of concavities and depressions in protein 3D structures along with the calculations of the area and volume of the predicted pockets.

## 2.12. Fast Fourier Transformation Based Docking

For the study of protein-protein interaction rigid body docking calculations were used in a background Fast Fourier transform (FFT) approach. Here for each molecule encoding of both surface shape and electrostatic charge and potential distributions was modeled using 3D expansions of real orthogonal spherical polar basis functions; which enables the representation of each property by a coefficient vector. The Van der Waals steric density model and the classical electrostatic theory was used for surface calculations and representations. The overlap of parametric functions were represented as expressions to generate the overall docking score as a function of the six degrees of freedom in a rigid body docking search. With suitable scaling factors, this docking score was then interpreted as gradually minimized interaction energy. The docking calculations were performed using HEX and were validated with PATCHDOCK server and filtered through Firedock.

## 2.13. Studies on Interactions Using DIMPLOT

The DIMPLOT program was used to plot the interactions across the protein-protein interface; *i.e.* the interface between the *CYR1* and *MYMIV-CP* as well as the interactions between the T9 cultivar *cyr1* and *MYMIV-CP*. The interactions plotted were hydrogen bonds and non-bonded contacts through Van der Waal's Interactions. They were extracted from the usual HBPLUS output by the program DIMER; and the plot was generated by LIGPLOT.

## 2.14. *In Silico* Mutagenesis and Calculation of Interaction Hotspots

*In silico* mutagenesis was performed using SPDB viewer and the proteins were then simulated using GROMACS. Separate models were generated for each interacting residue. This resulted in 3 different models of the *CYR1* protein. However, in case of T9 *CYR1* protein no residues were mutated since hydrogen-bonding pattern were not identified. Thus 3 distinct docking calculations were reperformed and hotspots were identified (Amino acid residue—the change which causes  $\Delta\Delta G > 2.5$  kcal/mol change in free energy of docking is accepted to be a hotspot) [10].

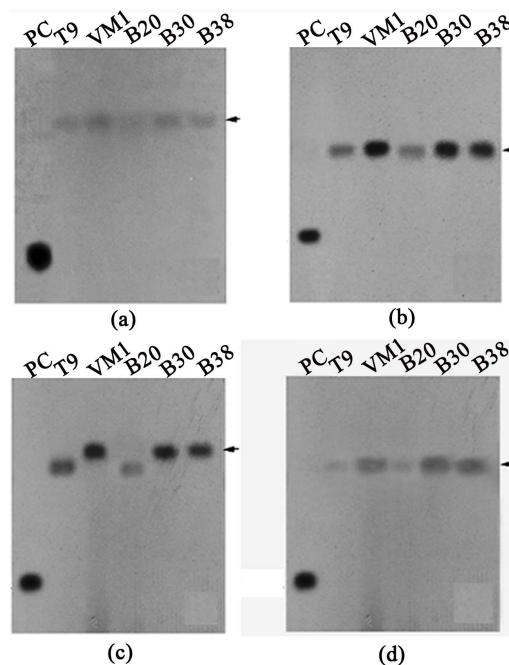
A single residue N112 was found to be present in the unaligned portion of the CYR1 protein which was involved in Van der Waal's contact with the cognate MYMIV-CP protein. The corresponding segment in the N terminal domain of the *cyr1* protein was truncated. To test the importance of this particular residue, it was mutated to alanine, simulated using the Amberff99sb force field and redocked with the MYMIV protein.

### 3. Results and Discussion

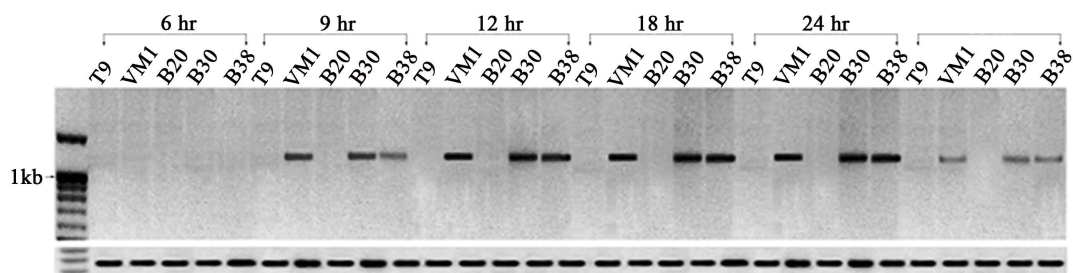
Southern hybridization with labeled "CYR1<sub>1235</sub>" fragment of the digested genomic DNA with all the four different restriction enzymes (*Bam*HI, *Eco*RI, *Hind*III and *Sma*I) revealed presence of single copy *CYR1* gene within the genomes of MYMIV-resistant lines of *V. mungo*, VM1, B30 and B38 (Figures 1(a)-(d)). Faint signals were also observed in the genomes of susceptible cultivar T9 and susceptible line B20. However there were differences in the migration of signals between MYMIV-resistant and -susceptible plants, especially in *Hind*III digested genomes of susceptible lines, which moved faster in the gel than that of the resistant lines. Thus Southern hybridization revealed presence of an allele of *CYR1* in the susceptible Cv. T9 and in the inbred susceptible line B20, which is referred to as *cyr1* allele in the following text.

In order to investigate expression profiles of transcripts of *CYR1* after MYMIV infection, reverse transcription PCR was conducted. *CYR1* transcripts were expressed in the resistant plants (VM1, B30 and B38) at 9 hpi after virus infection (Figure 2). Highest expression level was observed at 18 and 24 hpi and declined subsequently at 36 hpi. Transcript expression was undetectable in the susceptible plants, Cv. T9 and B20 (Figure 2). None of the mock inoculated and wounded plants showed *CYR1* transcript expression (data not shown).

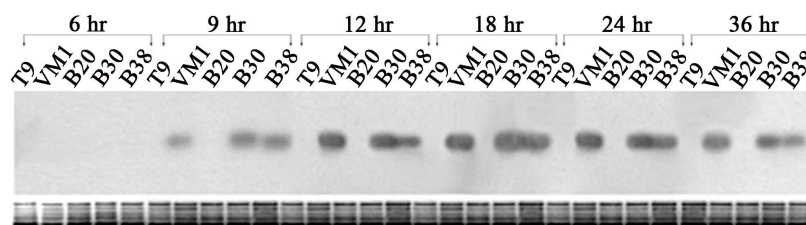
Subsequently RNA blot hybridization was performed to confirm *CYR1* transcript expression at different time points in resistant and susceptible lines of *V. mungo* after challenging with the virus. In none of these cases transcripts of *CYR1* could be detected from the MYMIV-susceptible T9 cultivar and in the susceptible inbred line, B20. Whereas, the expression pattern of *CYR1* is completely in agreement with the RT-PCR results. As expected, the house keeping actin gene was present in all the samples irrespective of their MYMIV-reaction status (Figure 3).



**Figure 1.** Southern hybridization analysis of genomic DNA of *Vigna mungo*, digested with *Bam*HI (a), *Eco*RI (b), *Hind*III (c) and *Sma*I (d), probed with  $\alpha^{32}$ -labeled *CYR1* fragment of candidate *R* genes. Lane number in each blot represents: 1) positive control (PC); 2) T9, MYMIV susceptible cultivar; 3) VM1, MYMIV resistant line; 4) B20, a susceptible inbred line; 5) and 6) B30 and B38 resistant inbred lines. The solid arrows indicate the hybridization signal of the target gene.



**Figure 2.** *CYR1* transcript accumulation at different hours after MYMIV infection in resistant and susceptible *V. mungo* plants (upper panel). T9 is the susceptible parent and VM1 is the resistant parent, whereas B20 is a susceptible inbred line and B30 and B38 are the resistant inbred lines of *V. mungo*. Actin gene amplification by RT-PCR has been shown (lower panel) as positive control from same RNA samples.



**Figure 3.** *CYR1* transcript expression by Northern hybridization at different hours after MYMIV infection in resistant and susceptible *V. mungo* plants. Twenty microgram aliquots of total RNA extracted from leaf tip tissues of experimental plants at 0, 3, 6, 9, 12, 18 and 24 hpi were used for northern assay. The blot was hybridized with  $\alpha^{32}$ -labeled *CYR1* marker fragment. Equal loading of RNA is shown in ethidium bromide-stained RNA gel (lower panel).

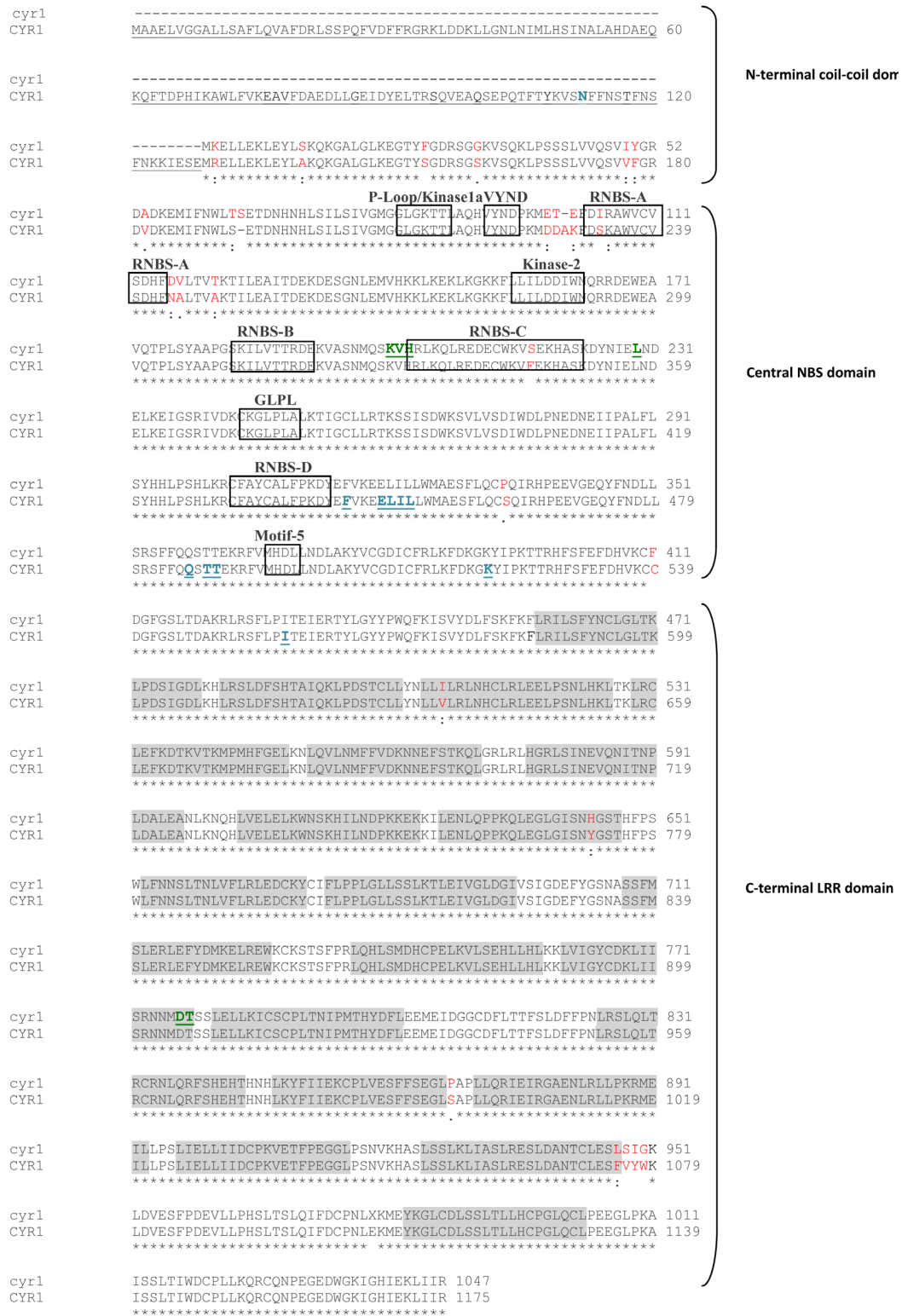
In a previous study we had shown that the candidate resistant gene *CYR1* is intron less [1], hence the allele *cyr1* was isolated from the genomic DNA of Cv. T9 by genome walking and submitted to NCBI (GenBank ID: KR350634).

The predicted amino acid sequence (5'-3' Translational Frame 1) of *cyr1* also shows homology with NBS and LRR domains like that of *CYR1* (Figure 4). Although *cyr1* has characteristic conserved motifs of CC-NBS-LRR class of disease resistance (R) protein, yet definite mismatches are noted within the sequence when compared with amino acid sequence of *CYR1*. A stretch of 128 amino acids is absent in N-terminal CC domain (Figure 4). Such truncated N-terminal CC-region indicates functional impairment with respect to the signal transduction presumably even after recognizing the effector protein of the virus. Role of N-terminal CC domain of R protein has been implicated in plant pathogen interaction by several researchers. Moffett *et al.* [11] have shown activation of R proteins in presence of coat protein of *Potato virus X* (PVX-CP), which they presumed due to the induction of a conformational change in the CC-type R proteins of potato Rx. They have further demonstrated that the isolated CC domain complemented a copy of Rx lacking NBS-LRR. CC domain complements NBS-LRR leading to pathogen-dependant HR during intramolecular interactions of Rx protein of potato [11] [12] and BS2 protein of *Nicotiana benthamiana* [13]. Similar inter-domain interactions appear to take place in TIR-NBS-LRR protein as well [14].

While, in RNBS-A motif, serine of *CYR1* is replaced by isoleucine in *cyr1*; and in RNBS-C phenylalanine residue is replaced by serine (Figure 4). In the C terminal of LRR domain valine, tyrosine and serine is replaced by isoleucine, histidine and phenylalanine, respectively. Also a short stretch of amino acid residues, FVYW, is replaced by LSIQ.

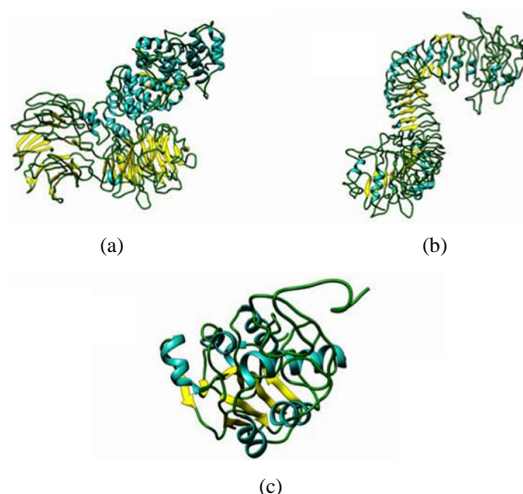
3-D structures of *CYR1*, *cyr1* and MYMIV-CP generated using comparative modeling are depicted in Figure 5. Distinct differences in the  $\beta$ -sheet conformations between *CYR1* and *cyr1* are evident. This is perhaps due to the truncated CC region of *cyr1*. Although structural alignment between these two plant proteins has conformational similarity (Figure 6).

Molecular recognition is one of the major facets behind the key biological processes that occur in an organism. The molecular switch that operates in plants to recognize pathogen invasion depends on the kind of pathogen [15]. In case of virus pathogens, host's NBS-LRR genes either induce HR or complete resistance as phenotypic

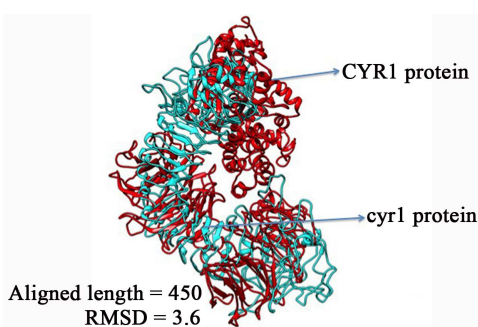


**Figure 4.** Alignment of of CYR1 predictive protein from the MYMIV-resistant genotype of *Vigna mungo* with *cyr1* from MYMIV-susceptible Cv. T9. The underlined region at the start of the alignment indicates the truncated portion of the N-terminal coiled coil domain of the *cyr1* sequence, the framed regions indicate the conserved motifs of central NBS domain and LRR repeats at the C-terminal region are shaded. Interacting residues of the *cyr1* protein are marked in green and those of the CYR1 protein are marked in blue.





**Figure 5.** 3-D structures of the interacting partners generated using comparative modelling. CYR1 protein (a); *cyr1* protein (b) and MYMIV-coat protein (c); Colour schemes according to secondary structural conformations: Yellow:  $\beta$ -sheet, Sea green:  $\alpha$ -helix, forest green: coils.

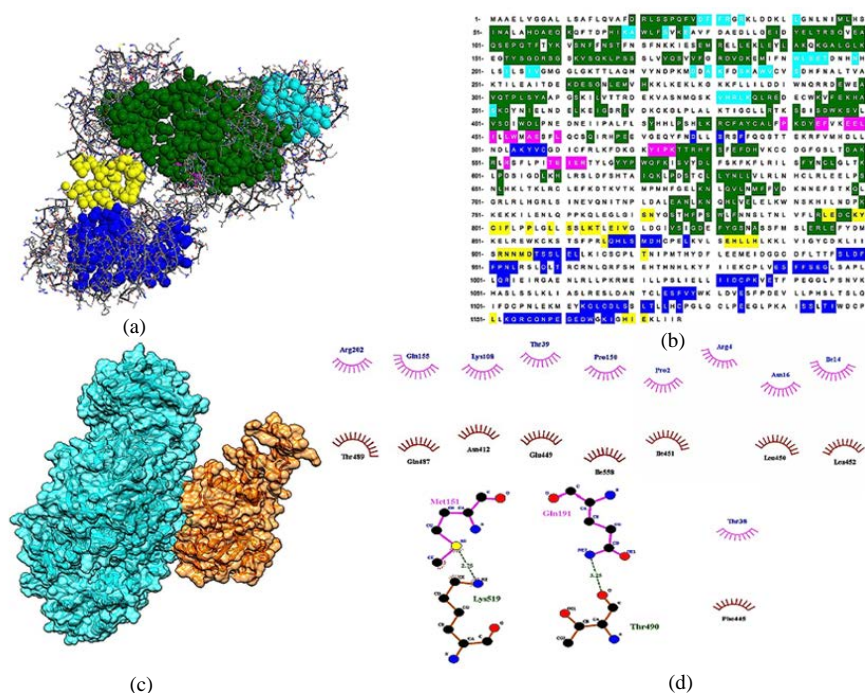


**Figure 6.** Structure alignment between CYR1 protein of MYMIV-resistant and *cyr1* of MYMIV-susceptible cultivar showing similar conformations.

response against virus infection [15]. The HSP90 (heat-shock protein 90) proteins are involved in the protein folding and modulate innate immune responsive proteins in plants [16]–[18]. In a previous study we have shown that in MYMIV-resistant genotype of blackgram HSP 90 was over-expressed at early hours of viral infection than in Cv. T9 [8]. This further indicates the possibility of functional inability of *cyr1* in recognizing MYMIV-CP, perhaps due to folding impairment.

Most molecular recognition processes are based on either electrostatic or geometric complementarities. Cellular proteins are involved in their interactions with other cellular proteins or small molecular agonists, antagonist cofactors or even drugs. For most of such interactions, there are either one or two or multiple (allosteric) binding sites, which represent a precise 3D conformation that enables the protein to get a good grip of its interacting molecule. However, the surrounding environment where the interaction is to take place also plays a big role and properties such as pH, ionic strength and ordered water have been reported to influence the binding process [19].

In this work the efficacy of interactions between the defense protein of blackgram and a viral pathogen, MYMIV-CP was investigated. It was found that the interactions of the two proteins were mediated in the MYMIV-resistant blackgram through two hydrogen bonded and nine Van der Waals contact. The total free energy of interactions was found to be  $-1535.1$  KJ/mol, indicative of a stable complex (Figure 7). The residues involved in the hydrogen bonded interactions—Lys519 and Thr490 were identified to be present in the largest pocket of the protein. When these residues were replaced with alanine following the principle of *in silico* mutagenesis, the stability of the formation of the complex suffered as indicated by the lack of formation of the hydrogen bond at the respective positions and the increase in the free energy of complex formation from



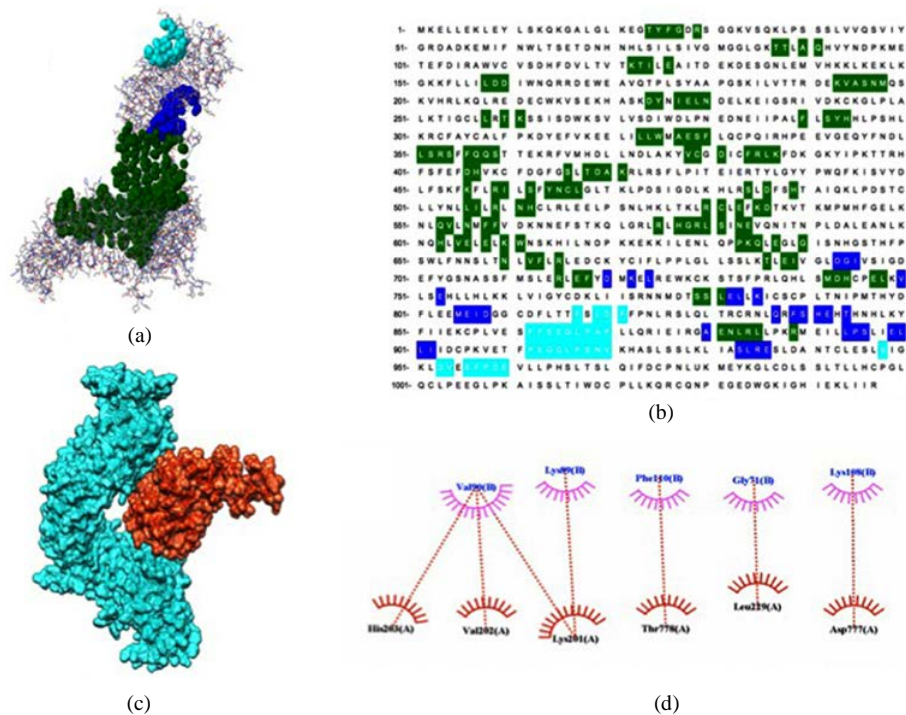
**Figure 7.** (a): Showing major pockets of CYR1; (b): Showing amino acids involved in the active pockets; (c): Macromolecular complex involving CYR1 and MYMIV coat-protein; (d): Van der Waal's interactions between CYR1 and MYMIV coat-protein.

–1535.1KJ/Mol to –1237.25 KJ/mol (Lys519Ala) and from –1535.1 KJ/Mol to –1146.3KJ/mol (Thr490Ala). These results indicate that both the residues were interacting hotspots and played significant role in regulating the interactions of the interacting partners. However, in case of *cyr1* and MYMIV-CP interactions, distinct differences were observed: the lack of formation of any intermolecular hydrogen bonds between the two proteins and a marked increase in the total free energy of interactions: 1291.35 KJ/mol. Of the six residues involved in Van der Waals interaction only Lys201 was identified to be present in a predicted pocket of the protein (Figure 8). A domain specific study of interacting residues led to the observation that two residues (marked in green) of the *cyr1* protein Asp777 and Thr778 part of the NBS-LRR region are involved in protein-protein interactions whereas only the Ile 558 (marked in blue) of the CYR1 protein NBS-LRR region is involved in Van der Waal's interaction (Figure 4).

The truncated N terminal coil coiled domain of the *cyr1* protein is probably responsible for the lack of conformational stability as an aspartate at position 112 is involved in CYR1-MYMIV interaction which is lacking for the *cyr1* protein. This residue when mutated to alanine resulted in alteration of the binding energy from 1535.1 KJ/Mol to –1357.25 KJ/mol, indicating that apart from the hydrogen bonded residues mentioned earlier this residue, is also an important hotspot for the protein-protein interaction and as none of the three hotspot residues are involved in protein binding in case of the *cyr1* protein, the efficacy of interaction is poor, resulting in susceptibility.

The present study further augments the observations of Maiti *et al.* [1], where we have studied the interactions of the isolated LRR domain of the CYR1 protein and the MYMIV-CP under the backdrop that the LRR domain mediates protein-protein and protein-ligand interactions and is involved in the specific recognition of pathogen-derived elicitors [20]-[22]. Since the work of Maiti *et al.* [1] was done using only the LRR domain, specific calculations and conformations of the macromolecular complex formation between the cognate interacting partners were not estimated.

Over the years several workers have indicated that true protein-protein interactions when simulated *in silico* is actually a successful energetic ensemble which has optimized representation and balance of both shape and physico-chemical complementarities, including Van der Waals electrostatic, hydrogen bonding, hydrophobic and solvent interactions [19] [23]. Since in case of *cyr1* and MYMIV interactions the complex formation is solely



**Figure 8.** (a): Showing major pockets of *cyr1*; (b): Showing amino acids involved in the active pockets; (c): Macromolecular complex involving *cyr1* and MYMIV coat-protein; (d): Van der Waal’s interactions between *cyr1* and MYMIV coat-protein.

dependent upon the Van der Waals contacts the stability stands compromised, which is not the case in the interactions between the CYR1 protein and MYMIV-CP. Thus, from the above data we can conclude that the lack of proper molecular recognition of the pathogen effector protein by *cyr1* of blackgram Cv. T9 leads to the establishment of host-pathogen compatible interaction. This study also confirms presence of the *cyr1* allele of a cognate MYMIV-resistance gene, *CYR1* in blackgram Cv. T9.

#### 4. Conclusion

Thus, the present study revealed that due to the truncation in the cognate “R” gene, that recognizes MYMIV effector protein, resulting in structural difference in the N-terminal of CC-domain of the protein *cyr1* the widely cultivated T9 variety of blackgram shows compatible interaction with MYMIV, *i.e.* susceptible to the virus.

#### Acknowledgements

This work was supported by the Council of Scientific and Industrial Research, New-Delhi, India for the Emeritus Scientist’s Scheme to AP (Sanction No. 21(0884)/12/EMR-II) and also by the Department of Biotechnology, India (Sanction no. BT/01/COE/06/03). We are thankful to the Director, Bose Institute for providing all infra-structural facilities.

#### References

- [1] Maiti, S., Paul, S. and Pal, A. (2012) Isolation, Characterization and Structure Analysis of a Non-TIR-NBS-LRR Encoding Candidate Gene from MYMIV-Resistant *Vigna mungo*. *Molecular Biotechnology*, **52**, 217-233. <http://dx.doi.org/10.1007/s12033-011-9488-1>
- [2] Kundu, A. and Pal, A. (2012) Identification and Characterization of Elite Inbred Lines with MYMIV-Resistance in *Vigna mungo*. *Field Crops Research*, **135**, 116-125. <http://dx.doi.org/10.1016/j.fcr.2012.07.006>
- [3] Feys, B.J., Moisan, L.J., Newman, M.A. and Parker, J.E. (2001) Direct Interaction between the *Arabidopsis* Disease Resistance Signaling Proteins, EDS1 and PAD4. *EMBOJ*, **20**, 5400-5411. <http://dx.doi.org/10.1093/emboj/20.19.5400>

- [4] 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. <http://dx.doi.org/10.1126/science.1136372>
- [5] Weaver, L.M., 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. <http://dx.doi.org/10.1111/j.1365-313X.2006.02834.x>
- [6] Swiderski, M.R., Birker, D. and Jones, J.D. (2009) The TIR Domain of TIR-NB-LRR Resistance Proteins Is a Signaling Domain Involved in Cell Death Induction. *Molecular Plant-Microbe Interactions*, **22**, 157-165. <http://dx.doi.org/10.1094/MPMI-22-2-0157>
- [7] Basak, J., Kundagramy, S., Ghose, T.K. and Pal, A. (2005) Development of Yellow Mosaic Virus (YMV) Resistance Linked DNA-Marker in *Vigna mungo* from Population Segregating for YMV Reaction. *Molecular Breeding*, **14**, 375-383. <http://dx.doi.org/10.1007/s11032-005-0238-6>
- [8] Kundu, A., Patel, A., Paul, S. and Pal, A. (2015) Transcript Dynamics at Early Stages of Molecular Interactions of MYMIV with Resistant and Susceptible Genotypes of the Leguminous Host, *Vigna mungo*. *PLoS One*, **10**, e0124687. <http://dx.doi.org/10.1371/journal.pone.0124687>
- [9] Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J.L., Dror, R.O. and Shaw, D.E. (2010) Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins*, **78**, 1950. <http://dx.doi.org/10.1002/prot.22711>
- [10] Ofra, Y. and Rost, B. (2007) Protein-Protein Interaction Hotspots Carved into Sequences. *PLOS Computational Biology*, **3**, e119. <http://dx.doi.org/10.1371/journal.pcbi.0030119>
- [11] 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. *The EMBO Journal*, **21**, 4511-4519. <http://dx.doi.org/10.1093/emboj/cdf453>
- [12] 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. <http://dx.doi.org/10.1105/tpc.106.042747>
- [13] Leister, T.R., 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. <http://dx.doi.org/10.1105/tpc.104.029637>
- [14] 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 Molecular Biology*, **61**, 31-45. <http://dx.doi.org/10.1007/s11103-005-5817-8>
- [15] Marone, D., Russo, M.A., Laidò, G., De, Leonardis, A.M. and Mastrangelo, A.M. (2013) Plant Nucleotide Binding Site—Leucine-Rich Repeat (NBS-LRR) Genes: Active Guardians in Host Defense Responses. *International Journal of Molecular Sciences*, **14**, 7302-7326. <http://dx.doi.org/10.3390/ijms14047302>
- [16] Hubert, D.A., Tornero, P., Belkhadir, Y., Krishna, P., Takahashi, A., Shirasu, K. and Dangl, J.L. (2003) Cytosolic HSP90 Associates with and Modulates the *Arabidopsis* RPM1 Disease Resistance Protein. *The EMBO Journal*, **22**, 5679-5689. <http://dx.doi.org/10.1093/emboj/cdg547>
- [17] 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. *The Journal of Biological Chemistry*, **279**, 2101-2108. <http://dx.doi.org/10.1074/jbc.M310029200>
- [18] Lu, R., Malcuit, T., 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. *The EMBO Journal*, **22**, 5690-5699. <http://dx.doi.org/10.1093/emboj/cdg546>
- [19] Henrich, S., Salo-Ahen, O.M., Huang, B., Rippmann, F.F., Cruciani, G. and Wade, R.C. (2010) Computational Approaches to Identifying and Characterizing Protein Binding Sites for Ligand Design. *Journal of Molecular Recognition*, **23**, 209-219.
- [20] Dangl, J.L. and Jones, J.D.G. (2001) Plant Pathogens and Integrated Defence Responses to Infection. *Nature*, **411**, 826-833. <http://dx.doi.org/10.1038/35081161>
- [21] Dodds, P.N., Lawrence, G.J., Catanzariti, A., Teh, T., Wang, C.I.A., Ayliffe, M.A., Kobe, B. and Ellis, J.G. (2006) Direct Protein Interaction Underlies Gene-for-Gene Specificity and Co-Evolution of the Flax L5/L6/L7 Resistance Genes and Flax Rust AvrL567 Avirulence Genes. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 8888-8893. <http://dx.doi.org/10.1073/pnas.0602577103>
- [22] Hammond-Kosack, K.E. and Jones, J.D.G. (1997) Plant Disease Resistance Genes. *Annual Review of Plant Physiology and Plant Molecular Biology*, **48**, 575-607. <http://dx.doi.org/10.1146/annurev.arplant.48.1.575>
- [23] Kahraman, A., Morris, R.J., Laskowski, R.A. and Thornton, J.M. (2007) Shape Variation in Protein Binding Pockets and Their Ligands. *Journal of Molecular Biology*, **368**, 283-301. <http://dx.doi.org/10.1016/j.jmb.2007.01.086>