

Molecular modeling and docking characterization of CzR1, a CC-NBS-LRR R-gene from *Curcuma zedoaria* Loeb. that confers resistance to *Pythium aphanidermatum*

Raj Kumar Joshi^{1*}, Satyabrata Nanda¹, Ellojita Rout¹, Basudeba Kar¹, Pradeep Kumar Naik² & Sanghamitra Nayak¹

¹Centre of Biotechnology, Siksha O Anusandhan University, Bhubaneswar-751003, India; ²Dept. of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Wakanaghat, Solan, Himachal Pradesh, India; Raj Kumar Joshi - Email: rajkumar.joshi@yahoo.co.in; Phone: 09437684176; *Corresponding author

Received May 25, 2013; Accepted May 27, 2013; Published June 29, 2013

Abstract:

Plant NBS-LRR R-genes recognizes several pathogen associated molecular patterns (PAMPs) and limit pathogen infection through a multifaceted defense response. CzR1, a coiled-coil-nucleotide-binding-site-leucine-rich repeat R-gene isolated from *Curcuma zedoaria* L exhibit constitutive resistance to different strains of *P. aphanidermatum*. Majority of the necrotrophic oomycetes are characterized by the presence of carbohydrate PAMPs β -glucans in their cell walls which interact with R-genes. In the present study, we predicted the 3D (three dimensional) structure of CzR1 based on homology modeling using the homology module of Prime through the Maestro interface of Schrodinger package ver 2.5. The docking investigation of CzR1 with β -glucan using the Glide software suggests that six amino acid residues, Ser186, Glu187, Ser263, Asp264, Asp355 and Tyr425 act as catalytic residues and are involved in hydrogen bonding with ligand β -(1,3)-D-Glucan. The calculated distance between the carboxylic oxygen atoms of Glu187-Asp355 pair is well within the distance of 5Å suggesting a positive glucanase activity of CzR1. Elucidation of these molecular characteristics will help in in silico screening and understanding the structural basis of ligand binding to CzR1 protein and pave new ways towards a broad spectrum rhizome rot resistance development in the cultivated turmeric.

Key words: *Curcuma zedoaria*, *Pythium aphanidermatum*, CC-NBS-LRR, molecular docking, β -(1, 3)-D-Glucan.

Background:

Curcuma longa L. (turmeric) of the family Zingiberaceae is one of the most important crop with great medicinal and economic significance [1]. India is the world's largest producer, and exporter of turmeric followed by China, Indonesia, Bangladesh and Thailand [2]. The International Trade Centre, Geneva, has estimated an annual growth rate of 10% in the world demand for turmeric. However, turmeric yield is not up to the demand because continuous domestication of the preferred genotypes coupled with their exclusive vegetative nature seems to have eroded the genetic base of this crop and as a result all of its

cultivars available today are equally susceptible to major fungal and bacterial diseases. Crop losses to the tune of 60% has been realized in the recent times mainly due to the infection by a necrotrophic oomycetic fungus *Pythium aphanidermatum* causing the rhizome rot disease in turmeric [2]. Due to several constraint associated with the utilization of chemical pesticides and host resistance breeding, a genetic transformation approach using alien genes could be the most likely solution towards development of rhizome rot resistance. Our lab has recently identified and characterized a coiled-coil-nucleotide-binding-site-leucine-rich repeat (CC-NBS-LRR) R-gene, CzR1 from wild

turmeric genotype *Curcuma zedoaria* that exhibit constitutive upregulated expressions in response to different strains of *P. aphanidermatum* [3]. The conceptual protein has 906 amino acids, a predicted molecular weight of 102473.57 Da and a calculated isoelectric point (pI) of 8.55.

Before utilization in long term turmeric improvement programmes, it is essential to determine the three dimensional (3D) model of *CzR1* to assess its structural integrity and biological relevance. A homology modeling based protein structure prediction of *CzR1* could be an efficient strategy taking into account the time consuming and cumbersome nature of X-ray diffraction or nuclear magnetic resonance spectroscopy (NMR) [4]. Further, several evidences suggest that various *R*-genes and pathogenesis related proteins play major role in neutralizing the oomycetic cell wall to kill the pathogen in line of defense signaling [5]. Beta-1,3-glucan, the major cell-wall constituents of necrotrophic oomycetes often acts as elicitors of defence reactions and serve as pathogen associated molecular patterns (PAMPs) [6]. Thus, the recognition of these PAMPs by *R*-genes are crucial to activate defense reactions in plants. However, there is no report on the interaction between *R*-gene and β -1,3-glucan to study the presence of a possible glucanase activity in *R*-gene dependent pathogen defense. Thus the present study was designed to predict the 3D structure of *CzR1* based on homology modeling and molecular docking of the *R*-protein with β -1, 3(D)-glucan.

Methodology:

A BLASTX search of the Protein Data Bank (PDB) using *CzR1* predicted protein as query showed high sequence identity (61%) with human APAF-1 protein (PDB ID: 1Z6T) and significant homology with coiled coil domain of *MLA10* gene of *Hordeum vulgare* (PDB ID: 3QFL) and leucine rich repeat domain from *Xanthomonas oryzae* (PDB ID: 4FCG). Models of *CzR1* protein were constructed based on the crystal structure coordinates of 1Z6T along with other two sequences at a resolution of 2.21 Å. The pair wise alignment of *CzR1* was done with 1Z6T as reference using the homology module of Prime through the Maestro interface (Schrodinger, LLC, and New York, USA). The pair wise alignment was improved manually by minor editing based on the secondary structure predictions as well as the pair wise superposition of the NB-ARC domain to develop the three dimensional structure of *CzR1*. The model was screened for unfavorable steric contacts and remodeled using a rotamer library database of Prime. Explicit hydrogen was added to the protein and the model was subjected to energy minimization using the force-field OPLS-2005 using 300 iterations in a simple minimization method. The steepest descent energy minimization was performed until the energy showed stability in the sequential repetition. *CzR1* model structure was evaluated for its backbone conformation by inspecting the Psi/Phi angles in a Ramachandran plot generated using the software PROCHECK version 3.5 [7]. Z score of *CzR1* and the template (1Z6T) was estimated using the software Prosa2003 [8]. Nevertheless, PROCHECK assured the reliability of the structure and the protein was subjected to VERIFY3D [9] available from NIH MBI Laboratory Servers. The root mean square deviation (RMSD) was calculated by superposing the predicted C α traces and backbone atoms of *CzR1* structure onto the template (1Z6T) in order to assess the the deviation of the modeled structure from the template. The

protein structure image of the model was illustrated using the software PyMOL.

Molecular docking was performed in order to determine the binding affinity of polymer of β -(1,3)-D-Glucan onto *CzR1* protein. The chemical structure β -(1,3)-D-Glucan was extracted from Chemical Book (<http://www.chemicalbook.com>) and its three dimensional structure was developed using the molecular builder tool of Maestro (Schrodinger package, ver 8.5). Initially different binding sites were predicted on *CzR1* protein using SiteMap (Schrodinger, ver 2.4), out of which only the binding site from the NB-ARC domain was considered for docking study. The receptor-grid file was generated at the centre of the predicted binding site using Glide (Schrodinger, ver 2.5). A bounding box of size 14Å x 14Å x 14Å was defined and centered on the mass center of each binding site in order to confine the mass center of the docked ligand. β -(1,3)-D-Glucan was then docked into each predicted binding site using Glide XP (extra precision) docking [10-11]. Out of the 5,000 poses that were sampled initially through exhaustive search of the torsional minima 800 poses per ligand were selected for energy minimization (conjugate gradients 1,000 steps). The 10 lowest-energy poses obtained were subjected to post docking minimization (Monte Carlo sampling based on torsional minima and refining the orientation of side groups of ligand). A single best conformation of the ligand was considered for further analysis and the binding affinity between the protein and ligand was estimated using the GlideScore function [12]. The mode of protein-ligand interaction was plotted using the program LIGPLOT [13].

Results & Discussion:

The *CzR1* protein from *C. zedoaria* with 906 amino acids contains NB-ARC domain (168 to 465) and LRR-1 domain (630 to 651) predicted from Pfam database [3]. Although *Mla1* protein from *Hordeum vulgare* shared 62% similarity with *CzR1*, it could not be used as a template due to non-availability of its structure information in the PDB database. Hence, the crystal structures of *MLA10* coiled coil domain (3QFL), human APAF-1 protein (1Z6T) and *Xanthomonas oryzae* leucine rich repeat domain (4FCG) were combinedly used as templates for developing a homology based model of *CzR1* (Figure 1A). The overall stereo chemical quality of the model was assessed by PROCHECK. The Ramachandran plot showed 96.6% amino acid residues from the main chain conformation of *CzR1* within the favored or allowed regions suggesting that majority of the amino acids are in a phi-psi distribution that is consistent with a right-handed α -helix. The G-factors, indicating the quality of the covalent bond length and overall bond angles, were -0.10° for dihedrals, 0.42° for covalent bonds, and -0.11° as overall value. Although, a calculated RMSD value of 0.448 Å (Figure 1B) was obtained by super positioning of the C α atoms of the *CzR1* and the template protein (1Z6T), the RMSD between the NB-ARC domain structures between both the proteins was only 0.284 Å. Prosa2003 analysis yielded a Z score of -4.32 for *CzR1* and -4.63 for the central template (1Z6T). The final model, which we took for further analysis, consisted of 631 amino acid residues. The overall main-chain and side-chain parameters, as evaluated by PROCHECK, are all very favorable. The assessment with VERIFY3D, which derives a "3D-1D" profile based on the local environment of each residue, described by the statistical preferences for: the area of the residue that is buried, the

fraction of side-chain area that is covered by polar atoms (oxygen and nitrogen) and the local secondary structure, also substantiated the reliability of the three dimensional structure.

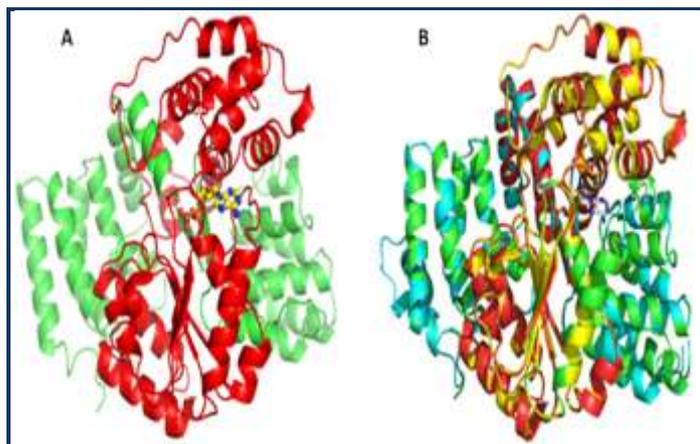


Figure 1: Homology model structure of *CzR1*. **A)** The NB-ARC domain of *CzR1* is represented in red colour and the LRRs were represented in green colour; **B)** Super positioning of *CzR1* and the template 1Z6T (yellow colour). The RMSD between both the structures was 0.448 Å. The RMSD between the NB-ARC domains between both the proteins was 0.284 Å.

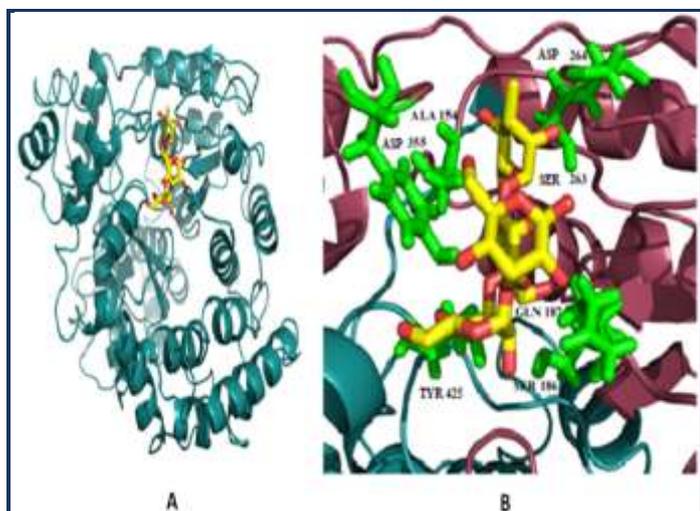


Figure 2: **A)** Molecular docking snapshot of β -(1,3)-D-Glucan bound to site 4 of NB_ARC domain in *CzR1*; **B)** Enlarge view of bound β -(1,3)-D-Glucan along with the interacting amino acids. β -(1,3)-D-Glucan is represented in yellow carbon colour while the binding site amino acids were represented in green colour

To examine the glucanase prospective of *CzR1*, we subjected it to molecular docking with the polymeric carbohydrate β -(1,3)-D-Glucan which is a major constituent of cell wall in most of the necrotrophic oomycetes [14]. Five different binding sites with different site scores were predicted with different physiochemical properties in Table 1 (see supplementary material). However, the binding site (site 4) that includes amino acids from the NB-ARC domain with good site score, volume, enclosure, hydrophilicity and donor/acceptor ratio was considered for molecular docking of β -(1,3)-D-Glucan. The docking show better docking score (-10.59 kcal/mol) and Emodel energy (-48.44 kcal/mol) of β -(1,3)-D-Glucan with the binding site 4 located at the NB-ARC domain of *CzR1* (Table

S2). β -(1,3)-D-Glucan shares higher values of electrostatic interaction (E_{coul} , -15.42 kcal/mol), hydrogen bonding energy (H_{bond} , -2.092 kcal/mol) and Van der Waals interaction (E_{vdw} , -33.03 kcal/mol) energies Table 2 (see supplementary material) predicted by Glide-XP because of the polar and non polar interactions with the binding site amino acids. The model produced by GlideXP predicts that six amino acid residues, Ser186, Glu187, Ser263, Asp264, Asp355 and Tyr425 are involved in hydrogen bonding with docked polymer β -(1,3)-D-Glucan (Figure 2A & 2B). The calculated distance between the carboxylic oxygen atoms of the Glu187-Asp355 pair is 4.38Å, which is well within the distance of 5Å required between the two catalytic residues for glucanase activity [15]. Furthermore, polymer of β -(1,3)-D-Glucan showed desirable hydrophobic interactions with the hydrophobic residues within the binding cavity (Figure 3). The favourable H-bonding, columbic, and hydrophobic interaction energies make polymer of β -(1,3)-D-Glucan highly active and selective for *CzR1*. Previous to our studies, there is no data on the possible glucanase or hydrolytic activity of the *R*-genes. However, many PR-proteins with anti-fungal activity exhibits such interdomain surface cleft for glycoside hydrolase enzymes and binding by (1,3)-b-D-glucan [16-17]. Further research on binding interactions between variable NBS-LRR class *R*-genes and different pathogen /microbe associated molecular patterns (PAMPs/MAMPs) is needed to get a decisive conclusion on the global functioning and regulation of such *R*-genes.

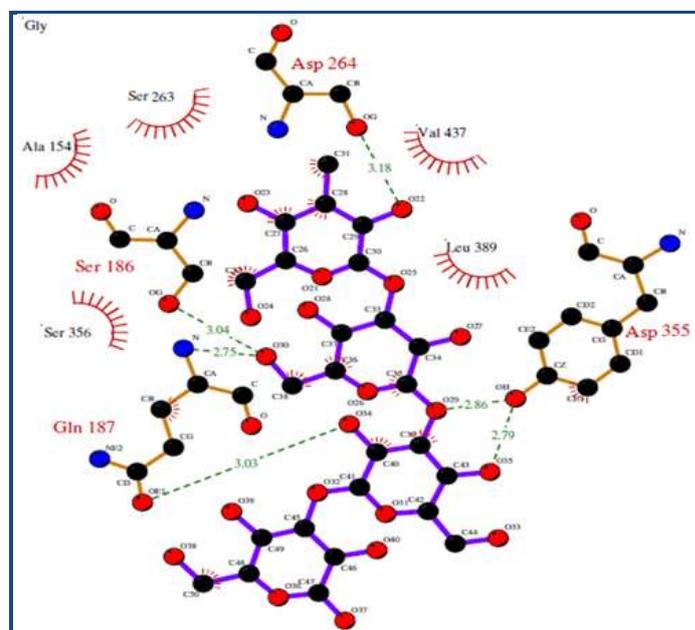


Figure 3: Ligand plot demonstrating the interacting amino acids of *CzR1* (site 4) with β -(1,3)-D-Glucan. The hydrogen bonds were represented with dotted green line along with the bond length in Å.

Conclusion:

In conclusions, molecular modeling procedures revealed the presence of a narrow cleft near the ATP-GTP binding site of the highly charged NBS domain of *CzR1*, a CC-NBS-LRR class *R*-gene from *C. zedoaria*. Molecular docking predicted high affinity between *CzR1* and (1,3)- β -D-glucan, a major component of the oomycete cell wall with favorable pairing of two negatively charged amino acid residues Glu187-Asp (355) within the

docked complex. This suggest that *CzR1* could be possibly having a β -1, 3-glucanase activity in vivo besides the regular ATP-GTP binding function in initiating pathogen defense. Revelation of these molecular characteristic of *CzR1* controlling the resistance of *C. zedoaria* against *P.aphanidermatum* together with studies on the significance of hydrolytic activity of *CzR1* will be highly informative towards the development of a broad spectrum rhizome rot resistance in the cultivated turmeric.

Acknowledgment:

SN gratefully acknowledges the financial assistance in the form of Institutional research fellowship from Siksha O Anusandhan University. ER acknowledges financial support in the form of junior research fellowship from Department of Science and Technology, Govt. of India. BK is thankful to Council of Scientific and Industrial research (CSIR), Govt. of India for providing senior research fellowship. The authors are grateful to Dr. Manoj Ranjan Nayak, President, Siksha O Anusandhan University for his encouragement and support.

References:

- [1] Ravindran PN *et al.* *Turmeric-the genus Curcuma* CRC Press 2007
- [2] Selvan *et al.* *Indian Spices- production and utilization.* 2002
- [3] Kar B *et al.* *Physiol Mol Plant Pathol.* 2013 **Doi:** 10.1016/j.pmpp.2013.05.003.
- [4] Yadav BS *et al.* *Exp Mol Pathol.* 2012 **92:** 7 [PMID: 22015804]
- [5] Asselbergh B *et al.* *Plant Physiol.* 2007 **144:** 1863 [PMID: 17573540]
- [6] Altenbach D & Robatzek S, *Mol Plant-Microbe Interact.* 2007 **20:** 1031 [PMID: 17849705].
- [7] Laskowski RA *et al.* *J Appl Crystallogr.* 1993 **26:** 283
- [8] Sippl MJ, *Proteins.* 1993 **17:** 355 [PMID: 8108378]
- [9] Eisenberg D *et al.* *Methods Enzymol.* 1997 **277:** 396 [PMID: 9379925]
- [10] Friesner *et al.* *J Med Chem.* 2004 **47:** 1739 [PMID: 15027865]
- [11] Halgren TA *et al.* *J Med Chem.* 2004 **47:** 1750 [PMID: 15027866]
- [12] Eldridge MD *et al.* *J Comput Aided Mol Des.* 1997 **11:** 425 [PMID: 9385547]
- [13] Wallace AC *et al.* *Protein Eng.* 1995 **8:** 127 [PMID: 7630882]
- [14] Latijnhouwers M *et al.* *Trends Microbiol.* 2003 **11:** 462 [PMID: 14557029]
- [15] Zechel DL & Withers SG, *Curr Opin Chem Biol.* 2001 **5:** 643 [PMID: 11738173]
- [16] Osmond RI *et al.* *Eur J Biochem.* 2001 **268:** 4190 [PMID: 11488912]
- [17] Ghosh R & Chakrabarti C, *Planta.* 2008 **228:** 883 [PMID: 18651170]

Edited by P Kanguane

Citation: Joshi *et al.* Bioinformation 9(11): 560-564 (2013)

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited

Supplementary material:

Table 1: Different binding sites predicted by SiteMap (Schrödinger Inc.) alongwith their physicochemical properties: site score, volume, enclosure, hydrophilicity and lower value of donor/acceptor ratio from CzR1 modeled structure. Site 4, located at the NB-ARC domain was used for the molecular docking of β -(1,3)-D-Glucan.

Site ID	Site score	Volume(\AA^3)	Exposure	Enclosure	Hydrophobic	Hydrophilic	Donor/Acceptor
1	1.169	272.34	0.402	0.951	0.565	1.287	0.736
2	1.169	272.34	0.402	0.951	0.565	1.287	0.736
3	1.089	409.88	0.606	0.731	1.423	0.667	0.863
4	1.001	389.31	0.685	0.718	0.571	0.873	1.966
5	0.945	235.98	0.556	0.731	0.298	0.893	2.974

Table 2: Glide XP docking pose of β -(1,3)-D-Glucan onto CzR1 and their energy scores. Glide XP docking generated more precise binding conformation of β -(1,3)-D-Glucan with good binding affinity.

Glide XPscore (kcal/mol)	Glide energy (kcal/mol)	Glide E_{vdw} (kcal/mol)	Glide E_{coul} (kcal/mol)	Glide H_{bond} (kcal/mol)	Glide E_{model} (kcal/mol)
-10.59	-48.44	-33.03	-15.42	-2.092	-72.598