Plant Omics Journal POJ

POJ 11(02):98-105 (2018) ISSN:1836-3644
doi: 10.21475/poj.11.02.18.1238

In silico identification of agriculturally important molecule(s) for defense induction against bacterial blight disease in Soybean (*Glycine max)*

Saril Mamgain, Shalini Dhiman, Rajesh Kumar Pathak, Mamta Baunthiyal*

Department of Biotechnology, Govind Ballabh Pant Institute of Engineering & Technology, Pauri Garhwal-246194, Uttarakhand, India

***Corresponding author: mamtabaunthiyal@yahoo.co.in**

Abstract

The productivity of *Glycine max* (Soybean), one of the economically important crops of India, is seriously affected by bacterial blight disease which is mainly caused by *Psedomonas syringae.* The disease results in significant yield losses in Soybean crops. Since no proven resistant source is available against bacterial blight, the only option remaining is to utilize biotechnological strategies which could lead to inhibition of pathogenic proteins of the *Psedomonas syringae* responsible for disease progression. Phytoalexins are well known to inhibit bacterial growth and trigger defense response against diseases in crop plants. The present study was conducted to identify the molecules which could inhibit the growth and development of bacteria. A few proteins were selected from literature analysis viz., Ornithine carbamoyl transferase 2, phaseolotoxin-insensitive, avirulence protein AvrRpt2, HarpinHrpZ, Sensor protein GacS, and Translation initiation factor IF-3 of *Psedomonas syringae* as possible molecular targets of phytoalexins. The molecular modeling of these proteins were done by using their amino acid sequence on Phyre2 and I-TASSER tool followed by model validation through energy minimization and Ramachandran plot analysis. Subsequently molecular docking was performed using some selected phytoalexins produced by members of *Brassicaceae*, *Fabaceae*, *Solanaceae*, *Vitaceae* and *Poaceae* family with each modeled protein structure by AutoDock vina. Based on the molecular docking study, we identified efficient defense molecules, which can be used for the development of agrochemicals for protection of *G. max* against infection of *Psedomonas syringae*.

Key words: *Glycine max*; *Pseudomonas syringae;* Bacterial blight*;* Molecular docking; Agriculturally important molecule.
Abbreviations: *G. Max_Glycine max*; *P. syringae_Psedomonas syringae;* BLAST_Basic local alignment search tool; PDB_Protein data bank; SPDB_Swiss PDB Viewer; ProSA_Protein Structure Analysis; ProQ_Protein model quality prediction.

Introduction

Soybean (*Glycine max*) is one of the most important crops in the World. It is a source of dietary proteins and oil. Soybeans are beneficial for health since it contains α-Linolenic acid, Iso-flavones, Lecithins, Lectins and Linoleic acid etc (Ahuja et al., 2012). The crop is greatly challenged by bacterial pathogen, *Pseudomonas syringae* that is a gram-negative rod-shaped bacterium, which infects a variety of plant species. In Soybean it causes bacterial blight disease on leaves as well as on stems, petioles, and pods (Park, 1986; Block et al., 2010). It is an early season disease, which appears in winters in the field on Soybean. Primarily the disease appears when wind or splashing water droplets spread *Pseudomonas* cells from crop plant on the soil surface to leaves. The *Pseudomonas* enters the Soybean leaves through stomata subsequently producing a toxin that is responsible for breakdown of Chlorophyll (Chen et al., 2000). From literature studies, some proteins were identified for their role in pathogenesis viz., Ornithine carbamoyl transferase 2, phaseolotoxin-insensitive avirulence protein AvrRpt2, HarpinHrpZ, Sensor protein GacS, and Translation initiation factor IF-3 (Hrabak et al., 1992; He et al., 1993; Dixit et al., 2011; Echeverri et al., 2012; González et al., 2012; Jeandet, 2015). Ornithine carbamoyltransferase 2, phaseolotoxin-insensitive plays a vital role in the continued existence and pathogenicity of *P. syringae*; Cysteine protease avirulence protein AvrRpt2 is required for the degradation of plant cell RIN4 and consequent activation of RPS2 during bacterial infection by thiol protease activity. The activation of RPS2 is enough for the initiation of hypersensitive response (HR) and resistance in plants. The cleavage of RIN4 by AvrRpt2 also interferes with RPM1-mediated resistance activated by either AvrRpm1 or AvrB. It inhibits PAMP (pathogen-associated molecular patterns) and blocks the plant callose deposition. The molecular mechanism of virulence is not known as yet, but this activity is independent of ethylene and salicylic acid response pathways as well as independent of RIN4 disappearance. Sensor protein GacS may be involved in lesion production, swarming and in the formation of extracellular protease, syringomycin as well as N-acyl-L-homoserine lactone (acyl-HSL) and is required for pathogenicity on bean plants. Harpin HrpZ are proteins

capable to elicit hypersensitive response (HR) in non-host plants and are required for pathogenesis in host plants. HrpZ forms types of ion-conducting pores permeable for cations. Such pore-forming activity may allow nutrient release and are responsible for delivery of virulence factors during bacterial colonization of host plants.

In comparison with animal immune systems, plants lack adaptive immune systems, they produce several short-lived molecules (generally within 72-96 hours) to protect themselves from the attack of abiotic and biotic stresses (Jeandet et al., 2014). These molecules have been known as phytoalexins, but as the intensity of these stresses increases, the defense responses rapidly decrease in time dependant manner (Kitten and Willis, 1996; Kelley et al., 2015; Pathak et al., 2013). Studies have demonstrated that the exogenous application of phytoalexins is able to induce defense in a concentration-dependent fashion (Mukherjee et al., 2011; Lamberth et al., 2013; Jeandet et al., 2014; Kumar et al., 2015). Computational approaches may facilitate understanding of the interactions of phytoalexins with selected target pathogenic proteins. Such knowledge may allow designing and development of novel agrochemicals against bacterial blight disease caused by *Pseudomonas syringae*. It will also allow better understanding of the plant–pathogen interaction at the molecular level (Kumar et al., 2015; Pathak et al., 2017a).

Results and Discussion

Primary sequence analysis and prediction of secondary structure

The amino acid sequence of the target proteins, viz., Ornithine carbamoyl transferase 2 phaseolotoxin-insensitive, avirulence protein AvrRpt2, HarpinHrpZ, Sensor protein GacS, Translation initiation factor IF-3 from *Psedomonas syringae*, which are probable molecular target against BB disease of *Glycine max,* were downloaded from Uniprot database (http://www.uniprot.org/) (Table 1). These sequences were subjected to BLASTp (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) program for identification of suitable template structure and method that is used for modeling of its three dimensional structure. The SOPMA online tools were employed for identification of amino acid residues involved in the formation of protein secondary structure i.e., helix, sheet, coil and turn, this information can be also utilized for the modeling of the protein structure (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html).

3D structure modeling, model quality assessment and validation

Advances in computational biology techniques to predict protein structure and function reduces the experimental cost and time. On average, 50–70% of a genome can be structurally modeled using computational techniques (Whalen et al., 1991). Widely used method for protein structure prediction involved comparing a sequence of interest with available database sequences, to build an evolutionary or statistical profile of that sequence and then scans this profile against a database profiles for known structures. This results in an alignment between two sequences, one of unknown structure and one of known structure at sequence level. On the basis of such comparison, a model of sequence using the information of known structure(s) sequence is built. When the similarity of sequence between the protein of interest and the protein(s) structure available in database is low, then detection of the relationship and the subsequent alignment may be enhanced if structural information is included to augment the analysis of sequence.

In the present study, Pyre2 and I-TASSER were employed to model 3D structure of target protein sequences (Yang et al., 2015). SPDB Viewer was used for stabilizing its stereochemical properties through energy minimization. The stability of the target protein model was validated by RAMPAGE server. The Ramachandran plot statistics of target protein models in range of 77.3 to 96.7% (favored regions), 2.2 to 15.9% (allowed regions) and 0.8 to 6.6 (outer regions) are shown in Table 2**.** The results of the RAMPAGE analysis revealed that the relatively low percentage of residues have phi/psi angles in the outer regions suggesting the acceptability of Ramachandran plots for target proteins. The stereochemical quality of the predicted models found to be satisfactory were taken for molecular docking studies (Pathak et al., 2013; 2016).

Ligand designing and molecular docking

The structure of selected phytoalexins was constructed using ChemSketch http://www.acdlabs.com/resources/freeware/chemsketch/. Each phytoalexins structure and target protein structures were imported to AutoDock tool for preparation of ligand.pdbqt and protein.pdbqt file. PDBQT is a type of file format that stores atomic coordinates, partial charges and all other information of macromolecules required for molecular docking in AutoDock vina. The each target protein model was entirely covered under the grid box and made configration file for molecular docking with each phytoalexins. For each compound, out of the many docking poses, only those having the highest docking score and relatively good hydrogen bond interaction were selected for further analysis. Arachidin2 was docked with Ornithine carbamoyltransferase2, phaseolotoxin-insensitive with binding free energy -6.4 kcal/mol; Cysteine protease avirulence protein AvrRpt2 with binding free energy -8.1 Kcal/mol; Sensor protein GacS with binding free energy –7.5 Kcal/mol; Harpin HrpZ with binding free energy -8.5 Kcal/mol and Translation initiation factor IF-3 with binding free energy -6.8 Kcal\mol respectively. Arachidine 3 was docked with all proteins with binding free energy -6.8, -8.0, -7.7, -8.4 and -6.7 Kcal/mol; Arachidin 1 was docked with all proteins with free binding energy -7.3, -8.2, -7.5, -8.5 and -6.8 Kcal/mol; Avenanthramide was docked with all proteins with binding free energy -6.7, -8.0, -7.3, -8.5 and -7.5 Kcal/mol; Brassilexin was docked with all proteins with binding free energy -5.2, -6.8, -5.5, -6.6 and -6.9 Kcal/mol; Camalexin was docked with all proteins with binding free energy -5.5, -6.8, -5.7, -6.7, -6.9 Kcal/mol; Capsidiol was docked with all proteins with binding free energy -5.5, -6.3, -5.6, -6.7, -5.5 Kcal/mol; E_viniferin was docked with all proteins with binding free energy -8.1, -8.9,

Table 1. Detail information regarding pathogenic protein sequences of *P. syringae.*

S.N.	Name of protein	Length	UniProtKB ID
1	Ornithine carbamoyltransferase 2, phaseolotoxin-insensitive	327	P68746
2	Cysteine protease avirulence protein AvrRpt2	255	Q6LAD6
3	Sensor protein GacS	907	P48027
4	Harpin HrpZ	341	P35674
5	Translation initiation factor IF-3	183	P0A133

Table 2. Ramachandran plot statistics of top *P. syringae* proteins model.

Protein name	Favored region (%)	Allowed region (%)	Outlier region (%)
Ornithine carbamoyltransferase 2, phaseolotoxin-insensitive	92.5	5.7	1.6
Cysteine protease avirulence protein AvrRpt2	92.5	3.7	3.7
Sensor protein GacS	95.9	3.3	0.8
Harpin HrpZ	77.3	15.9	6.8
Translation initiation factor IF-3	96.7	2.2	1.1

Table 3. Molecular docking studies of selected phytoalexins with pathogenic proteins: illustrates minimum binding free energy (Kcal/mol).

Receptors Ligands	Ornithine carbamoyltransferase 2, phaseolotoxin-insensitive	Cysteine protease avirulence protein AvrRpt2	Sensor protein GacS	Harpin HrpZ	Translation initiation factor IF-3
Arachidin2	-6.4	-8.1	-7.5	-8.5	-6.8
Arachidine3	-6.8	-8.0	-7.7	-8.4	-6.7
Arachidin1	-7.3	-8.2	-7.5	-8.5	-6.8
Avenanthramide	-6.7	-8.0	-7.3	-8.5	-7.5
Brassilexin	-5.2	-6.8	-5.5	-6.6	-6.9
Camelexin	-5.5	-6.8	-5.7	-6.7	-6.9
Capsidiol	-5.5	-6.3	-5.6	-6.7	-5.5
ε-viniferin	-8.1	-8.9	-8.2	-9.7	-5.9
Glyceollin I	-8.1	-7.7	-8.3	-9.3	-6.3
Glyceollin II	-7.6	-8.5	-7.7	-9.1	-6.5
Kauralexin A1	-7.3	-7.2	-7.8	-8.5	-5.5
KauralexinB1	-7.2	-7.2	-7.2	-8.6	-5.6
Luteolin	-7.5	-8.1	-7.5	-8.8	-7.2
Medicarpin	-6.8	-7.8	-7.0	-8.4	-7.1
Monilactone A	-8.5	-8.0	-6.9	-8.1	-5.5
Phytocassane A	-7.2	-7.5	-7.3	-8.9	-6.2
Pistin	-7.5	-8.2	-7.4	-8.2	-5.7
Resveratrol	-6.2	-7.4	-6.9	-7.6	-6.4
Rutalexin	-6.1	-7.1	-6.4	-7.5	-5.8
Sakuranetin	-6.9	-7.9	-6.9	-8.3	-7.0
Scopoletin	-6.0	-6.7	-6.2	-6.8	-6.7
Spirobrassin	-5.6	-6.3	-6.2	-7.3	-5.5
Wighteone	-7.0	-9.0	-7.3	-9.1	-7.5
Zealexin A1	-6.2	-7.7	-6.6	-7.7	-7.2
Zealexin B1	-6.1	-7.9	-6.9	-7.6	-7.9

Table 4. Amino acids residues of pathogenic proteins involved in interactions with top identified phytoalexins viz., E_viniferin, Glyceollin I and Monilactone A, minimum binding free energy corresponding to each phytoalexins are given.

S.N.	Proteins	E_viniferin		Glyceollin I		Monilactone A	
		Binding free energy (Kcal/mol)	Amino acid residues involved in protein-ligand interaction	Binding free energy (Kcal/mol)	Amino acid residues involved in protein-ligand interaction	Binding free energy (Kcal/mol)	Amino acid residues involved in protein-ligand interaction
1	Ornithine carbamoyltransferase 2, phaseolotoxin-insensitive	-8.1	Lys56, Arg109, Gln139, Asn168, Arg313,	-8.1	Thr61, Arg313	-8.5	Ser58, Arg60, Thr61, His136
2	Cysteine protease avirulence protein AvrRpt2	-8.9	Val173, Asp177, Ser206	-7.7	No H bond interactions	-8.0	Ser133
3	Sensor protein GacS	-8.2	Gly295, His298, Gln301	-8.3	Ile360	-6.9	Lys275
4	Harpin HrpZ	-9.7	Ser122, Asp125, Asn337	-9.3	Ser240, Thr305, Ala304, Asp308	-8.1	Leu177
5	Translation initiation factor IF-3	-5.9	Glu104, Gln108, Asp150, Asp151	-6.3	Asp151	-5.5	Gln108

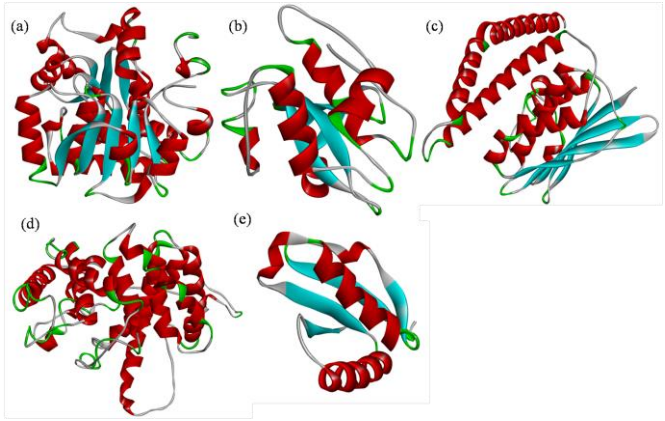

Fig 1. 3D Structure of pathogenic proteins of *Psedomonas syringae* (a) Ornithine carbamoyltransferase 2, phaseolotoxin-insensitive, (b) Cysteine protease avirulence protein AvrRpt2, (c) Sensor protein GacS, (d) Harpin HrpZ, and (e) Translation initiation factor IF-3 modeled through computational methods using amino acid sequence.

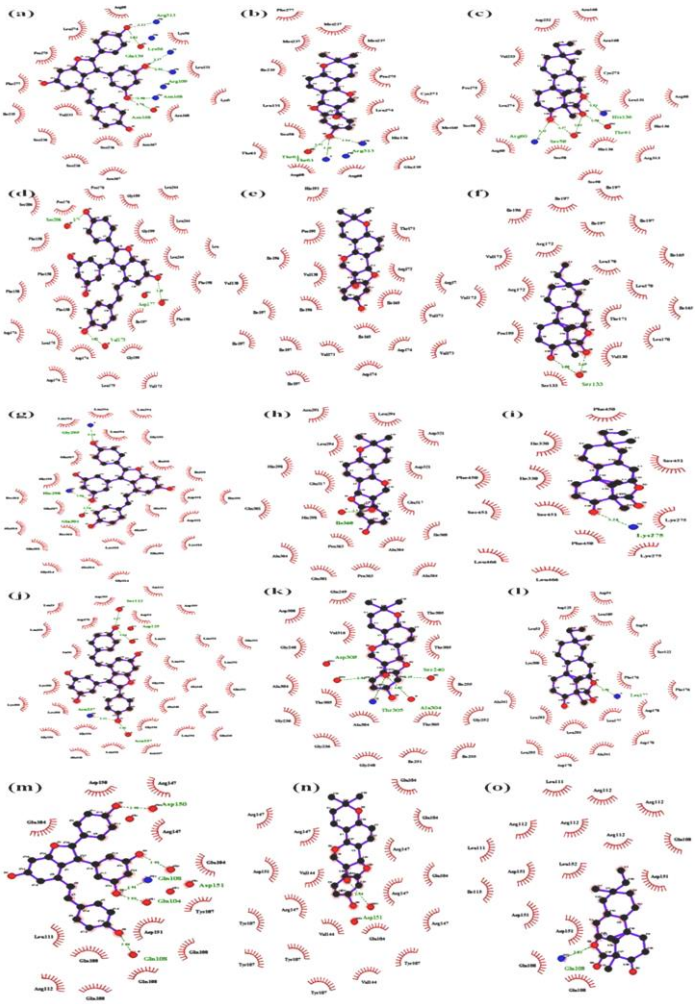

Fig 2. 2D interaction diagram of the docked structures of top identified ligands with pathogenic proteins generated by Ligplot (a) Ornithine carbamoyltransferase 2, phaseolotoxin-insensitive docked with E_viniferin; (b) Ornithine carbamoyltransferase 2, phaseolotoxin-insensitive docked with Glyceollin I; (c) Ornithine carbamoyltransferase 2, phaseolotoxin-insensitive docked with Monilactone A; (d) Cysteine protease avirulence protein AvrRpt2 docked with E_viniferin; (e) Cysteine protease avirulence protein AvrRpt2 docked with Glyceollin I; (f) Cysteine protease avirulence protein AvrRpt2 docked with Monilactone A; (g) Sensor protein GacS docked with E_viniferin; (h) Sensor protein GacS docked with Glyceollin I; (i) Sensor protein GacS docked with Monilactone A; (j) Harpin HrpZ docked with E_viniferin; (k) Harpin HrpZ docked with Glyceollin I; (l) Harpin HrpZ docked with Monilactone A; (m) Translation initiation factor IF-3 docked with E_viniferin; (n) Translation initiation factor IF-3 docked with Glyceollin I; (o) Translation initiation factor IF-3 docked with Monilactone A; showing H-bond$_{(green)}$ and Hydrophobic$_{(red)}$ interactions.

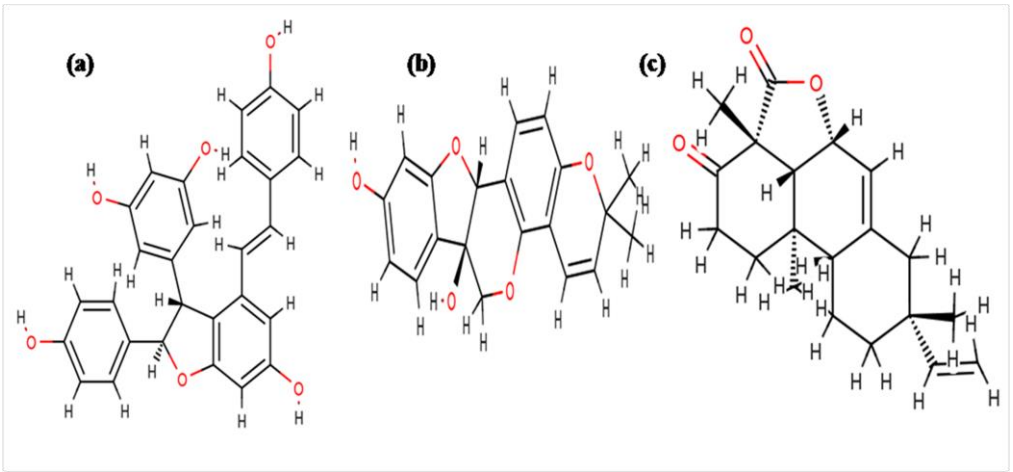

Fig 3. Structure of (a) ε-viniferin (b) Glyceollin I (c) Monilactone A: Agriculturally important molecules identified through molecular docking studies.

-8.2, -9.7 and -5.9 Kcal/mol; Glyceollin I was docked with all proteins with binding free energy -8.1, -7.7, -8.3, -9.3, -6.3 Kcal/mol; Glyceollin II was docked with all proteins with binding free energy -7.6, -8.5, -7.7, -9.1 and -6.5 Kcal/mol; Kauralexin A1 was docked with all proteins with binding free energy -7.3, -7.2, -7.8, -8.5 and -5.5 Kcal/mol; Kauralexin B1 was docked with all proteins with binding free energy -7.2, -7.2, -7.2, -8.6 and -5.6 Kcal/mol; Luteolin was docked with all proteins with binding free energy -7.5, -8.1, -7.5, -8.8 and -7.2 Kcal/mol; Medicarpin was docked with all proteins with binding free energy -6.8, -7.8, -7.0, -8.4 and -7.1 Kcal/mol; Monilactone A was docked with all proteins with binding free energy -8.5, -8.0, -6.9, -8.1 and -5.5 Kcal/mol; Phytocassane A was docked with all proteins with binding free energy -7.2, -7.5, -7.3, -8.9 and -6.2 Kcal/mol; Pistin was docked with all proteins with binding free energy -7.5, -8.2, -7.4, -8.2 and -5.7 Kcal/mol; Resveratrol was docked with all proteins with binding free energy -6.2, -7.4, -6.9, -7.6 and 6.4 Kcal/mol; Rutalexin was docked with all proteins with binding free energy -6.1, -7.1, -6.4, -7.5 and -5.8 Kcal/mol; Sakuranetin was docked with all proteins with binding free energy -6.9, -7.9, -6.9, -8.3 and -7.0 Kcal/mol; Scopoletin was docked with all proteins with binding free energy -6.0, -6.7, -6.2, -6.8 and -6.7 Kcal/mol; Spirobrassinin was docked with all proteins with binding free energy -5.6, -6.3, -6.2, -7.3 and -5.5 Kcal/mol; Wighteone was docked with all proteins with binding free energy -7.0, -9.0, -7.3, -9.1 and -7.5 Kcal/mol; Zealexin A1 was docked with all proteins with binding free energy -6.2, -7.7, -6.6, -7.7 and -7.2 Kcal/mol; Zealexin B1 was docked with all proteins with binding free energy -6.1, -7.9, -6.9, -7.6 and -7.9 Kcal/mol (Fig. 2; Table 3).

The world's population is increasing rapidly and could reach to 9 billion people in year 2050. Agriculture is a major driver for securing food and nutritional security but crop plants are challenged with several pathogenic microorganisms that directly attack on crops, reducing agricultural productivity. Bioinformatics has immense potential to identify novel agrochemicals that can protect crop plants and thus increase agricultural production. Bioinformatics based discovery and designing is relatively new in agrochemical industries. The development of scytalone dehydratase inhibitors for rice blast as fungicides is one of the most detailed examples (Walter, 2002). Bioinformatics-based identification of agriculturally important molecules may reduce use of hazardous bactericides (Wiederstein et al., 2007; Pathak et al., 2017b).

Identification of Agriculturally important lead molecule(s)

A response to several biotic and abiotic stresses phytoalexins E_viniferin is produced by the members of Vitaceae family; Glyceollin I is produced by the members of Fabaceae family and Monilactone A is produced by members of Poaceae family (Ahuja et al., 2012), These phytoalexins demonstrated greatest affinity towards pathogenic proteins of *Pseudomonas syringae* as compared to other phytoalexins selected in present study during docking simulation. E_viniferin showed H-bond interactions with Ornithine Carbamoyltransferase 2, phaseoltoxin-insensitive amino acid residue Lys56, Arg109, Gln139, Asn168 and Arg313 with binding free energy -8.1 Kcal/mol; Cystein protease avirulence protein AvrRpt2 amino acid residue Val173, Asp177, Ser206 with binding free energy -8.9 Kcal/mol; Sensor protein GacS amino acid residue Gly295, His298 and Gln301 binding free energy -8.2 Kcal/mol; Harpin HrpZ amino acid residue Ser122, Asp125, and Asn337 with binding free energy -9.7 Kcal/mol and Translation initiation factor IF-3 amino acid residue, Glu104, Gln108, Asp150 and Asp151 with binding free energy -5.9 Kcal/mol. Glyceollin I showed H-bond interaction with Ornithine carbamoyl transferase 2, phaseolotoxin insensitive amino acid residue Thr61, Arg313 with binding free energy -8.1 Kcal/mol; Cysteine protease avirulence protein AvrRpt2 with binding free energy -7.7 Kcal/mol, however no hydrogen bond interaction were detected during Ligplot analysis; Sensor protein GacS amino acid residue Ile360 with binding free energy -8.3 Kcal/mol; Harpin HrpZ amino acid residues Ser240, Thr305, Ala304, Asp308 with binding free energy -9.3 Kcal/mol and Translation initiation factor IF-3 amino acid residue Asp151 with binding free energy -6.3 Kcal/mol. Monilactone A showed H-bond interaction with Ornithine carbamoyltransferase 2, phaseolotoxin insensitive amino acid residue Ser58, Arg60, Thr61 and His136 with binding free energy -8.5 Kcal/mol; Cysteine protease avirulence protein AvrRpt2 amino acid residue Ser133 with binding free energy -8.0 Kcal/mol; Sensor protein GacS amino acid residue Lys275 with binding free energy -8.2 Kcal/mol; Harpin HrpZ amino acid residue Leu177 with binding free energy -9.7 Kcal/mol and Translation initiation factor IF-3 amino acid residue Gln108 with binding free energy -5.5 Kcal/mol (Fig. 3; Table 4). The results of present study suggest that the phytoalexin *ε*-viniferin, Glyceollin I and Monilactone A may decrease the effect of pathogen on Soybean crop against Bacterial Blight disease.

Materials and Methods

Retrieval of target sequences

The amino acid sequence of the target pathogenic proteins of *P. syringae* viz., Ornithin Carbomoyltransferase 2, phaseolotoxin-insensitiive; Cysteine protease avirulence protein AvrRpt2; Sensor protein GacS; Harpin HrpZ; and Translation initiation factor IF-3 were retrieved from UniProt database (http://www.uniprot.org/) (Table 1).

Secondary structure prediction

Amino acid sequence of target proteins were subjected to SOPMA server (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) for identification of amino acid residue involved in the formation of secondary structure i.e. Helix, Sheet, Turn and Coil.

3D structure prediction and visualization

Target protein sequence of *P. syringae* was subjected to NCBI BLASTp (Altschul et al., 1990) against PDB database (http://www.rcsb.org/pdb/home/home.do)for identification of suitable template as well as determined method for modeling of 3D model. In addition to BLASTp search, Pyre2 and I-TASSER were employed to model 3D structure of target protein sequences (Fig. 1). Discovery studio visualizer was used for the visualization of 3D models

(http://accelrys.com/products/collaborative-science/biovia-discovery-studio/).

Quality assessment and evaluation of predicted model

The structural refinements through energy minimization of each predicted models were performed by SPDB viewer (http://spdbv.vital-it.ch/refs.html). The overall quality of each models were done by ProSA (Protein Structure Analysis) (Sawada et al., 1997) and ProQ (Protein model quality prediction) (Trott and Olson, 2010). RAMPAGE server was used to analyze the Ramachandran plot of the predicted protein models (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php) (Table 2).

Ligand designing and molecular docking

The structure of each selected phytoalexins was constructed using ChemSketch (http://www.acdlabs.com/resources/freeware/chemsketch/). Molecular docking was performed using phytoalexins with pathogenic proteins by AutoDock vina (Wallner and Elofsson, 2003), which requires the three dimensional structure of ligands and receptors and uses a sophisticated gradient optimization method for molecular docking.

Conclusion

The present computational study provides an insight about the interactions between phytoalexins and pathogenic protein of *Psedomonas syringae.* It explains the mode of inhibition of bacterial activity. The phytoalexins *ε*-viniferin, Glyceollin I and Monilactone A may be utilized for further studies to validate its role for curtailing the incidence of bacterial blight disease of Soybean and other agriculturally important diseases of crop plants. This information may be beneficial for scientific community to develop novel molecules using phytoalexins for sustainable agriculture that directly replace the use of hazardous bactericides, thus ensuring food and nutrition security of the rapidly growing world population.

Acknowledgements

Authors are grateful to TEQIP-II for financial assistance and Bioinformatics Laboratory, Department of Biotechnology, G. B. Pant Institute of Engineering & Technology, Pauri Garhwal, Uttarakhand, India for providing research facilities. One of the authors RKP would like to thank CSIR-India for providing senior research fellowship.

Compliance with Ethical Standards

The authors declare that they have no competing financial interests.

References

Ahuja I, Kissen R, Bones AM (2012) Phytoalexins in defense against pathogens. Trends Plant Sci. 17(2): 73-90.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Boi. 215(3): 403-410.

Block A, Guo M, Li G, Elowsky C, Clemente TE, Alfano JR (2010) The *Pseudomonas syringae* type III effector HopG1 targets mitochondria, alters plant development and suppresses plant innate immunity. Cell Microbiol. 12(3): 318–330.

Chen Z, Kloek AP, Boch J, Katagiri F, Kunkel BN (2000) The *Pseudomonas syringae* avrRpt2 gene product promotes pathogen virulence from inside plant cells. Mol Plant Micro Int. 13(12): 1312–1321.

Dixit AK, Antony JIX, Sharma NK, Tiwari RK (2011) Soybean constituents and their functional benefits. Proteins. 661(2): 367–383.

Echeverri F, Torres F, Quiñones W, Escobar G, Archbold R (2012) Phenylphenalenone phytoalexins, will they be a new type of fungicide. Phytochem Rev. 11(1): 1-12.

González AJ, Fernández AM, José MS, González-Varela G, Rodicio MR (2012) A pseudomonas viridiflava-related bacterium causes a dark-reddish spot disease in *Glycine max*. App Environ Microbiol. 78(10): 3756–3758.

He SY, Huang HC, Collmer A (1993) *Pseudomonas syringae pv. syringae harpinPss*: a protein that is secreted via the Hrp pathway and elicits the hypersensitive response in plants. Cell. 73(7): 1255–1266.

Hrabak EM, Willis DK (1992) The lemA gene required for pathogenicity of Pseudomonas syringae pv. syringae on bean is a member of a family of two-component regulators. J Bacteriol. 174(9): 3011–3020.

Jeandet P (2015) Phytoalexins: current progress and future prospects. Mol. 20(2): 2770-2774.

Jeandet P, Hébrard C, Deville MA, Cordelier S, Dorey S, Aziz A, Crouzet J (2014) Deciphering the role of phytoalexins in plant-microorganism interactions and human health. Mol. 19(11): 18033-18056.

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ (2015) The phyre2 web portal for protein modeling, prediction and analysis. Nat Prot. 10(6): 845-858.

Kitten T, Willis DK (1996) Suppression of a sensor kinase-dependent phenotype in *Pseudomonas syringae* by ribosomal proteins L35 and L20. J Bacteriol. 178(6): 1548–1555.

Kumar A, Pathak RK, Gupta SM, Gaur VS, Pandey D (2015) Systems biology for smart crops and agricultural innovation: filling the gaps between genotype and phenotype for complex traits linked with robust agricultural productivity and sustainability. OMICS. 19(10): 581-601.

Lamberth C, Jeanmart S, Luksch T, Plant A (2013) Current challenges and trends in the discovery of agrochemicals. Sci. 341(6147): 742-746.

Mukherjee S, Szilagyi A, Roy A, Zhang Y (2011) Genome-wide protein structure prediction. in multiscale approaches to protein model. Springer New York. 255-279.

Park EW (1986) Effects of bacterial blight on soybean yield. Plant Dis. 70(3): 14.

Pathak RK, Baunthiyal M, Pandey N, Pandey D, Kumar A (2017a) Modeling of the jasmonate signaling pathway in *Arabidopsis thaliana* with respect to pathophysiology of *Alternaria* blight in *Brassica.* Sci Rep. 7: 16790.

Pathak RK, Baunthiyal M, Shukla R, Pandey D, Taj G, Kumar A (2017b) *In Silico* identification of mimicking molecules as defense inducers triggering jasmonic acid mediated immunity against *Alternaria blight* disease in *Brassica species*. Front Plant Sci. 8.

Pathak RK, Giri P, Taj G, Kumar A (2013) Molecular modeling and docking approach to predict the potential interacting partners involved in various biological processes of mapk with downstream wrky transcription factor family in *Arabidopsis thaliana*. Int J Comput Bioinfo In Silico Model. 2: 262-268.

Pathak RK, Taj G, Pandey D, Arora S, Kumar A (2013) Modeling of the mapk machinery activation in response to various abiotic and biotic stresses in plants by a system biology approach. Bioinform. 9(9): 443.

Pathak RK, Taj G, Pandey D, Kasana VK, Baunthiyal M, Kumar A (2016) Molecular modeling and docking studies of phytoalexin(s) with pathogenic protein(s) as molecular targets for designing the derivatives with anti-fungal action on *Alternaria spp.* of *Brassica*. Plant Omics. 9 (3), 172-182.

Sawada H, Takeuchi T, Matsuda I (1997) Comparative analysis of *Pseudomonas syringae* pv. actinidiae and pv. phaseolicola based on phaseolotoxin-resistant ornithine carbamoyltransferase gene (argK) and 16S-23S rRNA intergenic spacer sequences. Appl Env Microbio. 63(1): 282–288.

Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Computat Chem. 31(2): 455-461.

Wallner B, Elofsson A, (2003) Can correct protein models be identified?. Prot Sci. 12(5): 1073-1086.

Walter MW (2002) Structure-based design of agrochemicals. Nat Product Rep. 19: 278

Whalen MC, Innes RW, Bent AF, Staskawicz BJ (1991) Identification of *Pseudomonas syringae* pathogens of *Arabidopsis* and a bacterial locus determining avirulence on both *Arabidopsis* and soybean. Plant Cell. 3(1): 49–59.

Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 35(suppl 2): W407-W410.

Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y (2015) The I-TASSER Suite: protein structure and function prediction. Nat Meth. 12(1): 7-8.