RESEARCH ARTICLE

Study of Molecular Interaction Between *Arabidopsis* Aquaporin and *Pseudomonas syringae* pv. *syringae* Harpin ($HrpZ_{Pss}$) Through Molecular Docking Tools

Kishori Lal, Vinay K. Singh, Debashish Dey*

DOI: 10.18811/ijpen.v9i04.06

Abstract

Harpins constitute one of the unique group of elicitor proteins secreted through the type 3 secretion system during plant-pathogen interactions and induce various responses like the formation of pores in the membrane, hypersensitive response, and systemic acquired resistance in non-host plants. Harpins from different bacterial sources elicit different responses in plants which may be due to their structural differences. As of today, the complete mechanisms of action of different harpins are lacking. The present study aimed to investigate the protein-protein-mediated functional association between members of the aquaporin PIP family from *Arabidopsis thaliana* At PIP1;3 and harpin ($HrpZ_{Pss}$) from *Pseudomonas syringae* pv. *syringae* through an *in-silico* molecular docking approach using ZDOCK. We identified a motif (HINPAVTFG) present in AtPIP1;3 as a functional signature sequence. Most of the residues from this signature sequence showed a positive and close interaction with the $HrpZ_{Pss}$ protein. The quality assessment and residues-wise fluctuation of the selected protein were analyzed with the help of molecular docking. The stability of predicted models of docked complexes (AtPIP1;3, $HrpZ_{Pss}$ and $HrpZ_{Pss}$-AtPIP1;3 complexes) were checked by molecular dynamics simulations using WEBGRO. The RMSD values of AtPIP1;3, $HrpZ_{Pss,}$ and $HrpZ_{Pss}$-AtPIP1;3 complexes were calculated at 50 ns. The stable conformation of the AtPIP1;3 and $HrpZ_{Pss}$ were observed at 14.1 and 13.2 ns, respectively, while the stability of the $HrpZ_{Pss}$-AtPIP1;3 complex was achieved at 20 ns. The present *in-silico* analysis demonstrates a positive interaction between harpin ($HrpZ_{Pss}$) and aquaporin (AtPIP1;3) which would help in understanding the harpin-induced signaling responses in non-host plants.

Keywords: *AtPIP1;3*, $HrpZ_{Pss}$, *In-silico*, Molecular dynamics simulations, Protein-protein interactions, Signature sequence.

Highlights

- Structure prediction and analysis of harpin effector protein from *Pseudomonas* and *Arabidopsis* aquaporin protein using *in-silico* tools.
- The protein-protein interaction between the harpin effector protein and aquaporin receptor was detected using the ZDOCK docking server and prediction of interacting residues and binding hot spots.
- Stability of predicted models of docked complexes (AtPIP1;3, $HrpZ_{Pss}$ and $HrpZ_{Pss}$-AtPIP1;3 complexes) was checked by molecular dynamics simulations using the WEBGRO tool.
- We identified a motif (HINPAVTFG) present in AtPIP1;3 as a functional signature sequence that directly participates in the protein-protein interaction with the harpin effector protein.

International Journal of Plant and Environment (2023); **ISSN:** 2454-1117 (Print), 2455-202X (Online)

Introduction

Harpins are a unique group effector protein produced by several gram-negative plant-pathogenic bacteria that are transported directly into the host cytoplasm during plant-pathogen interactions (Wei *et al.,* 1992; He *et al.,* 1993; Choi *et al.,* 2013). Harpins are a part of the type III secretion system (T3SS) secretome, which plays crucial roles in host-pathogen interactions by stimulating or interfering with host cellular processes (Galan and Collmer 1999). Its unique physicochemical properties, such as being a heat-stable, glycine-rich protein with very few or no cysteine make them distinctive from other proteins (Wang and Zhang 2007). It induces the efflux of cations, resulting in extracellular media alkalization and the influx of calcium across the plasma membrane of cells, which is essential for the initiation of defense mechanisms in non-host plants.

Harpins have been reported from different phytopathogenic bacterial sources, namely HrpN and HrpW from *Erwinia amylovora*; PopA1 and PopW from *Ralstonia solanacearum*; HrpZ1 and its orthologs from *Pseudomonas syringae*; Hpa1 and its orthologs from *Xanthomonas* species (He *et al.,* 1993; Li *et al.,* 2011; Choi *et al.,* 2012; Madhuri *et al.,* 2012; Dey *et al.,* 2014; Fu *et al.,* 2014; Nadendla *et al.,* 2018; Yang *et al.,* 2022). Harpins are called multifunctional proteins and have been broadly classified into four groups based on their sequence similarity and their domain organization (Choi *et al.,* 2013). Although harpins have diverse functions, they elicit these effects on plants remains to be determined. Therefore, further detailed studies are required to have a clear understanding of the mechanism of action of the harpins in the plant's system.

Laboratory of Plant Biotechnology, School of Biotechnology, Banaras Hindu University, Varanasi, Uttar Pradesh, India

***Corresponding author:** Debashish Dey, Laboratory of Plant Biotechnology, School of Biotechnology, Banaras Hindu University, Varanasi, Uttar Pradesh, India, Email: deybiotech@gmail.com

How to cite this article: Lal, K., Singh, V.K., Dey, D. (2023). Study of Molecular Interaction Between *Arabidopsis* Aquaporin and *Pseudomonas syringae* pv. *syringae* Harpin ($HrpZ_{Pss}$) Through Molecular Docking Tools. International Journal of Plant and Environment. 9(4), 343-349.

Submitted: 07/10/2023 **Accepted:** 15/11/2023 **Published:** 28/12/2023

Harpins act like an immunity booster that may activate disease resistance signaling (Bauer *et al.,* 1995; Gaudriault *et al.,* 1998). Both the exogenous application of purified harpins by spraying on plants or transgenic expression of harpins are known to activate various beneficial responses in plants such as increased photosynthesis and plant growth promotion, induction of systemic acquired resistance (SAR), and defense responses against a wide variety of pathogens like bacteria, viruses, fungi, and insects. The foliar application of harpin (PopW) induced resistance in tobacco plants against the Tobacco mosaic virus (TMV) (Arlat *et al.,* 1994). Transgenic expression of harpins such as $HrpZ_{Pss}$ led to the development of male sterile plants (Li *et al.,* 2010) and disease-resistant plants in tobacco (Noel *et al.,* 2002). Similarly, plants expressing $Hpa1_{Xoo}$ from *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) developed disease resistance against both bacterial and fungal pathogens (Kim *et al.,* 2003) and insects (Kim *et al.,* 2004). Therefore, harpins act as multifunctional proteins, playing diverse roles such as modulating photosynthesis, plant growth promotion, increasing crop yield etc. besides playing a role in plant-pathogen interactions (Choi *et al.,* 2013).

Previously, it was reported that Harpin (Hpa1) from *Xoo*, a causal pathogen of bacterial blight in rice, interacts with the aquaporin protein (AtPIP1;4) of *Arabidopsis* and leads to enhanced photosynthesis rates as well as substrate transport (Li *et al.,* 2015). A study on rice demonstrated that Hpa1 from the T3 translocator of *Xoo* has helped in the delivery of effector proteins from bacterial cells into the plant cells' cytosol, by making interaction with plasma membrane-bound aquaporin (OsPIP1;3) protein from rice (Li *et al.,* 2019). Further *in-silico* study of the protein-protein interactions uncovered the interfaces of the Hpa1 protein and an aquaporin from a rice plant (OsPIP1;3) (Patolia *et al.,* 2023).

Harpin ($HrpZ_{Pss}$) protein, identified from *Pseudomonas syringae* pv. *syringae* is encoded by the *hrpZ* gene and belongs to the second group (HrpZ1 group) of harpin categorization (Choi *et al.,* 2013). It is an extracellular protein of ~34.7 kDa, which induces HR upon infiltration of harpin protein into the dorsal surface of non-host plants such as tobacco (Li *et al.,* 2011). Previous reports have shown that the $HrpZ_{Pss}$ protein forms ion-conducting pores in plant cell membranes, indicating that it would function on the plasma membrane of plants. (Lee *et al.,* 2001). Multiple reports have been published illustrating the multifunctional roles of harpins, but there is very limited information on the investigation of harpin-receptor interactions, so an approach has been made to investigate the protein-protein interactions to unravel the host-pathogen interactions.

Interaction of harpin with transmembrane protein leads to playing various roles in crosstalk between plant and pathogen. A major gap in these interaction studies is plant sensing linked to various cellular responses and sensors of plants for Harpin like $HrpZ_{Pss}$. This gap might be overcome by in-silico bioinformatics-based approaches by utilizing various computational tools. Based on the literature survey, there are very limited reports available on the interaction study between different groups of harpins and aquaporins. In the present study, we intend to find which residues of $HrpZ_{Pss}$ from *Pseudomonas syringae* pathogen functionally interact with the residues of AtPIP1;3 proteins. We used a computational approach to find out the interacting residues using KFC Server which are involved in $HrpZ_{Pss}$-AtPIP1;3 interaction, ZDOCK server 3.0.2 used to dock the $HrpZ_{Pss}$ and AtPIP1;3 proteins (Pierce *et al.,* 2011; Zhu *et al.,* 2011). Our results suggest that this interaction will be helpful in the future to understand the crosstalk between plants and pathogens and the various responses of plant cells against the pathogen effector molecules.

Materials and Methods

Structural Prediction and Analysis of *P. syringae* pv. *syringae* Harpin ($HrpZ_{Pss}$) Protein and *Arabidopsis thaliana* PIPs (AtPIP1;1, AtPIP1;2, AtPIP1;3, AtPIP1;4 and AtPIP1;5)

The amino acid sequences of $HrpZ_{Pss}$ (NCBI gene back ID: AAA25839.1) and AtPIP1;1 (TAIR ID: AT3G61430), AtPIP1;2 (TAIR ID: AT2G45960), AtPIP1;3 (TAIR ID: AT1G01620), AtPIP1;4 (TAIR ID: AT4G00430) and AtPIP1;5 (TAIR ID: AT4G23400) were retrieved from the NCBI and TAIR databases respectively for further studies. To predict the 3D structure of the proteins, the amino acid sequences of all the proteins were submitted to the LOMETS server (Zheng *et al.,* 2022). Finally, the predicted 3D models were refined by using the ModRefiner server. The structural analysis of the predicted proteins was done using ERRAT, VERIFY 3D, and PROCHECK tools (Bowie *et al.,* 1991; Colovos *et al.,* 1993; Laskowski *et al.,* 1996; Benkert *et al.,* 2011). Quality assessment was done based on structural verifications. The ProtParam tool of the Expasy server was used for the calculation of the physical and chemical properties of selected proteins. Solvent accessibility of the proteins was calculated using the RePROF predictor of the PredictProtein server.

Protein-Protein Interaction Studies Between Harpin ($HrpZ_{Pss}$) and AtPIP1;3

In order to investigate the protein-protein docking affinity and complex formation, a molecular docking study was performed on the widely used ZDOCK webserver. In the ZDOCK web server, docking was performed by submitting the entire amino acid sequences of both proteins, $HrpZ_{Pss}$ as well as AtPIP1;3. Docking was performed by rigid-based docking between selected proteins with novel pair-wise statistical energy potential. Knowledge-based FADE and Contacts server (KFC) was used to investigate the interacting residues and binding hot spots prediction within protein-protein interaction using binding contacts. The PDB complex structure of $HrpZ_{Pss}$-AtPIP1;3 was submitted to the KFC webserver (Zhu *et al.,* 2011). The interacting residues within the complex model ($HrpZ_{Pss}$-AtPIP1;3) were visualized by BIOVIA Discovery Studio 2021. The interacting residues of *Arabidopsis* PIP1;3 (AtPIP1;3) were compared with the functional signature identified through the PROSITE-Expasy server. The HawkDock Server was used to calculate the binding free energy of the protein-protein complex (Weng *et al.,* 2019). The Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) method is applied to predict the binding free energy of a docked complex. MM/GBSA is a fast and effective method to calculate the desolvation potentials based on solvent-accessible surfaces.

Molecular Dynamics Simulation of Harpin ($HrpZ_{Pss}$), AtPIP1;3 and Harpin-AtPIP1;3 Complexes

Molecular Dynamics Simulation (MDS) for each protein structure and protein-protein complexes were carried out using the WEBGRO Macromolecular Simulations server where MDS of ($HrpZ_{Pss}$, AtPIP1;3 and Harpin-AtPIP1;3 Complexes) were prepared at 50-ns in MD equilibration and run parameters. All the MD simulations were done using a protein in water simulation in the WEBGRO Macromolecular Simulations server. All the MD simulation systems were solvated using a forcefield GROMOS96 43a1 and SPC water model in a triclinic box. For neutralizing the system and providing the physiological conditions for simulation, salt-type NaCl ions were added as a solvent. The configurational space and the number of local minima were high, so energy minimization is required to eliminate the poor connection and steric conflicts, particularly simulations were subjected by approaching the steepest descent route and setting energy minima to 5000 steps. Furthermore, systematic moves down the equilibration and MD were done by equilibration type NVT/NPT, and the entire equilibration was done at a simulation time of 50 ns with position restraint. WEBGRO server was used for simulation tools for the MD study. Further root mean square deviation (RMSD) and fluctuations during the entire process were analyzed based on obtained values of root mean square fluctuation (RMSF) besides computing the Radius of Gyration (Rg) and hydrogen bonds (H-bonds).

Results and Discussions

Analysis of Predicted Protein Structure

The three-dimensional (3D) structures of both proteins were predicted by the LOMETS server. Fig. 1 shows the 3D models of $HrpZ_{PSS}$ and AtPIP1;3 predicted in this study. The already available alpha fold structure ID of harpin ($HrpZ_{Pss}$) protein of the *Pseudomonas syringae* pv. *syringae* is AF-P35674-F1 whereas the available alpha fold structure ID of aquaporin (AtPIP1;3) of the *A. thaliana* is AF-P61837-F1.

The determination of physicochemical properties of both the proteins "harpin ($HrpZ_{Pss}$) and aquaporin (AtPIP1;3) were done by comparative analysis of the amino acid profile of both the proteins and is shown in Fig. 2. The comparative study revealed that the major part, approximately 47% of the harpin protein is made up of amino acid residues such as alanine, glycine, leucine and serine, whereas there is complete lack of some amino acid residues or presence of very few residues such as cystine, tyrosine, histidine and tryptophan. In the case of aquaporin, the major part, approximately 44% of the protein is formed of alanine, glycine, valine, leucine, and isoleucine amino acids.

The model assessments of the predicted structures were done by using the QMEAN 4.3.1 of the SWISS-MODEL server. The QMEAN 4.3.1 server was used to estimate the global quality and local quality estimations. The quality of the predicted model based on the QMEAN score is shown in Fig.3. The local quality estimate is shown in Fig. 3 (a and d) where the x-axis shows the residue number, while the y-axis reflects the expected similarity to the native structure. Fig. 3 (b and e) shows the quality of the structural models of $HrpZ_{Pss}$ and AtPIP1;3 respectively. Fig. 3 (c and f) shows the plot derived from the normalized QMEAN score (y-axis) against the scores (x-axis) obtained from the non-redundant set of high-resolution PDB structures.

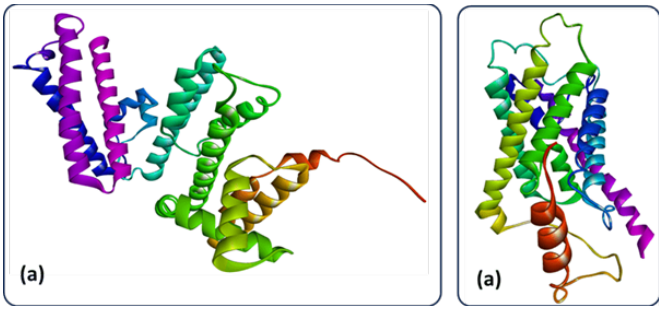

Fig. 1: Three-Dimentional structural models of the HrpZ and AtPIP1;3 predicted through LOMETS server. (a) Predcited 3D models of $HrpZ_{Pss}$ protein having major part of the protien forming alpha helix. (b) Predicted 3D models of AtPIP1;3.

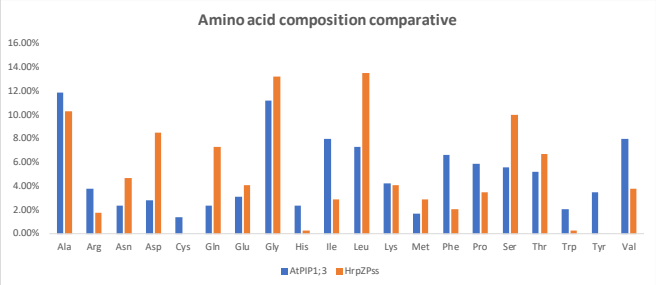

Fig. 2: Plot of physicochemical properties of both the proteins "harpin ($HrpZ_{Pss}$) and aquaporin (AtPIP1;3)". Comparative analysis of the amino acid profile of both the proteins were shown and indicated with different color (blue = aquaporin (AtPIP1;3) and orange = harpin ($HrpZ_{Pss}$)).

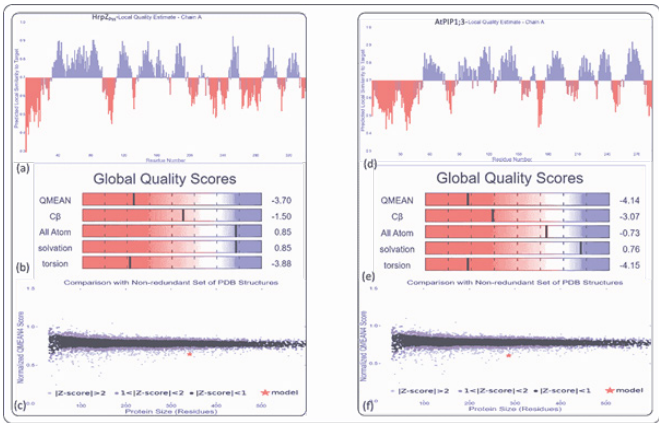

Fig. 3: Validation of the *in silico* predicted model of $HrpZ_{Pss}$ and AtPIP1;3 by QMEAN webserver. The estimation of Z-scores, an indication of $HrpZ_{Pss}$ model quality as a global score against a set of PDB structures based on QMEAN scoring function. a) Structural verification by QMEAN resulted -3.70 QMEAN score (red star) for the predicted harpin model ($HrpZ_{Pss}$). b) Harpin Global Quality Score ($HrpZ_{Pss}$) represented in gray and black points (c) Comparison of Harpin ($HrpZ_{Pss}$) with a non-redundant set of high-resolution experimental structures. d) Estimation of the AtPIP1;3 model quality based QMEAN score as a global score (Z-score). e) Global quality factor for AtPIP1;3 proteins with QMEAN score (-4.14) (d) (red star) (f) for prediction model of AtPIP1;3 and AtPIP1;3 comparison set of PDB structures (gray and black points).

Energy Minimization and Validation

The SAVES server was used for the evaluation of the stereochemical quality of the structures of $HrpZ_{Pss}$ and AtPIP1;3 and further evaluated and validated based on the analysis of the Ramachandran plots with the aid of Ψ/ϕ angles. After analysis of the Ramachandran plot for both the 3D models $HrpZ_{Pss}$ and AtPIP1;3 showed that 92.2 % and 87.6 % residues resided in the allowed regions, respectively. While the 1.1 % residues in $HrpZ_{Pss}$ were present in disallowed regions as well as 0.9 % residues of AtPIP1;3 was present in the disallowed regions where it signified the reliability of the predicted models. Ramachandran plot details for both proteins are provided in Fig. 4.

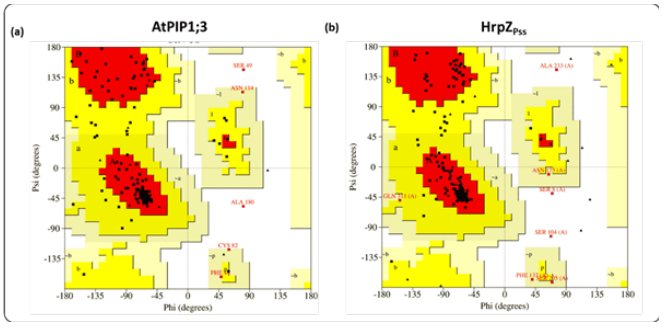

Fig. 4: Ramachandran plots of AtPIP1;3 (a) and $HrpZ_{Pss}$ (b) proteins.

Protein-Protein Docking

The classical method of studying protein-protein interaction necessitates the availability of existing three-dimensional structures of proteins under study. Protein-protein docking is the primary condition to study an enzymatic reaction or protein molecular affinity. The protein-protein interaction study was done using the ZDOCK server and a functional signature sequence (HINPAVTFG) of the AtPIP1;3 was identified through the PROSITE Expasy-server which was involved in the interaction. In this study, we found the interacting residues of *Arabidopsis* PIP1;3 (AtPIP1;3) as Ser^{46}, Ser^{47}, Trp^{48}, Trp^{51}, Arg^{52}, Gly^{54}, Ile^{55}, Ala^{56}, Phe^{58}, Ile^{59}, Phe^{62}, Leu^{63}, Leu^{65} Tyr^{66}, Leu^{70}, Met^{73}, Pro^{115}, Thr^{118}, Phe^{119}, Gly^{120}, Phe^{122}, Leu^{123}, Leu^{129}, Ala^{132}, Leu^{133}, Tyr^{134}, Tyr^{135}, Ile^{136}, Val^{137}, Met^{138}, Gln^{139}, Cys^{140}, Leu^{141}, Gly^{142}, Ala^{143}, Ile^{144}, Cys^{145}, Gly^{146}, Ala^{147}, Gly^{148}, Val^{149}, Val^{150}, Gly^{152}, Phe^{153}, Pro^{236}, Ala^{237}, Arg^{238}, Ser^{239}, Leu^{240}, Gly^{241}, Ala^{242}, Ile^{244}, Trp^{252}, His^{255}, Trp^{256}, Phe^{258}, Trp^{259}, Val^{260}, Phe^{263}, Ile^{264}, Ala^{267} involved in interaction with residues of harpin ($HrpZ_{Pss}$) protein of the pathogen *Pseudomonas syringae* pv. *syringae* (Fig. 5, 6) and residues were Met^{1}, Gln^{2}, Ser^{3}, Leu^{4}, Ser^{5}, Leu^{6}, Asn^{7}, Ser^{8}, Ser^{9}, Ser^{10}, Leu^{11}, Gln^{12}, Thr^{13}, Pro^{14}, Ala^{15}, Met^{16}, Val^{41}, Ala^{44}, Glu^{45}, Glu^{46}, Leu^{47}, Met^{48}, Arg^{49}, Asn^{50}, Gly^{51}, Gln^{52}, Leu^{53}, Asp^{54}, Asp^{55}, Ala^{73}, Gly^{74}, Gly^{75}, Gly^{76}, Ile^{77}, Phe^{137}, Glu^{139}, Asp^{140}, Asp^{141}, Met^{142}, Pro^{143}, Met^{144}, Asn^{146}, Lys^{147}. The interacting residues of *A. thaliana* PIP1;3 were analyzed and it was found that the HINPAVTFG motif is participating in the interaction and a major part of the signature sequence of AtPIP1;3 showed an interaction with the $HrpZ_{Pss}$ protein residues (Table 1 and 2). The HawkDock Server was used for determining the binding free energy of the protein-protein complex. The Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) method is a fast and effective method to calculate the binding free energy. The free energy decomposition analysis using the MM/GBSA method yielded the final estimated binding free energy of the protein-protein complex which is -80.79 (kcal/mol).

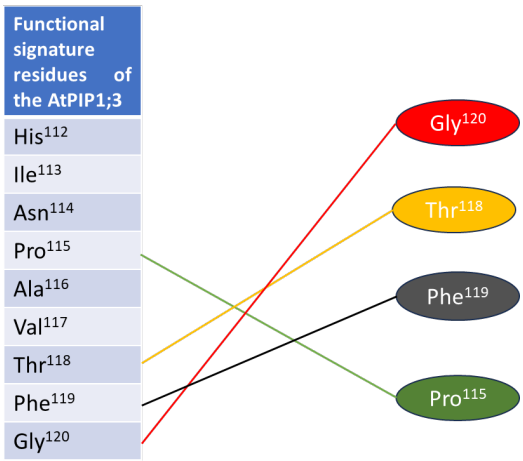

Fig. 5: The major residues (Pro^{115}, Thr^{118}, Phe^{119} and Gly^{120}) from the functional signature sequences (His^{112}, Ile^{113}, Asn^{114}, Pro^{115}, Ala^{116}, Val^{117}, Thr^{118}, Phe^{119}, and Gly^{120}) of *A. thaliana* aquaporin (AtPIP1;3) involved in the protein-protein interaction with harpin ($HrpZ_{Pss}$) protein.

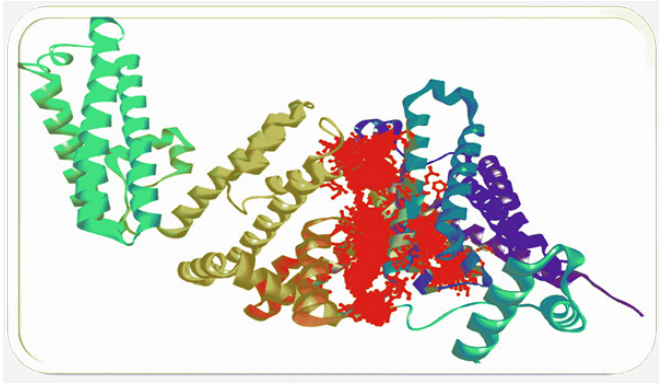

Fig. 6: Structural overviews of AtPIP1;3-Harpin ($HrpZ_{Pss}$) interacting complex using KFC server. The interacting residues were predicted through the Knowledge-based FADE and Contacts (KFC) server in AtPIP1;3-$HrpZ_{Pss}$ complex by recognizing structural features indicative of important binding contacts. The interacting regions of both the protein were highlighted in red color at the center in AtPIP1;3-$HrpZ_{Pss}$ complex.

Simulation of AtPIP1;3 and Harpin ($HrpZ_{Pss}$); and AtPIP1;3-Harpin Complex Using Molecular Dynamics

The MD simulations of the AtPIP1;3, $HrpZ_{Pss}$, and $HrpZ_{Pss}$-AtPIP1;3 complexes were carried out to detect the stability and conformational changes at 50 ns using WEBGRO Macromolecular Simulations server. To decipher the molecular interaction of the $HrpZ_{Pss}$-AtPIP1;3 complex, the MD simulation of the post-docking complex has been evaluated. The values of RMSD, RMSF and H-bonds of the complexes were compared with values of same for $HrpZ_{Pss}$, AtPIP1;3 and $HrpZ_{Pss}$ in complexes with AtPIP1;3.

Stability Analysis of AtPIP1;3, Harpin ($HrpZ_{Pss}$), and AtPIP1;3-Harpin ($HrpZ_{Pss}$) Docked Complexes

The stability of $HrpZ_{Pss}$ and AtPIP1;3 proteins in the individual form and in complex form of $HrpZ_{Pss}$-AtPIP1;3 was evaluated followed by analysis and calculation of the RMSD values for the C backbone of simulations performed at 50 ns. In the simulation process, we

Table 1: Analysis of interacting residues and binding hot spots of AtPIP1;3 using KFC webserver

Protein model	*Residues involved in interaction*
AtPIP1;3	Ser^{46}, Ser^{47}, Trp^{48}, Trp^{51}, Arg^{52}, Gly^{54}, Ile^{55}, Ala^{56}, Phe^{58}, Ile^{59}, Phe^{62}, Leu^{63}, Leu^{65} Tyr^{66}, Leu^{70}, Met^{73}, Pro^{115}, Thr^{118}, Phe^{119}, Gly^{120}, Phe^{122}, Leu^{123}, Leu^{129}, Ala^{132}, Leu^{133}, Tyr^{134}, Tyr^{135}, Ile^{136}, Val^{137}, Met^{138}, Gln^{139}, Cys^{140}, Leu^{141}, Gly^{142}, Ala^{143}, Ile^{144}, Cys^{145}, Gly^{146}, Ala^{147}, Gly^{148}, Val^{149}, Val^{150}, Gly^{152}, Phe^{153}, Pro^{236}, Ala^{237}, Arg^{238}, Ser^{239}, Leu^{240}, Gly^{241}, Ala^{242}, Ile^{244}, Trp^{252}, His^{255}, Trp^{256}, Phe^{258}, Trp^{259}, Val^{260}, Phe^{263}, Ile^{264}, Ala^{267}

Table 2: Analysis of interacting residues and binding hot spots of $HrpZ_{Pss}$ using KFC webserver.

Protein model	*Residues involved in interaction*
$HrpZ_{Pss}$	Met^{1}, Gln^{2}, Ser^{3}, Leu^{4}, Ser^{5}, Leu^{6}, Asn^{7}, Ser^{8}, Ser^{9}, Ser^{10}, Leu^{11}, Gln^{12}, Thr^{13}, Pro^{14}, Ala^{15}, Met^{16}, Val^{41}, Ala^{44}, Glu^{45}, Glu^{46}, Leu^{47}, Met^{48}, Arg^{49}, Asn^{50}, Gly^{51}, Gln^{52}, Leu^{53}, Asp^{54}, Asp^{55}, Ala^{73}, Gly^{74}, Gly^{75}, Gly^{76}, Ile^{77}, Phe^{137}, Glu^{139}, Asp^{140}, Asp^{141}, Met^{142}, Pro^{143}, Met^{144}, Asn^{146}, Lys^{147}.

tried to derive the RMSD values over the whole duration of the simulation in order to detect any structural variance between the initial conformation of the protein to any changes in protein conformation over time. RMSD values are an indication of stability of the simulation process with a low RMSD value suggesting a more stable simulation. Values of RMSD versus time curve are shown in Fig. 7 for the AtPIP1;3, $HrpZ_{Pss}$ and $HrpZ_{Pss}$-AtPIP1;3 complexes by plotting time (ns) in the x-axis *vs* RMSD values (nm) in the y-axis. RMSD values were 0.63, 0.64, and 1.07 nm for AtPIP1;3, $HrpZ_{Pss}$ and $HrpZ_{Pss}$-AtPIP1;3 respectively. As per the RMSD analysis of the C backbone, the $HrpZ_{Pss}$-AtPIP1;3 complex displayed higher flexibility as compared to aquaporin (AtPIP1;3) protein and harpin ($HrpZ_{Pss}$) protein. The flexibility of aquaporin (AtPIP1;3) and harpin ($HrpZ_{Pss}$) proteins were observed to be almost similar.

Residue-wise Fluctuation Analysis of AtPIP1;3, Harpin ($HrpZ_{Pss}$), and AtPIP1;3-Harpin Complexes

During the entire 50 ns MD simulation, the RMSF plot was generated to analyze the residue-wise fluctuations during the simulation process (Fig. 8, 9). The calculation of the RMSF values of AtPIP1;3 was done at 50 ns and compared with RMSF values of the complex of AtPIP1;3-$HrpZ_{Pss}$ and led to the identification of the flexible residues in presence and absence of $HrpZ_{Pss}$ interaction. Remarkably, the residue from 104-108 (Cys^{104}, Thr^{105}, Ala^{106}, Gly^{107}, and Ile^{108}), 151-159 (Lys^{151}, Gly^{152}, Phe^{153}, Gln^{154}, Pro^{155}, Asn^{156}, Pro^{157}, Tyr^{158}, and Gln^{159}), and 207-210 (Val^{207}, Pro^{208}, Leu^{209} and Ile^{210}) showed fluctuations in AtPIP1;3 unbounded form. Whereas the residues from 124-129 (Ala^{124}, Arg^{125}, Leu^{126}, Lys^{127}, Ser^{128}, and Leu^{129}) and 168-175 (Val^{168}, Ala^{169}, His^{170}, Gly^{171}, Tyr^{172}, Thr^{173}, Lys^{174}, and Gly^{175}) showed fluctuations in AtPIP1;3-$HrpZ_{Pss}$ complex forms. The fluctuated residues in Fig. 8,9 are marked with rectangular boxes. Similarly, the flexibility of $HrpZ_{Pss}$ was identified individually as well as when present in the complex with AtPIP1;3 (Fig. 9, region marked as rectangular). The average RMSF values were 0.27 and 0.28 nm for the alone AtPIP1;3, and for AtPIP1;3-$HrpZ_{Pss}$ complex, respectively. However, in the case of $HrpZ_{Pss}$ the average RMSF values of alone and in complex with AtPIP1;3 were 0.36 and 0.30, respectively. The residue-wise fluctuations in AtPIP1;3 were almost similar in both individual and complex (AtPIP1;3-$HrpZ_{Pss}$) forms. Unlike AtPIP1;3, $HrpZ_{Pss}$ showed higher values of RMSF fluctuations when present in the complex with AtPIP1;3, compared to $HrpZ_{Pss}$ alone. The following figures of RMSF revealed that the AtPIP1;3 exhibited a similar pattern of residue-wide fluctuations in both

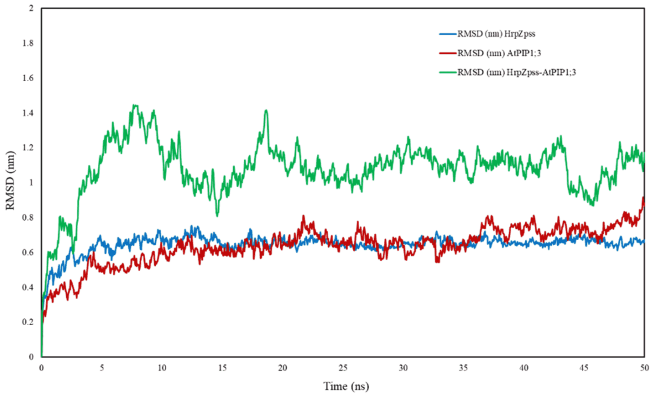

Fig. 7: Molecular dynamics simulation of AtPIP1;3, Harpin ($HrpZ_{Pss}$) proteins and AtPIP1;3-Harpin ($HrpZ_{Pss}$) bound complexes, computed on residue-wise RMSD deviations (nm). RMSD deviation plot of AtPIP1;3 and Harpin ($HrpZ_{Pss}$) Cα backbone atoms.

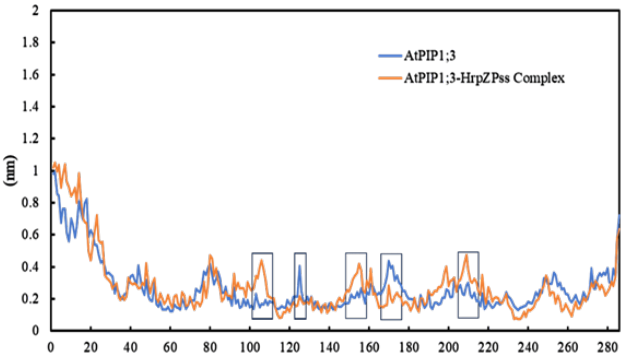

Fig. 8: Molecular Dynamics Simulations graph of individual AtPIP1;3 protein and in AtPIP1;3-$HrpZ_{Pss}$ complexes. RMSF deviation (nm) plot of AtPIP1;3 (blue color). RMSF deviation (nm) plot of AtPIP1;3-$HrpZ_{Pss}$ complexes (brown color).

forms, while the $HrpZ_{Pss}$ showed more fluctuations in AtPIP1;3 bounded forms. Remarkably, the residue from 264 to 283 (Ala^{264}, Gly^{265}, Gly^{266}, Gly^{267}, Leu^{268}, Gly^{269}, Thr^{270}, Pro^{271}, Val^{272}, Asn^{273}, Thr^{274}, Pro^{275}, Gln^{276}, Thr^{277}, Gly278, Thr279, Ser^{280}, Ala^{281}, Asn^{282} and Gly^{283}) and 271 to 283 (Pro^{271}, Val^{272}, Asn^{273}, Thr^{274}, Pro^{275}, Gln^{276}, Thr^{277}, Gly278, Thr279, Ser^{280}, Ala^{281}, Asn^{282} and Gly^{283}) in $HrpZ_{Pss}$ alone and bounded forms exhibited fluctuations with higher RMSF values (Fig. 9, region marked as rectangular).

$HrpZ_{Pss}$, secreted by *Psuedomonas syringae* pv. *syringae* is well

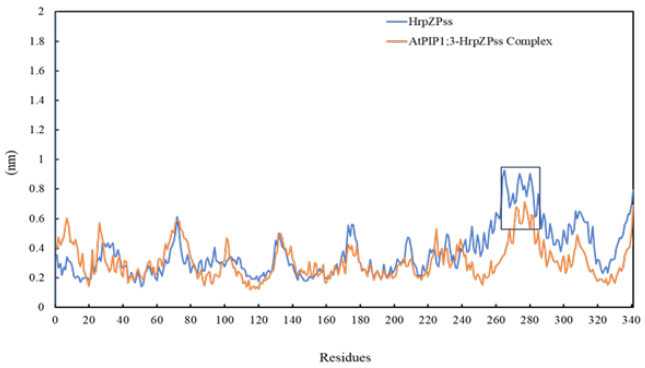

Fig. 9: Molecular Dynamics Simulations of individual $HrpZ_{Pss}$ protein and docked complex of AtPIP1;3-$HrpZ_{Pss}$. RMSF deviation (nm) plot of $HrpZ_{Pss}$ (blue color). RMSF deviation (nm) plot of AtPIP1;3-$HrpZ_{Pss}$ complexes (brown color).

known for the induction of defense signaling in non-host plants. AtPIP1;3 is a member of the plasma membrane intrinsic protein subfamily of *A. thaliana* PIP1. It is reported that harpin interacts with aquaporin proteins present on the membrane of *Arabidopsis* plants and initiates downstream harpin responses (Li *et al.,* 2015). To understand the exact mechanism of hairpin-aquaporin interaction, it is important to have an in-depth understanding of the type of molecular interaction ensuing between the interface of effector and receptor molecules. There are no experimentally determined crystal molecular structures available for any of the two proteins (AtPIP1;3 and $HrpZ_{Pss}$). Hence in this study, we used computational methods and algorithms to develop models for both the proteins and further to understand their molecular interactions. We investigated the protein-protein interaction of all five AtPIPs (AtPIP1;1, AtPIP1;2, AtPIP1;3, AtPIP1;4 and AtPIP1;5) with $HrpZ_{Pss}$ protein. Out of the five AtPIPs, we found a functional signature sequence motif (His^{112}, Ile^{113}, Asn^{114}, Pro^{115}, Ala^{116}, Val^{117}, Thr^{118}, Phe^{119}, and Gly^{120}) present only in AtPIP1;3 protein. Out of the nine residues of the functional signature sequence of AtPIP1;3, the major four residues (Pro^{115}, Thr^{118}, Phe^{119}, and Gly^{120}) were involved in the protein-protein interaction with $HrpZ_{Pss}$ protein. Whereas the interacted residues of *Arabidopsis* aquaporin (AtPIP1;3) such as Ser^{46}, Ser^{47}, Trp^{48}, Trp^{51}, Arg^{52}, Gly^{54}, Ile^{55}, Ala^{56}, Phe^{58}, Ile^{59}, Phe^{62}, Leu^{63}, Leu^{65} Tyr^{66}, Leu^{70}, Met^{73}, Pro^{115}, Thr^{118}, Phe^{119}, Gly^{120}, Phe^{122}, Leu^{123}, Leu^{129}, Ala^{132}, Leu^{133}, Tyr^{134}, Tyr^{135}, Ile^{136}, Val^{137}, Met^{138}, Gln^{139}, Cys^{140}, Leu^{141}, Gly^{142}, Ala^{143}, Ile^{144}, Cys^{145}, Gly^{146}, Ala^{147}, Gly^{148}, Val^{149}, Val^{150}, Gly^{152}, Phe^{153}, Pro^{236}, Ala^{237}, Arg^{238}, Ser^{239}, Leu^{240}, Gly^{241}, Ala^{242}, Ile^{244}, Trp^{252}, His^{255}, Trp^{256}, Phe^{258}, Trp^{259}, Val^{260}, Phe^{263}, Ile^{264}, Ala^{267} residues involved (enlisted in Table 1) in interaction with residues of harpin ($HrpZ_{Pss}$) protein of the pathogen *Pseudomonas syringae pv. syringae* and residues were Met^{1}, Gln^{2}, Ser^{3}, Leu^{4}, Ser^{5}, Leu^{6}, Asn^{7}, Ser^{8}, Ser^{9}, Ser^{10}, Leu^{11}, Gln^{12}, Thr^{13}, Pro^{14}, Ala^{15}, Met^{16}, Val^{41}, Ala^{44}, Glu^{45}, Glu^{46}, Leu^{47}, Met^{48}, Arg^{49}, Asn^{50}, Gly^{51}, Gln^{52}, Leu^{53}, Asp^{54}, Asp^{55}, Ala^{73}, Gly^{74}, Gly^{75}, Gly^{76}, Ile^{77}, Phe^{137}, Glu^{139}, Asp^{140}, Asp^{141}, Met^{142}, Pro^{143}, Met^{144}, Asn^{146}, Lys^{147}, (enlisted in Table 2). Therefore, further *in-silico* studies were done with AtPIP1;3 and $HrpZ_{Pss}$. The results of the docking studies to predict an interaction between AtPIP1;3 and $HrpZ_{Pss}$ by the *in-silico* tools with high accuracy suggest that this approach is quite helpful in studying the protein-protein interaction.

Previous reports highlighted the *in-silico* as well as *in-vivo* interaction of harpins from different microorganisms but there are no reports that showed the interacting residues of AtPIP1;3 protein with the residues of Harpin ($HrpZ_{Pss}$) from *Pseudomonas syringae* pv. *syringae* which belongs to the second group (HrpZ1 group) of harpin categorization (Choi *et al.,* 2013). This study found a motif that participates in the interaction and is a major part of the signature sequence of PIP1;3 protein from *Arabidopsis* plant. Docking studies revealed that the residues of the signature sequence of AtPIP1;3 were also involved in this protein-protein interaction (AtPIP1;3-$HrpZ_{Pss}$).

In the current study, the structural and functional relationships of AtPIP1;3 were analyzed using computational approaches including protein modeling, protein-protein molecular docking as well as molecular dynamics simulation studies. Structural models of both $HrpZ_{Pss}$ and AtPIP1;3 proteins were predicted using the LOMETS server and validated using ZDOCK docking and KFC webserver which revealed the amino acid residues involved in the molecular interaction at the interface of these two proteins. Subsequently, Molecular Dynamics simulations of $HrpZ_{Pss}$, AtPIP1;3, and $HrpZ_{Pss}$-AtPIP1;3 complex were also performed to check the stability of protein as an individual and in the complex form.

Previous *in-vitro* and *in-vivo* studies reported that Hpa1 interacts with PIP1;4 (AtPIP1;3) protein of *A. thaliana* and Hpa1 also showed the interaction with OsPIP1;3 protein of rice. Some *in vivo* studies also reported that Hpa1 interaction with the plant aquaporins leads to the activation of multiple hormones-associated pathways and the induction of various defenses-related signaling genes (Ji *et al.,* 2020). While the complete mechanisms of action of harpins is yet to be revealed.

Conclusion

Based on the current *in silico* study, we report that a total of 43 amino acid residues of the $HrpZ_{Pss}$ effector protein from *Pseudomonas syringae* pv. *syringae* are involved in the molecular interaction with the 61 amino acid residues of AtPIP1;3 receptor protein present on the plasma membrane of the *A. thaliana* plant. We demonstrate the importance of 61 amino acids of the AtPIP1;3 protein which includes the 9 amino acids functional signature sequence of aquaporin protein which directly participates in the protein-protein interaction with harpin effector protein ($HrpZ_{Pss}$). Apart from improving our present understanding of the crosstalk between $HrpZ_{Pss}$ and AtPIP1;3, a pathogen and host protein, respectively, it is also a demonstration of vigorous use of bioinformatics tools to detect and predict the protein-protein interaction through *in-silico* studies when the information about the crystal structure of the interacting proteins are lacking.

Acknowledgements and Funding Support

DD acknowledges the financial support in the form of 'UGC-BSR-Startup Research Grant' from UGC, Govt. of India, and 'Seed Grant' from Banaras Hindu University (BHU), Varanasi under the Institute of Eminence (IoE) scheme of Govt. of India. The financial support from the Department of Biotechnology (DBT), Govt. of India to the School of Biotechnology, BHU, and to KL in the form

of a DBT-Junior Research Fellowship (JRF) is also acknowledged. DD also acknowledges the Bioinformatics support from the BTIS-Net-Sub-Distributed Information Centre at the School of Biotechnology, Banaras Hindu University, Varanasi, India.

Author Contributions

Debashish Dey: conceived the research work. Debashish Dey and Vinay Kumar Singh designed the experiments. Kishori Lal and Vinay Kumar Singh: performed the experiments. Debashish Dey acquired the funding. All three authors contributed in writing and approving the paper draft for publication.

Conflict Of Interest

None

References

Arlat, M., Van Gijsegem, F., Huet, J. C., Pernollet, J. C., & Boucher, C. A. (1994). PopA1, a protein which induces a hypersensitivity-like response on specific Petunia genotypes, is secreted via the Hrp pathway of Pseudomonas solanacearum. The EMBO journal, 13(3), 543-553.

Bauer, D. W., Wei, Z. M., Beer, S. V., & Collmer, A. (1995). Erwinia chrysanthemi Harpin~ E~ c~ h: an elicitor of the hypersensitive response that contributes to soft-rot pathogenesis. Molecular Plant Microbe Interactions, 8, 484-484.

Benkert, P., Biasini, M., & Schwede, T. (2011). Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics, 27(3), 343-350.

Bowie, J. U., Lüthy, R., & Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three-dimensional structure. Science, 253(5016), 164-170.

Choi, M. S., Heu, S., Paek, N. C., Koh, H. J., Lee, J. S., & Oh, C. S. (2012). Expression of hpa1 gene encoding a bacterial harpin protein in Xanthomonas oryzae pv. oryzae enhances disease resistance to both fungal and bacterial pathogens in rice and Arabidopsis. The Plant Pathology Journal, 28(4), 364-372.

Choi, M. S., Kim, W., Lee, C., & Oh, C. S. (2013). Harpins, multifunctional proteins secreted by gram-negative plant-pathogenic bacteria. Molecular plant-microbe interactions, 26(10), 1115-1122.

Colovos, C., & Yeates, T. O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. Protein science, 2(9), 1511-1519.

Dey, D., Uma, B., Padmaja, G., & Podile, A. R. (2014). Pathogen-induced expression of harpinpss increases resistance in tobacco against Fusarium oxysporum f. sp. nicotianae. Journal of Plant Pathology, 96(2), 335-342.

Fu, M., Xu, M., Zhou, T., Wang, D., Tian, S., Han, L., ... & Zhang, C. (2014). Transgenic expression of a functional fragment of harpin protein Hpa1 in wheat induces the phloem-based defence against English grain aphid. Journal of experimental botany, *65*(6), 1439-1453.

Galán, J. E., & Collmer, A. (1999). Type III secretion machines: bacterial devices for protein delivery into host cells. Science, 284(5418), 1322-1328.

Gaudriault, S., Brisset, M. N., & Barny, M. A. (1998). HrpW of Erwinia amylovora, a new Hrp-secreted protein. FEBS letters, *428*(3), 224-228.

He, S. Y., Huang, H. C., & 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.

Ji, Z. L., Yu, M. H., Ding, Y. Y., Li, J., Zhu, F., He, J. X., & Yang, L. N. (2020). Coiled-coil n21 of hpa1 in Xanthomonas oryzae pv. Oryzae promotes plant growth, disease resistance and drought tolerance in non-hosts via eliciting hr and regulation of multiple defense response genes. International Journal of Molecular Sciences, 22(1), 203.

Kim, J. G., Park, B. K., Yoo, C. H., Jeon, E., Oh, J., & Hwang, I. (2003). Characterization of the Xanthomonas axonopodis pv. glycines Hrp pathogenicity island. Journal of bacteriology, 185(10), 3155-3166.

Kim, J. G., Jeon, E., Oh, J., Moon, J. S., & Hwang, I. (2004). Mutational analysis of Xanthomonas harpin HpaG identifies a key functional region that elicits the hypersensitive response in nonhost plants. Journal of bacteriology, 186(18), 6239-6247.

Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., & Thornton, J. M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. Journal of biomolecular NMR, 8, 477-486.

Lee, J., Klessig, D. F., & Nürnberger, T. (2001). A harpin binding site in tobacco plasma membranes mediates activation of the pathogenesis-related gene HIN1 independent of extracellular calcium but dependent on mitogen-activated protein kinase activity. The Plant Cell, 13(5), 1079-1093.

Li, J. G., Cao, J., Sun, F. F., Niu, D. D., Yan, F., Liu, H. X., & Guo, J. H. (2011). Control of Tobacco mosaic virus by PopW as a result of induced resistance in tobacco under greenhouse and field conditions. Phytopathology, 101(10), 1202-1208.

LI, J. G., LIU, H. X., Cao, J., CHEN, L. F., Gu, C., Allen, C., & GUO, J. H. (2010). PopW of Ralstonia solanacearum, a new two-domain harpin targeting the plant cell wall. Molecular plant pathology, 11(3), 371-381.

Li, L., Wang, H., Gago, J., Cui, H., Qian, Z., Kodama, N., ... & Dong, H. (2015). Harpin Hpa1 interacts with aquaporin PIP1; 4 to promote the substrate transport and photosynthesis in Arabidopsis. Scientific Reports, 5(1), 17207.

Li, P., Zhang, L., Mo, X., Ji, H., Bian, H., Hu, Y., ... & Dong, H. (2019). Rice aquaporin PIP1; 3 and harpin Hpa1 of bacterial blight pathogen cooperate in a type III effector translocation. Journal of experimental botany, 70(12), 3057-3073.

Madhuri, B., Raut, S., Dey, D., Nazneen, A., Uma, B., & Podile, A. R. (2012). Tapetum-specific expression of harpin Pss causes male sterility in transgenic tobacco. Biologia plantarum, 56, 628-634.

Nadendla, S. R., Rani, T. S., Vaikuntapu, P. R., Maddu, R. R., & Podile, A. R. (2018). HarpinPss encapsulation in chitosan nanoparticles for improved bioavailability and disease resistance in tomato. Carbohydrate polymers, 199, 11-19.

Noël, L., Thieme, F., Nennstiel, D., & Bonas, U. (2002). Two novel type III-secreted proteins of Xanthomonas campestris pv. vesicatoria are encoded within the hrp pathogenicity island. Journal of Bacteriology, 184(5), 1340-1348.

Patoliya, J., Thaker, K., Rabadiya, K., Patel, D., Jain, N. K., & Joshi, R. (2023). Uncovering the Interaction Interface Between Harpin (Hpa1) and Rice Aquaporin (OsPIP1; 3) Through Protein–Protein Docking: An In Silico Approach. Molecular Biotechnology, 1-13.

Pierce, B. G., Hourai, Y., & Weng, Z. (2011). Accelerating protein docking in ZDOCK using an advanced 3D convolution library. PloS one, 6(9), e24657.

Weng, G., Wang, E., Wang, Z., Liu, H., Zhu, F., Li, D., & Hou, T. (2019). HawkDock: a web server to predict and analyze the protein–protein complex based on computational docking and MM/GBSA. Nucleic acids research, 47(W1), W322-W330.

Wang, X., Li, M., Zhang, J., Zhang, Y., Zhang, G., & Wang, J. (2007). Identification of a key functional region in harpins from Xanthomonas that suppresses protein aggregation and mediates harpin expression in E. coli. Molecular biology reports, 34, 189-198.

Wei, Z. M., Laby, R. J., Zumoff, C. H., Bauer, D. W., He, S. Y., Collmer, A., & Beer, S. V. (1992). Harpin, elicitor of the hypersensitive response produced by the plant pathogen Erwinia amylovora. Science,257(5066), 85-88.

Yang, B., Yang, S., Zheng, W., & Wang, Y. (2022). Plant immunity inducers: From discovery to agricultural application. Stress Biology, 2(1), 5.

Zheng, W., Wuyun, Q., Zhou, X., Li, Y., Freddolino, P. L., & Zhang, Y. (2022). LOMETS3: Integrating deep learning and profile alignment for advanced protein template recognition and function annotation. Nucleic acids research, 50(W1), W454-W464.

Zhu, X., & Mitchell, J. C. (2011). KFC2: a knowledge-based hot spot prediction method based on interface solvation, atomic density, and plasticity features. Proteins: Structure, Function, and Bioinformatics, 79(9), 2671-2683.