RHIZOBITOXINE ENHANCES NODULATION BY INHIBITING ETHYLENE SYNTHESIS OF *BRADYRHIZOBIUM ELKANII* FROM *LESPEDEZA* SPECIES: VALIDATION BY HOMOLOGY MODELING AND MOLECULAR DOCKING STUDY

Ramachandran Vijayan\(^1\), Pitchai Palaniappan\(^2\), Tongmin SA\(^2\), and Padmanaban Elavarasi\(^3\), and Natesan Manoharan*\(^*\)

\(^1\)Department of Marine Science, Bharathidasan University, Tiruchirapalli 620 024, India.
\(^2\)Department of Agricultural Chemistry, Chungbuk National University, Cheongju, Chungbuk 361-763, South Korea.
\(^3\)Department of Chemistry, Periyar University, Salem 636 011, India.

**ABSTRACT**

*Bradyrhizobium elkanii* produces rhizobitoxine, a plant toxin which causes chlorosis in soybeans. Recent studies has shown that rhizobitoxine plays an important role for establishing symbiosis between legumes and *Bradyrhizobium elkanii*. rhizobitoxine is an inhibitor of ethylene biosynthesis thereby enhances the nodulation. As there are no experimentally determined structure for the *B. elkanii* reported in the literature, Therefore, we have modeled the protein structure by homology modeling and also docked with potent inhibitor rhizobitoxine having the GOLD score of 67.21. This binding allowed us to predict that this rhizobitoxine might be capable of inhibiting *B. elkanii* and thus enhancing nodulation. A significant finding of this study provides new insight into the structure and functional features are helpful to design the site-directed mutagenesis experiments and to govern the rhizobial infection of legumes.

**Key words:** rhizobitoxine, protein modeling, *in silico* docking, symbiosis.
INTRODUCTION
Symbiotic and non-symbiotic bacteria have been isolated from the root nodules of a wide range of legumes [1,2,34]. These non-symbiotic bacteria are endophytes living inside nodules and did not cause visible damage to the plants. Some endophytic bacteria exert several beneficial effects on host plants, such as stimulation of plant growth[1], nitrogen fixation [5], and induction of resistance to plant pathogens [6]. These bacteria can be inoculated to promote plant growth and biocontrol [7]. Some of the nodule-forming legumes can grow in stressful environments, and among the legumes, there are many important crop plants such as soybean, peanut, and common bean. During the past 3 years, new legume nodulating bacteria of the Alphaproteobacteria, i.e., Methylobacterium [8], Devosia [9], Blastobacter [10], and Ochrobactrum [11] genera have been discovered. Legume species of the genus *Lespedeza* are annual or perennial wild plants, of which about 140 species are herbs or shrubs. Most of these species are native to eastern Asia and some are indigenous to the south-eastern United States. Most *Lespedeza* sp. are drought-enduring plants and are held in high esteem as foliage, green manure crops or honey resources and in preventing soil erosion [12]. On some deforested hills in north-eastern China, *Lespedeza* sp. are the main pioneer plants after the trees have been cut; hence, they may be potential plants for reforestation. Previous studies on cross-nodulation indicated that *Lespedeza* plants belonged to the cowpea miscellany and showed that these plants have marked host specificity relative to the effectiveness of nodulation [13]. Because of the potential value of *Lespedeza* sp. in sustainable agriculture and reforestation, we studied previously the diversity and nodulation efficiency of strains isolated from nodules of *Lespedeza* sp [14].

Rhizobitoxine [2-amino-4-(2-amino-3-hydropropoxy)-trans-but-3-enoic acid] [Figure 1] is an ethylene synthesis inhibitor that is produced by the legume symbiont *B. Elkanii* [15]. Because it induces foliar chlorosis of soybeans, rhizobitoxine has been regarded as a plant toxin [16,17, 18]. Enzymatic studies have revealed that rhizobitoxine inhibits beta-cystathionase in the methionine biosynthesis pathway [19,20,21]. Aminoethoxy-Vinyl-Glycine (AVG), is an structural analogue of rhizobitoxine, has been used as inhibitor in the enzymatic studies of ACC synthase [22]. Therefore, rhizobitoxine is expected to be a potent inhibitor of ACC synthase. Indeed, ethylene has been suggested to be a component of the signaling pathway controlling the rhizobial infection of legumes [23,24].

To our knowledge, the potential agricultural importance of *B. elkanii* from *Lespedeza* sp. has
not been modeled. In the present work, an attempt has been made to achieve a model of the tertiary structure of the *B. elkanii* from *Lespedeza sp.* using a homology modeling approach [25,26,27]. The model was further used for the identification of substrate binding site. Finally, docking analysis was carried out to reveal the binding mechanism of the *B. elkanii* with rhizobitoxine.

**MATERIALS AND METHODS**

**Homology Model construction**

The amino acid sequence of *B. elkanii* from *Lespedeza sp.* (AFM94307) was retrieved from the NCBI database (www.ncbi.nlm.nih.gov). Sequence database search was carried out with blastp tool in NCBI (www.ncbi.nlm.nih.gov/blast) to identify the homologue of known structure from the PDB for template proteins of 3VGI. The coordinates of 3VGI channel (PDB ID: 3GVI, 2.25Å resolution) were selected as a template to build the initial *B. elkanii* structure.

![Fig. 1: The chemical structure of rhizobitoxine.](image)

The sequence alignments were performed with the ClustalW program server [28,29] (http://www.ebi.ac.uk/Tools/clustalw2/index.html) [Figure 2]. The models were built with MODELLER 9v1 [30,31,32]. This program implements comparative protein structure modeling by satisfying spatial restraints in terms of probable density functions. The best quality of the refined model were submitted to an energy-minimization using Swiss-pdb viewer [33] in order to remove bad contacts derived from the homology modeling [30,31].
Manoharan et al. | World Journal of Pharmacy and Pharmaceutical Sciences

**Fig. 2:** Sequence alignment of template protein (3GVI) and target protein (*B. elkanii*).

Highly conserved residues between *B. elkanii* and 3GVI families are depicted in black*. 

The energy minimized structure were validated by a combination of these programs such as PROCHECK [34] ProSA-web (https://prosa.services.came.sbg.ac.at/prosa.php) and ERRAT version 2.0 (http://nihserver.mbi.ucla.edu/ERRATv2/). The quality of the model was assessed using PROCHECK to confirm the quality of the stereochemistry of the protein structure [35]. The ProSA and ERRAT were also implemented to evaluate the predicted structures. The ProSA is employed in the refinement and validation of experimental protein structures, structure prediction, and modeling [36]. ERRAT is a protein structure verification algorithm that is especially well-suited for evaluating the progress of crystallographic model building and refinement [37].

**Binding Site Prediction**

To predict the binding site, LIGSITEcsc algorithm [39] and CASTP algorithm [38] LIGSITE which uses connolly surface and defines surface solvent surface events. The algorithm proceeds as follows: the protein is projected onto a 3D grid with a step size of 1.0Å; grid points are labeled as protein, surface, or solvent using certain rules. A grid point is marked as protein if there is at least one atom within 1.6Å. After the solvent excluded surface is calculated the surface vertices coordinates are stored. A sequence of grid points, which starts and ends with surface grid points and which has solvent grid points in between, is called a surface solvent surface event. If the number of surface solvent surface events of a solvent grid exceeds a minimal threshold of 6, then this grid is marked as pocket. Finally, all pocket grid points are clustered according to their spatial proximity. The clusters are ranked by the
number of grid points in the cluster. The top three clusters are retained and their centers of mass are used to represent the predicted pocket sites [38]. And also the binding pockets of *B. elkanii* were identified using CASTP [38] (Computed Atlas of Surface topography of proteins); a program for identifying and characterizing protein active sites, binding sites, and functional residues located on protein surfaces and voids buried in the interior of proteins by measuring concave surface regions on three-dimensional structures of proteins. It also measures the area and volume of pocket or void by solvent accessible surface model (Richards' surface) and by molecular surface model (Connolly's surface). It can also be used to study surface features and functional regions of proteins. The binding pocket obtained by in silico studies on *B. rhizobium elkanii* was consistent with the known biochemical studies (http://www.ebi.ac.uk/pdbe srv/view/entry/3gvi/ligands.html).

**Structural Superposition of B.rhizobium with the template**

RMSD (c-alpha atom) of the modeled protein with respect to the template, 3GVI was calculated, using RASH Structural Superposition (http://sysimm.ifrec.osaka-u.ac.jp/rash/).

**Molecular Docking studies**

Rhizobiotoxine was extracted in the CHEMBANK database (http://chembank.broadinstitute.org/) in 2D-structure data file (SDF) format were converted into 3D-MOL2 file with the program OpenBABEL2.3.1[40] (Figure 1). The rhizobiotoxine were further energy minimized for 100 steps with Swiss-PDB viewer with steepest descent and conjugated gradient algorithms[39,40] GOLD uses a genetic algorithm to explore the full range of ligand conformational flexibility with partial flexibility of the protein[42]. The Docking was performed using GOLD Software (Genetic Optimization Ligand Docking). GOLD uses a genetic algorithm to explore the full range of ligand conformational flexibility with partial flexibility of the protein. Docking procedure consisted of three interrelated components; a) identification of binding site b) a search algorithm to effectively sample the search space (the set of possible ligand positions and conformations on the protein surface) and c) a scoring function. The GOLD fitness function consisted of four components: a) protein-ligand hydrogen bond energy (external H-bond); b) protein-ligand vanderwals(vdw) energy(external vdw); c) ligand internal vdw energy (internal vdw); d) ligand torsional strain energy(internal torsion). Default settings for GOLD docking were adopted. The modeled structure was a homo tetramer, so we chose randomly D chain of the predicted structure for docking. Ten poses were kept for each ligand, and the one yielding the best score was used.
for further analysis.

RESULTS AND DISCUSSION

Sequence alignment and model building

The target sequence was compared with the related family for more identity and similarity using BLAST search. In the results of BLAST search against PDB, the Crystal structure of Lactate/malate dehydrogenase from Brucella melitensis (PDB: 3GVI), has a high level of sequence similarity and the identity of B. elkanii protein with 88.6%. Sequence of the B. elkanii from Lespedeza sp. with the template was carefully analyzed. Secondary structure prediction[Figure 3] was made by Split server 4.0[41], which uses hydrogen bond energy and main chain dihedral angles to recognize helix, coils and strands.

![Fig. 3: Secondary structure prediction. Blue line: Beta preference. (BET index); Red line: Transmembrane helix preference. (THM index); Gray line: Modified hydrophobic moment index. (INDA index); Violet boxes (below abscissa): Predicted transmembrane helix position. (DIG index).](image)

The three dimensional structure of the protein was constructed using MODELLER8V2[30,31,32] as shown in Figure 4. The Python script was executed to calculate a model with all non-hydrogen atoms by satisfaction of spatial restraint. The model structure were inspected visually and any bumps introduced during the modeling process were removed by manual rotation of torsions. Explicit hydrogens were added to the protein and was subjected to energy minimization using Swiss-PDB viewer with steepest descent and conjugated gradient algorithms31. Energy minimization and relaxation of the loop regions was performed using 300 iterations in a simple minimization method. Again the steepest descent was carried out until the energy showed stability in the sequential repetition. The B. elkanii
model structure exhibits a homo tetramer and it belongs to malate dehydrogenase family.

![Ribbon representation of modeled structure](image1)

**Fig. 4:** Ribbon representation of our modeled structure *B. elkanii* from *Lespedeza sp.*

**Protein- model quality estimation**

The PROCHECK, ProSA-Web, and ERRAT programs were used to validate the *B. elkanii* structure. The geometry of the *B. elkanii* model was then evaluated with Ramachandran plot calculations computed with the PROCHECK program. This revealed that the backbone phi and psi dihedral angles of *B. elkanii* are 93.2%, 5.8%, 0.8% and 0.2% of the residues are located within the most favorable, additionally allowed, and generously allowed regions, respectively of the Ramachandran plot (Figure 5).

![Ramachandran plot](image2)

**Fig. 5:** Ramachandran plot of *B. elkanii* obtained by PROCHECK.
Ramachandran plot analysis revealed that the residues are found to be in the core regions of the Ramachandran plot predicting the *B. elkanii* structure to be highly reliable for further docking simulation studies. This good stereo-chemical quality is not surprising for the high sequence identity (88.6%) between the template and the target, which are shown in Figure 2. Totally, 99.1% of the residues were in favored and allowed regions. Hence, X-ray structure of 3GVI at 2.0Å resolution has also 0.8% of its residues in disallowed regions. Most bond lengths, bond angles and torsion angles were in the range of values expected for a naturally folded protein, it can be said that our *B. elkanii* structure satisfies criteria of a good model.

Further, we checked the overall model quality of *B. elkanii* model was calculated using ProSA-web [Figure 6a, 6b]. The proper Z-score values (within -6.05) were observed in the modeled structure compared with that of the template (-10). As the result, our predicted structure are within the range of scores typically found for native proteins of similar size which are experimentally determined protein chains in current PDB.34

![Fig. 6a: B. elkanii model was evaluated with ProSA and the score value -6.05.](image)

![Fig. 6b: B. elkanii model was validated with ProSA and compared the sequence with the ProSA energy.](image)
Furthermore, ERRAT is protein structure verification algorithm for non-bonded atomic interactions between different atom types. The recommended value of ERRAT is greater than 50 for higher quality model, and the higher scores indicate the better quality [37]. In the case of our predicted model, ERRAT value was 100 for *B. elkanii*, the model were found within the range of a high quality model. All the validation results revealed that our constructed model of *B. elkanii* were the extremely best quality (Figure 7).

![Error value graph](image)

**Fig. 7:** *B. elkanii* structure was examined by ERRAT algorithm.

**Binding site Analysis**

LIGSITE and CASTp programs were used to search the protein binding sites by locating cavities in the *B. elkanii* structure. Through comparing the conserved residues in family of the studied protein and combining the *in silico* search results, the binding sites of *B. elkanii* were predicted. Those results were used to guide the following docking experiment. In order to investigate the interaction among *B. elkanii* with rhizobitoxin, the binding site on the *B. elkanii* model was determined by combining results obtained from LIGSITE [39] and CASTp[38]. Active sites are are shown in Figure 8.

![Binding mode](image)

**Fig. 8:** Binding mode of *B. elkanii* was predicted by LIGSITE and CASTp.
Structural comparison of active site region with template

*B. elkanii* model structure was used to compare binding site region with template, 3GVI [Figure 9]. In order to compare active site region of *B. elkanii* with that of template, the two structures were superimposed by using multiple alignments from Local ASH, followed by genetic-algorithm optimization and a residue equivalence defined by \( \exp[-(d/4)^2] \). RMSD value for *B. elkanii*-3GVI was calculated as 0.063Å as shown in the Figure 9. Moreover, the homology modeled structure were completely identical, and it showed that our model can fit exactly with the template.

![Fig. 9: Structural superimposition of model structure *B. elkanii* with the template.](image)

Protein-model deposition

The atomic coordinates of the theoretical model of *B. elkanii* from *Lespedeza* sp. have been deposited with Protein model database[43] (PMDB) which can be accessed with the code: PM0079076.

Gold docking

In order to further confirm the binding site of rhizobitoxine on *B. elkanii* from *Lespedeza* sp, one of the most appropriate methods to explore its predicted structural features is through docking studies with the known inhibitor rhizobitoxine. The crystal structure of protein was taken from the Protein Data Bank (entry PDB code: 3GVI). So, we considered the active site residues predicted from the LIGSITE and CASTp has been used for binding with rhizobitoxine that binds to *B. elkanii* from *Lespedeza* sp. with excellent binding affinity(GOLD score: 62.71). The best docking score are shown in the Figure 7. The docking result shows that the rhizobitoxine is located in the main hydrophobic pocket of the protein and also showed that rhizobitoxine mainly interacted with *B. elkanii* through H-bonding.
interactions. It can be seen in Figure 10.

![Docking pose of rhizobitoxine with B. elkanii.](image)

**Fig. 10: Docking pose of rhizobitoxine with B. elkanii.**

**Interactions with rhizobitoxine**

Hydrogen bonds play an important role for structure and function of biological molecules, especially for the enzyme catalysis. Accelrys Studio [44] was used for plotting the *B. elkanii* and rhizobitoxine interactions are shown in Figure 11. The interactions shown are those mediated by hydrogen bonds and hydrophobic contacts.

![Hydrogen bond interactions](image)

**Figure 11: Hydrogen bond interactions of rhizobitoxine with B.elkanii were predicted by LIGPLOT.**

Hydrogen bonds are indicated by dashed lines between the atoms involved, while hydrophobic contacts are represented by an arc with spokes radiating towards the ligand atoms they contact. The contacted atoms are shown with spokes radiating back. The
following hydrogen bond interactions were observed between rhizobitoxine and \textit{B. elkanii} [Figure 11]. The O atom of VAL55 of the active site of the receptor forms hydrogen bond with H1 of rhizobitoxine. The OG1 atom of VAL55 of the active site of the receptor forms hydrogen bond with H2 of rhizobitoxine. Gln128, Arg139, Val34, Ser 37, Ala38, Arg41, residues make electrostatic contacts with the inhibitor.

**CONCLUSION**

In the present work the 3D model of \textit{B. elkanii} was developed using homology modeling and validated by bioinformatics techniques. The model was then validated and further used for docking analysis with well-known inhibitor rhizobitoxine. The resulting docking solutions were analyzed for binding pattern and conformational analysis study. This information could also aid in the understanding of its catalytic mechanism. The detailed 3D structure, interaction information and the key residues identified are helpful for guiding the site-directed mutagenesis investigation and to understand the signaling pathway controlling the rhizobial infection of legumes.

**REFERENCES**


17. Owen LD, Wright, DA. Rhizobial-induced chlorosis in soybean: isolation, production in
30. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints.


39. Huang B, Schroeder M. LIGSITEcsc; predicting ligand binding sites using the Connolly surface and degree of conservation. BMC Structural Biology, 2006; 6: 19.


43. Castrignanò T, De Meo PD, Cozzetto D, Talamo IG, Tramontano A. The PMDB Protein Model Database. Nucleic Acids Research, 2006; 34: D306-309.