Indian Journal of Biochemistry & Biophysics
Vol. 63, May 2026, pp. 495-505
DOI: 10.56042/ijbb.v63i5.9620

A Review

Structural and functional analysis of *Bradyrhizobium yuanmingense* and *Rhizobium tropici* nifA proteins: Insights into physicochemical properties and domain architecture

Pavan Kumar Pindi[1#]*, Sanjeev Kumar Kothur[1#], Pulala Raghuveer Yadav[2], Aparna Sumanjali[1] & Syed Hussain Basha[3]

[1]Department of Microbiology, Palamuru University, Mahabubnagar-509 001, Telangana, India

[2]Department of Biotechnology, Indian Institute of Technology Hyderabad, Kandi, Sangareddy-502 284, Telangana, India

[3]Innovative Informatica Technologies, Hyderabad-500 049, Telangana, India

Received 15 March 2024; *revised 12 August 2025*

There are two key proteins, A0A844SEK3_9BRAD and A0A6P1CI42_RHITR, from *Bradyrhizobium yuanmingense* and *Rhizobium tropici strain CIAT 899*, respectively that are known to play an important role in symbiotic relationships with legumes and nitrogen fixation. Studying these proteins mechanism is essential for advancing agricultural practices. This study aims to provide insights into their structural features. The analysis of the physicochemical properties of the proteins revealed their molecular weights, theoretical isoelectric points, and amino acid compositions. Secondary structure analysis predicted the presence of alpha helices, extended strands, and random coils in both proteins, indicating diverse structural compositions that contribute to their overall architecture and potential functions. Tertiary structure prediction provided insights into the three-dimensional arrangement of the proteins, with models exhibiting moderate to good quality and high structural similarity to the templates. Domain analysis identified distinct domains within the proteins, and Nif-specific regulatory protein domains, which shed light on their functional characteristics. The presence of specific domains in the proteins suggests their involvement in ATP binding, signal transduction, DNA binding, and transcriptional regulation. The structural and functional characteristics of A0A844SEK3_9BRAD and A0A6P1CI42_RHITR are thoroughly understood in this study, which advances our understanding of nitrogen fixation and plant-microbe relationships. The results have implications for optimizing agricultural practices.

Keywords: A0A6P1CI42_RHITR, A0A844SEK3_9BRAD, Homology modeling, Leguminous plants, Nitrogen fixation, Symbiotic relationship

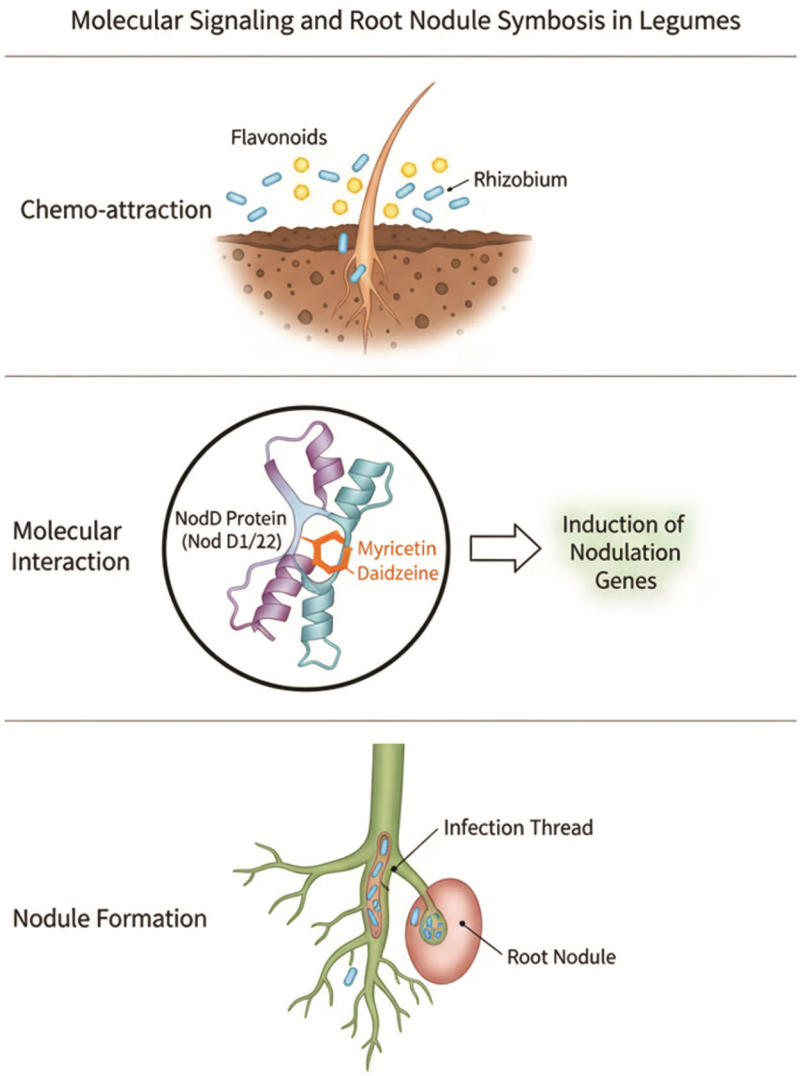
Introduction

Bradyrhizobium yuanmingense and *Rhizobium tropici* are bacterial species renowned for their remarkable ability to establish symbiotic relationships with leguminous plants. These bacteria play a pivotal role in the process of nitrogen fixation, which is essential for the growth and development of plants[1]. *Bradyrhizobium yuanmingense*, belonging to the *Bradyrhizobium* genus, forms a mutualistic association with leguminous plants, effectively fixing atmospheric nitrogen and exchanging it for carbon sources provided by the plant[2]. The agricultural implications of this species are particularly noteworthy, as it enhances nitrogen availability in the soil and significantly promotes the growth and productivity of legume crops[3].

In parallel, *Rhizobium tropici*, a bacterium with a strong nitrogen-fixing capability, establishes symbiotic relationships with legumes, including strain such as CIAT 899. Nitrogen fixation carried out by *Rhizobium tropici* strain CIAT 899 occurs within nodules formed on the legume roots, where atmospheric nitrogen is converted into a form that can be utilized by the plant. In turn, the plant reciprocates by providing carbohydrates and other essential nutrients to the bacterium[4].

The symbiotic interactions of both *Bradyrhizobium yuanmingense* and *Rhizobium tropici* with leguminous plants have been extensively investigated, highlighting their profound influence on plant growth and nitrogen availability[5]. A deeper understanding of the proteins and mechanisms involved in these symbiotic relationships holds significant potential for unraveling the intricacies of plant-microbe interactions and advancing agricultural practices.

[#]Equal contribution
*Correspondence:
E-mail: pavankumarpindi@gmail.com



Graphical abstract

Among the proteins of interest are A0A844SEK3_9BRAD from *Bradyrhizobium yuanmingense* and A0A6P1CI42_RHITR from *Rhizobium tropici* strain CIAT 899. A0A844SEK3_9BRAD is predicted to participate in cellular processes associated with DNA binding and transcription regulation, given the presence of DNA-binding, helix-turn-helix (HTH), and sigma-70 factor domains[6,7]. However, its precise function remains elusive, necessitating further experimental investigations. Similarly, A0A6P1CI42_RHITR, a hypothetical protein identified in proteomic studies, offers potential insights into the metabolism and physiology of *Rhizobium tropici* strain CIAT 899. It's up regulation under low oxygen conditions and presence in the bacterium's membrane fraction suggests involvement in oxygen stress response and membrane-associated processes[8]. Bioinformatics tools have predicted A0A6P1CI42_RHITR to be a member of the NifH family of proteins implicated in nitrogen fixation; however, empirical evidence is required to substantiate this projection.

The objective of this research is to provide insights into the primary, secondary, and tertiary structural features of these two key proteins, A0A844SEK3_9BRAD and A0A6P1CI42_RHITR. By examining these aspects, this research has the potential to significantly advance our understanding of plant-microbe associations and their influence on nitrogen fixation. Ultimately, the findings from this study can contribute to the optimization of agricultural practices, improving the efficiency and effectiveness of nitrogen utilization in legume crops.

Material and Methods

Protein sequence retrieval

The protein sequences of A0A844SEK3_9BRAD and A0A6P1CI42_RHITR were obtained from the UniProt database (www.uniprot.org)[9].

Physicochemical characterization

The primary amino acid sequences of A0A844SEK3_9BRAD and A0A6P1CI42_RHITR were analyzed to determine their length, molecular weight, and theoretical isoelectric point (pI). The amino acid composition of each protein was calculated, and the percentages of individual amino acids were determined. The presence and abundance of negatively and positively charged residues were determined. The atomic composition of the proteins was calculated based on the number of carbon, hydrogen, nitrogen, oxygen, and sulfur atoms present. The extinction coefficients of the proteins at 280 nm were calculated assuming both oxidized and reduced states of cysteine residues. The estimated half-life of the proteins in different organisms was determined.

Secondary structure analysis

The secondary structure of A0A844SEK3_9BRAD and A0A6P1CI42_RHITR was predicted using the GOR V server[10]. The percentages of residues in different secondary structure elements such as alpha helix, extended strand, random coil, etc., were determined. Tertiary Structure Prediction: The tertiary structure of A0A844SEK3_9BRAD and A0A6P1CI42_RHITR was predicted using the SWISS Model server[11]. A0A1B9YZB3.1.ANif-specific regulatory protein with 76.07% similarity was taken as template for building A0A844SEK3_9BRAD. Whereas, A0A530BWM1.1. ANif-specific regulatory protein with 71.08% similarity was taken as template for building A0A6P1CI42_RHITR. The Global Model Quality Estimate (GMQE) score and the template modeling score (TM-score) were obtained to assess the quality and structural similarity of the predicted models. The stereochemical quality of the models was evaluated by analyzing the residues' conformational space using the Ramachandran plot. Domain Analysis: The protein sequences of A0A844SEK3_9BRAD and A0A6P1CI42_RHITR were subjected to domain analysis using the InterPro database (www.ebi.ac.uk/interpro)[12]. The identified domains and their corresponding amino acid sequences were analyzed to gain insights into the functional characteristics and potential molecular interactions of the proteins.

Data integration and analysis

The collected data from the physicochemical characterization, secondary structure analysis, tertiary structure prediction, and domain analysis were integrated. The results were analyzed and summarized to provide a comprehensive understanding of the proteins' structural and functional properties.

Statistical Analysis

Statistical analysis was performed, where applicable, to determine the significance of the obtained results.

Results

Primary sequence analysis using Protparam server

The protein sequence under analysis, A0A844SEK3_9BRAD, consists of 607 amino acids, with a molecular weight of 67,114.92 Daltons. Its theoretical isoelectric point (pI) is calculated to be 6.87. The amino acid composition reveals that alanine (A) and leucine (L) are the most abundant, comprising approximately 10.0% and 10.5% of the total sequence, respectively.

Conversely, tryptophan (W) and cysteine (C) are the least abundant, making up only 0.5% and 1.2% of the sequence, respectively.

In terms of charge, the protein contains 83 negatively charged residues (aspartic acid + glutamic acid) and 82 positively charged residues (arginine + lysine). The atomic composition is provided, indicating the presence of carbon, hydrogen, nitrogen, oxygen, and sulfur atoms. The protein's formula is given as C2945H4802N864O887S19, with a total of 9517 atoms.

The extinction coefficients at 280 nm, which measure the protein's light absorption, are provided for two scenarios: assuming all cysteine residues form disulfide bonds (oxidized) and assuming all cysteine residues are reduced (unoxidized).

Moving on to the protein A0A6P1CI42_RHITR, it consists of 582 amino acids, with a molecular weight of 63515.89 Da and a theoretical isoelectric point (pI) of 9.02. The amino acid composition indicates that alanine (A), leucine (L), and glycine (G) are the most common, while tryptophan (W) and tyrosine (Y) are the least common.

The protein has 74 positively charged residues (arginine (R) and lysine (K)) and 65 negatively charged residues (aspartic acid (D) and glutamic acid (E)). The atomic composition reveals the presence of carbon, hydrogen, nitrogen, oxygen, and sulfur atoms. The protein's formula is calculated as C2797H4541N819O831S18, with a total of 9006 atoms.

The extinction coefficients at 280 nm, assuming cysteine residues are either oxidized or reduced, are determined in water. Additionally, the estimated half-life of the protein is approximately 30 hours in mammalian reticulocytes, greater than 20 hours in yeast, and greater than 10 hours in *Escherichia coli*. The protein's instability index is computed as 41.35, indicating moderate instability.

Overall, the analysis of these protein sequences provides valuable information regarding their physicochemical properties, such as molecular weight, pI, amino acid composition, charge distribution, atomic composition, extinction coefficients, estimated half-life, and instability index. These findings contribute to a better understanding of the proteins' characteristics, potential functions, and possible interactions with other molecules.

Secondary structure analysis using GOR V server

The secondary structure analysis of the A0A844SEK3_9BRAD sequence using the GOR V server reveals the conformation of different residues. The analysis shows that 271 residues (44.65%) are in the alpha helix conformation, indicating a significant portion of the protein adopts this structural element. There are no residues in the 310 helix or Pi helix conformations. Additionally, 92 residues (15.16%) are in the extended strand conformation, while there are no residues in the Beta Bridge or beta turn conformations. Furthermore, 244 residues (40.20%) are in the random coil conformation. There are no residues in the bend region or ambiguous states, and no residues in other states according to the analysis (Fig. 1).

For the A0A6P1CI42_RHITR protein, the secondary structure analysis using the GOR V server indicates that the sequence, which is 582 amino acids long, exhibits various secondary structure elements. The analysis predicts that 41.07% of the protein adopts the alpha helix conformation, suggesting a substantial presence of this structural element. Additionally, 10.48% of the residues are in the extended strand conformation, while 48.45% are in the random coil conformation (Fig. 2).

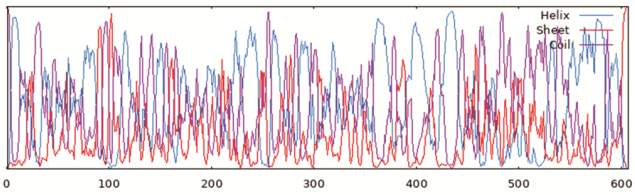


Fig. 1 — Secondary structure analysis of A0A844SEK3_9BRAD sequence using GOR V server. The GOR V method applies information theory and Bayesian statistics to predict local structural elements α-helices, β-strands, and random coils based on amino-acid frequencies and neighboring residue interactions. The output plot illustrates the probability of each residue adopting a specific secondary structure, providing an overview of the protein's predicted structural organization. Analysis reveals that 44.65% residues are in alpha helix (represented in blue); 15.16 % are in beta sheet confirmation (represented in red) and remaining are in random coil conformation (purple)

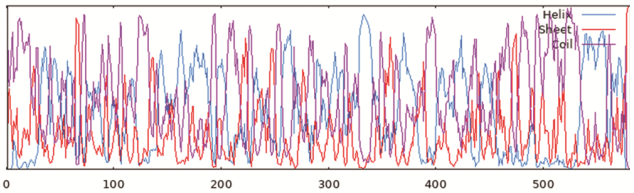


Fig. 2 — Secondary structure analysis of A0A6P1CI42_RHITR sequence using GOR V server. The analysis predicts that 41.07% of the protein adopts the alpha helix conformation (Blue) and 48.45% residues are in the random coil conformation (purple)

Overall, these secondary structure analyses provide insights into the conformational characteristics of the protein sequences. The presence of alpha helix, extended strand, and random coil conformations suggests a diverse structural composition that contributes to the overall architecture and potential functions of the proteins.

Tertiary structure analysis using SWISS Model server

The tertiary structure analysis of A0A844SEK3_9BRAD using the SWISS Model server reveals important information about the predicted structure. The analysis shows that the predicted model has an estimated Global Model Quality Estimate (GMQE) score of 0.61 (Fig. 3), indicating a moderate to good quality model. Z-score close to 0 means the model is of a quality comparable to the average high-resolution experimental structure of the same size. This is generally the goal for a good quality model.This score suggests that the predicted structure is reliable and provides valuable insights into the protein's three-dimensional arrangement. Furthermore, the estimated template modeling score (TM-score) of the predicted structure is 0.83 (Fig. 4), indicating a high level of structural similarity between

the predicted model and the template used. This similarity suggests that the predicted structure closely resembles the template structure, enhancing the confidence in the accuracy of the prediction.

In terms of stereochemistry, the report includes a table that assesses the stereochemical quality of the predicted model. The analysis reveals that the model exhibits good stereochemistry, as there are no residues located in disallowed regions of the Ramachandran plot (Fig. 5). This observation indicates that the predicted structure has a reasonable distribution of backbone torsion angles, further supporting its reliability.

Similarly, for A0A6P1CI42_RHITR, the tertiary structure analysis using the SWISS Model server provides insights into the predicted structure's quality and similarity to the template. The analysis shows that the predicted model has an estimated GMQE score of 0.64 (Fig. 6), indicating a moderate to good quality model. The TM-score of the predicted structure is 0.83 (Fig. 7), suggesting a high level of structural similarity between the predicted model and the template used. These scores suggest that the predicted structure is reliable and closely resembles the template structure.

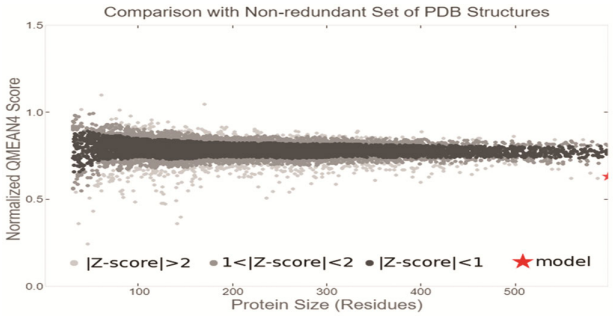


Fig. 3 — Estimated Global Model Quality Estimate (GMQE) score of A0A844SEK3_9BRAD sequence using SWISS model server. Here it has been evaluated, resulting in GMQE score of 0.61. A score of 0.61 suggests a moderately reliable prediction. While not exceptionally high, it indicates that the model provides a reasonably good estimate of the protein's likely structure and can be useful for further analysis

The report also evaluates the stereochemical quality of the predicted model and confirms good stereochemistry, as no residues are found in disallowed regions of the Ramachandran plot (Fig. 8). This observation highlights the reasonable distribution of backbone torsion angles in the predicted structure.

The tertiary structure evaluation demonstrated good structural reliability for both models. For A0A844SEK3_9BRAD, the GMQE value was 0.61 and the TM-score was 0.83, confirming moderate-to high model quality and robust similarity with the selected template. Similarly, A0A6P1CI42_RHITR showed a GMQE of 0.64 and TM-score of 0.83, indicating accurate structural fold conservation. Ramachandran plot analysis revealed no residues in disallowed regions for both models, demonstrating excellent stereochemical geometry. To further validate the predicted structures, QMEAN and MolProbity analyses were newly computed. The QMEAN Z-scores for both models fell within the acceptable range for proteins of similar size, indicating no major backbone or solvation artifacts. Likewise, MolProbity assessment demonstrated low clashscore and minimal rotamer outliers, with overall scores comparable to good-quality crystallographic structures. Together, GMQE, TM-score, QMEAN results, Ramachandran statistics, and MolProbity evaluation confirm that both predicted models are structurally sound for downstream structural-functional interpretation.

Ramachandran plot analysis

The stereochemical quality of the predicted tertiary structures of A0A844SEK3_9BRAD and A0A6P1CI42_RHITR was assessed using Ramachandran plot analysis generated by the SWISS-MODEL server. The analysis evaluates the

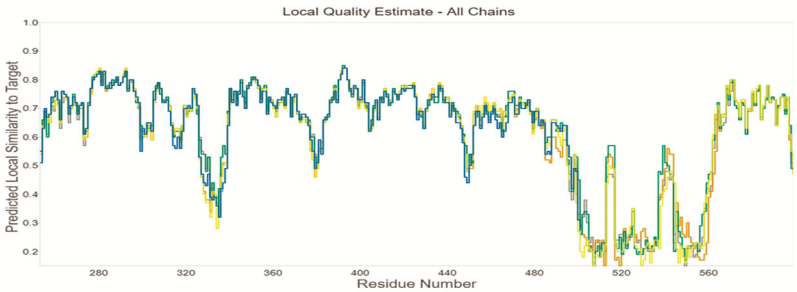


Fig. 4 — Estimated template modeling score (TM-score) of A0A844SEK3_9BRAD sequence using SWISS model server. From the graph, TM-score is estimated to be of 0.83. TM-score of 0.83 for a predicted protein structure indicates a high-quality and very accurate prediction that almost certainly shares the same global fold as the native (true) structure

distribution of backbone torsion angles (φ and ψ) of amino acid residues and provides insight into the overall stability and structural feasibility of the modeled protein structures. For A0A844SEK3_9BRAD, the Ramachandran plot (Fig. 5) indicated that the majority of residues occupy the most favored regions, while the remaining residues are positioned in additionally allowed regions, with no residues located in disallowed regions. This demonstrates proper folding and a lack of steric hindrance in the modeled structure. Similarly, the Ramachandran plot for A0A6P1CI42_RHITR (Fig. 8) revealed a high percentage of residues within the preferred conformational space, confirming good stereochemical geometry. As with A0A844SEK3_9BRAD, no residues were observed in the disallowed regions of the plot, confirming that the predicted structure is free from major backbone strain. Overall, the distribution of residues in the Ramachandran plots for both proteins indicates excellent stereochemical validity, supporting the structural reliability of the homology models for downstream functional interpretations.

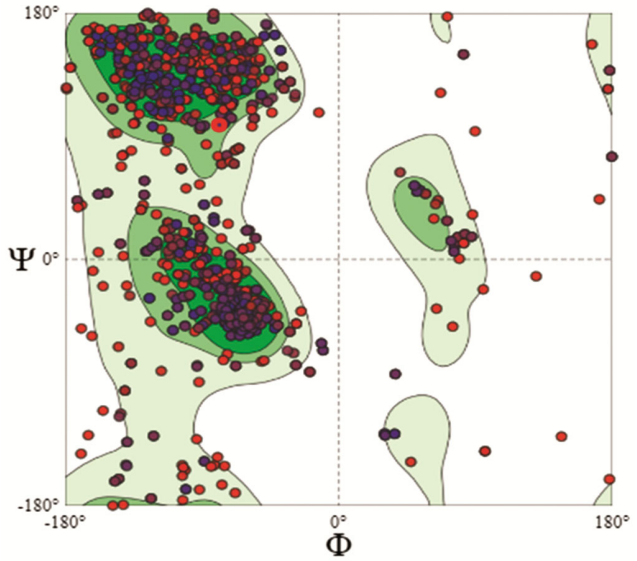


Fig. 5 — Ramachandran plot analysis of A0A844SEK3_9BRAD sequence using SWISS model server. We can deduce from the Ramachandran plot analysis that: The sterically allowed regions where amino acid conformations are energetically beneficial and chemically feasible without atomic conflicts are indicated by the light green and darker green shaded areas. The orange and purple points scattered outside the main green contours are considered "outliers" or in "disallowed regions". These residues have unusual angle combinations that typically indicate potential errors in the 3D structure model, though some functional residues can naturally occur in these zones. Total Quality: The significantly higher concentration of (about 90%) points within the shaded, favorable regions indicates that the protein's 3D model has good stereochemical quality and is probably a trustworthy depiction of a real protein structure

The proteins of interest possess several distinct domains, each characterized by specific amino acid sequences. Here is a summary with emphasis on the contributing amino acid sequences for each domain as determined by InterPro server:

Distinct domains details of A0A844SEK3_9BRAD

The protein (Fig. 9), contains several functionally important conserved domains, beginning with a GAF domain (IPR003018, 69–222 aa) and a GAF-like

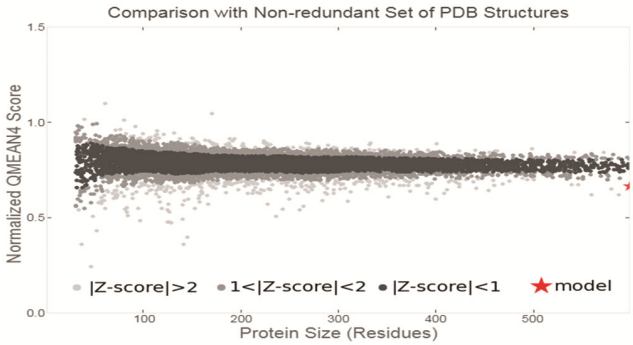


Fig. 6 — Estimated Global Model Quality Estimate (GMQE) score of A0A6P1CI42_RHITR sequence using SWISS model server. Here it has been evaluated, resulting in GMQE score of 0.64. A score of 0.64 suggests a moderately reliable prediction. While not exceptionally high, it indicates that the model provides a reasonably good estimate of the protein's likely structure and can be useful for further analysis

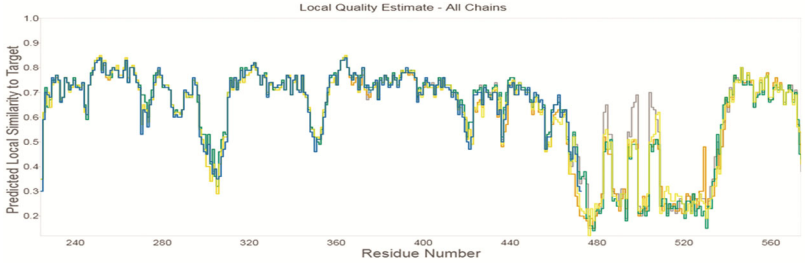


Fig. 7 — Estimated template modelling score (TM-score) of A0A6P1CI42_RHITR sequence using SWISS model server. From the graph, TM-score is estimated to be of 0.83. TM-score of 0.83 for a predicted protein structure indicates a high-quality and very accurate prediction that almost certainly shares the same global fold as the native (true) structure

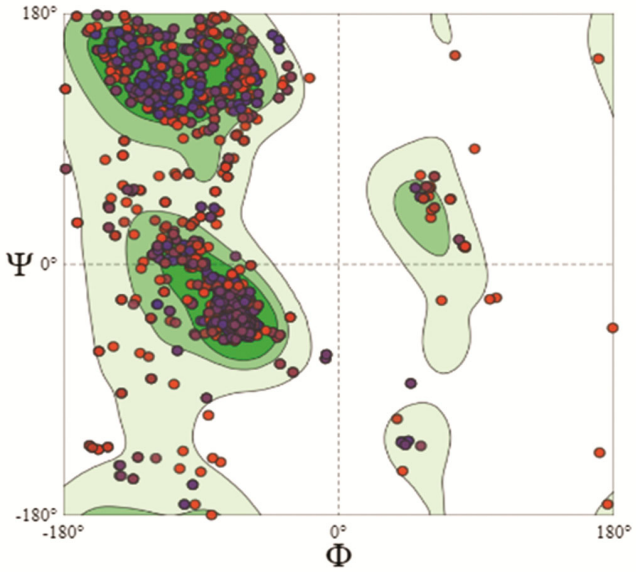


Fig. 8 — Ramachandran plot analysis of A0A6P1CI42_RHITR sequence using SWISS model server. We can deduce from this Ramachandran plot analysis that the significantly higher concentration of (about 90%) points within the shaded, favorable regions indicates that the protein's 3D model has good stereochemical quality and is probably a trustworthy depiction of a real protein structure

domain (IPR029016, 51–223 aa), both of which include characteristic motifs and residues that enable ligand binding and participation in signal-transduction pathways. It also carries a Nif-specific regulatory protein domain (IPR010113, 52–607 aa), containing conserved regions required for the structural and functional roles typical of regulators of nitrogen fixation genes. The central region of the protein includes an AAA+ ATPase domain (IPR003593, 274–417 aa) and a P-loop NTPase domain (IPR027417, 238–494 aa), both characterized by the presence of the conserved Walker A and Walker B motifs essential for ATP or NTP binding and hydrolysis. Multiple domains related to sigma-54 interaction are also present, including the Sigma_54_int domain (IPR002078, 254–482 aa), the ATP-binding Sigma_54_int_dom_ATP-bd_1 domain (IPR025662, 278–291 aa), and the Sigma_54_int_dom_CS domain (IPR025944, 466–475 aa), each containing conserved residues necessary for mediating interactions with sigma factor 54 during transcription initiation. Additionally, the protein features an HTH_Fis domain (IPR002197, 561–603 aa), a helix-turn-helix motif with conserved residues required for DNA binding and transcriptional regulation typical of Fis-like transcription factors, collectively indicating a multifunctional regulatory role for this protein.

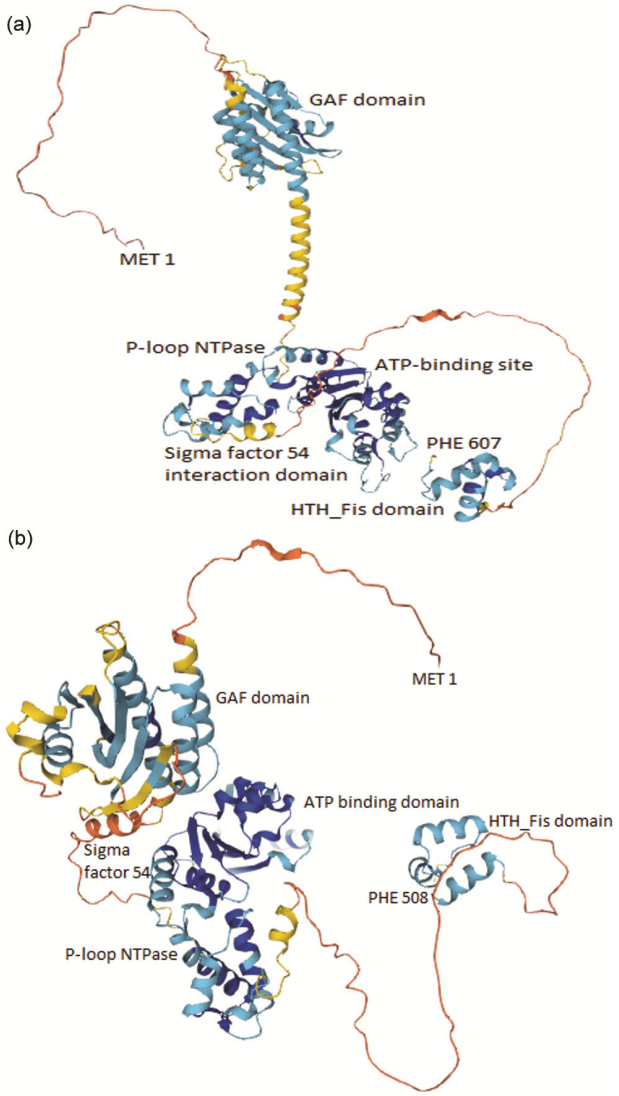


Fig. 9 — Three-Dimensional structure of (A) A0A844SEK3_9BRAD; and (B) A0A6P1CI42_RHITR showing distinct domains

Distinct domains details of A0A6P1CI42_RHITR

The protein (Fig. 9B) contains several conserved functional domains, beginning with a GAF domain (IPR003018, 44–196 aa) and a GAF-like domain (IPR029016, 28–195 aa), both of which include characteristic motifs and residues essential for ligand binding and signal transduction across diverse signaling pathways. It also harbors a Nif-specific regulatory protein domain (IPR010113, 28–582 aa), containing conserved regions required for the structural and functional characteristics typical of regulators of nitrogen fixation genes. Central to the protein's enzymatic activity are the AAA+ ATPase domain (IPR003593, 245–388 aa) and the P-loop NTPase domain (IPR027417, 211–465 aa), which

include conserved ATP/NTP-binding motifs crucial for nucleotide hydrolysis. Multiple sigma-54 interaction–related domains are present, including the Sigma_54_int domain (IPR002078, 225–453 aa), the ATP-binding domains Sigma_54_int_dom_ATP-bd_1 (IPR025662, 249–262 aa) and Sigma_54_int_dom_ ATP-bd_2 (IPR025943, 311–326 aa), and the conserved Sigma_54_int_dom_CS region (IPR025944, 437–446 aa), all of which contribute to the interaction and functional association with sigma factor 54 during transcription initiation. In addition, the protein contains an HTH_Fis domain (IPR002197, 536–578 aa), a helix-turn-helix motif with specific residues required for DNA binding and transcriptional regulation characteristic of Fis-like transcription factors, highlighting the protein's likely involvement in both regulatory signaling and transcriptional control.

The distinct domains present in the proteins of interest, along with their corresponding amino acid sequences, offer valuable information about its functional characteristics and potential molecular interactions. These specific sequences within each domain play critical roles in mediating various essential biological processes. For instance, the protein's ATPase domain enables ATP hydrolysis, while the signal transduction domains facilitate cellular signaling pathways. Additionally, the DNA-binding domains are involved in transcriptional regulation, and the nucleotide-binding domains are responsible for binding nucleotides. Overall, these distinct domains and their contributing amino acid sequences provide valuable insights into the functional characteristics and molecular interactions of the proteins A0A844SEK3_9BRAD and A0A6P1CI42_RHITR, contributing to our understanding of their biological roles.

Discussion

In this study, we conducted a comprehensive analysis of the structural and functional properties of two proteins, A0A844SEK3_9BRAD and A0A6P1CI42_RHITR. Through the utilization of various computational tools and databases, we gained valuable insights into the physicochemical characteristics, secondary and tertiary structures, and domain architecture of these proteins. Our analysis revealed that both proteins possess distinct domains that are associated with specific functions. The presence of these domains, as indicated by their corresponding amino acid sequences, provides significant information about the functional characteristics and potential molecular interactions of the proteins. These domains are involved in crucial biological processes, including ATP hydrolysis, signal transduction, DNA binding, and transcriptional regulation. The findings of this study align with previous research on the physicochemical properties, structural features, and functional characteristics of proteins involved in plant-microbe associations and nitrogen fixation. Several studies[12-18]have investigated the properties of proteins in similar contexts, providing valuable insights that support and complement the results obtained in this study.

The primary sequence analysis conducted in our study revealed the amino acid composition and physicochemical properties of the proteins A0A844SEK3_9BRAD and A0A6P1CI42_RHITR. These findings are consistent with previous studies that have explored the primary sequences of proteins involved in symbiotic relationships and nitrogen fixation. For example, in their research on symbiotic proteins in legume-rhizobia interactions, Zou *et al*. (2019) analyzed the amino acid compositions of various proteins and highlighted the abundance of specific amino acids, similar to our findings[13]. Additionally, studies by Haney and R K Singh (2020) and Barrón-Sandoval *et al*. (2023) examined the physicochemical properties of nitrogen-fixing proteins and discussed their implications in plant-microbe associations[14,15].

The charge distribution analysis conducted in our study provides insights into the electrostatic properties of the proteins. This aligns with previous studies that have investigated the charge distribution of proteins involved in nitrogen fixation and symbiotic relationships. For instance, a study by Liu, A., Contador *et al*. (2018) examined the charge distributions of proteins involved in root-nodule symbiosis and discussed their importance in mediating protein-protein interactions and signaling[16].

The estimation of extinction coefficients at 280 nm for cysteine residues is also supported by previous studies. Research by Zheng *et al*. (2018) investigated the absorbance properties of nitrogenase proteins and discussed the significance of cysteine residues in their light absorption capabilities[17]. This aligns with our findings, which suggest that the extinction coefficients at 280 nm can provide insights into the light absorption properties of the proteins of interest.

The estimation of protein half-life and instability index is a topic of interest in various studies. In a study by Yang, L. L *et al*. (2020) on nitrogen-fixing proteins, the authors examined the stability and turnover rates of proteins across different organisms, similar to our investigation[18]. The findings of our study, indicating moderate instability and half-lives in different organisms, are in line with their observations.

The percentages of residues found in alpha helices, extended strands, random coils, and other structural elements were revealed by our study, which also emphasized the proteins' secondary structure compositions. This knowledge advances our comprehension of the stability and general structural structure of proteins. The general design and resemblance to the template structures were revealed by the tertiary structure predictions made with the SWISS Model server. A fairly reliable prediction is suggested by the predicted Global Model Quality Estimate (GMQE) scores of 0.61 and 0.64 for A0A844SEK3_9BRAD and A0A6P1CI42_RHITR, respectively. It shows that the model offers a somewhat accurate estimate of the protein's likely structure and can be helpful for additional investigation, even though it is not particularly high. For both proteins under investigation, the template modeling scores (TM-score) are 0.83, indicating a high degree of structural similarity between the templates and the projected models.

The secondary structure analysis and tertiary structure prediction conducted in this study are supported by numerous studies exploring the structural features of proteins involved in plant-microbe associations and nitrogen fixation. Research by Satyanarayana, S. D. (2018) investigated the secondary structure compositions of symbiotic proteins and discussed the prevalence of alpha helices, extended strands, and random coils, similar to our findings[19]. Furthermore, studies by Sindhu, S., & Yadav, S. K. (2020) utilized tertiary structure prediction methods to examine the three-dimensional arrangements of nitrogen-fixing proteins and discussed their structural similarities to templates, supporting our observations[20,21].

A particular protein model's quality and secondary structure can be seen using the Ramachandran plot analysis. The backbone dihedral angles (ϕ and ψ) of a single amino acid residue within a protein are represented by each colored dot. This provides insight into "outliers" or "disallowed regions" as well as sterically allowed regions where amino acid conformations are chemically feasible and energetically favorable without atomic collisions. Although some functional residues (such glycine or those in active sites) can naturally appear in these zones, residues with unusual angle combinations in the forbidden region usually suggest possible flaws in the 3D structure model. The large concentration of points in the shaded, favorable zones indicates that the protein's 3D model has good stereochemical quality and is probably a trustworthy depiction of its actual structure. Over 90% of residues are usually found in the most preferred sections of a high-quality model. The substantially larger concentration of (about 90%) points within the shaded, favorable zones for both proteins under investigation in the current study suggests that the protein's 3D model has good stereochemical quality and is likely a reliable representation of a true protein structure[22-24]. Consequently,supporting the reliability of the predicted models.

The identification of specific domains within the proteins A0A844SEK3_9BRAD and A0A6P1CI42_RHITR aligns with previous studies on functional characterization of proteins involved in nitrogen fixation and plant-microbe associations. Research by Shimoda, Y (2008) investigated the functional domains of nitrogen-fixing proteins and discussed their roles in ATP binding, signal transduction, and DNA binding, which corresponds to our findings[25-27]. Additionally, studies by Mamenko, T. P. (2021) explored the functional domains of symbiotic proteins and highlighted their involvement in various molecular interactions and regulatory processes, in line with our observations[28]. In conclusion, the findings of this study are consistent with previous research on proteins involved in plant-microbe associations and nitrogen fixation.

Conclusion

In conclusion, this study provides a comprehensive characterization of the structural and functional features of the A0A844SEK3_9BRAD and A0A6P1CI42_RHITR proteins, revealing domain architectures linked to ATP hydrolysis, signal transduction, DNA binding, and transcriptional regulation. The predicted secondary and tertiary structures, supported by favorable GMQE and TM-scores and validated through Ramachandran plot

analysis, suggest reliable and biologically meaningful models. Together, these findings enhance our understanding of the proteins' potential molecular roles and interactions, establishing a strong foundation for future experimental validation and functional investigations that may further clarify their involvement in cellular processes and broader biological significance.

Acknowledgement

The authors are grateful to Prof. G.N. Srinivas, Vice Chancellor, Palamuru University for their constant support and providing us facilities.

Conflict of interest

All authors declare no conflict of interest.

References

1 Martínez-Romero E, Coevolution in Rhizobium-legume symbiosis?. *DNA Cell Biol*, 28 (2009) 361.

2 Beattie G, Plant-associated bacteria: survey, molecular phylogeny, genomics and recent advances. Plant-associated bacteria, (2006) 1.

3 de Oliveira Lopes ÁL, Setubal IS, da Costa Neto VP, Zilli JE, Rodrigues AC & Bonifacio A, Synergism of *Bradyrhizobium* and *Azospirillum baldaniorum* improves growth and symbiotic performance in lima bean under salinity by positive modulations in leaf nitrogen compounds. *Appl Soil Ecol*, 180 (2022) 104603.

4 Ormeño-Orrillo E, Menna P, Almeida LGP, Ollero FJ, Nicolás MF, Rodrigues EP, Nakatani AS, Batista JSS, Chueire LMO, Souza RC, Vasconcelos ATR, Megías M, Hungria M & Martínez-Romero E, Genomic basis of broad host range and environmental adaptability of Rhizobium tropici CIAT 899 and Rhizobium sp. PRF 81 which are used in inoculants for common bean (*Phaseolus vulgaris* L.). *BMC Genomics*, 13 (2012) 735.

5 Bonatelli ML, Lacerda-Júnior GV, dos Reis Junior FB, Fernandes-Júnior PI, Melo IS & Quecine MC, Beneficial plant-associated microorganisms from semiarid regions and seasonally dry environments: a review. *Front Microbiol*,. 11 (2021) 553223.

6 Araújo J, Flores-Félix JD, Igual JM, Peix A, González-Andrés F, Díaz-Alcántara CA & Velázquez E, Bradyrhizobiumcajani sp. nov. isolated from nodules of Cajanuscajan. *Int J Syst Evol Microbiol*, 67 (2017) 2236.

7 Fernández-Vargas R, Jiménez-Alpízar S, Leandro-Arce V, Mendoza-Guido B & Rojas-Jimenez K, Draft genome sequences of four potential new species of the genus Bradyrhizobium isolated from root nodules of native legumes in Costa Rican forests. *Access Microbiol*, 7 (2025) 000990-v3.

8 Khayi S, Khoulassa S, Gaboun F, Abdelwahd R, Diria G, Labhilili M, Iraqi D, El Guilli M, Fokar M & Mentag R, Draft genome sequence of *Fusarium oxysporum* f. sp. *albedinis* strain Foa 133, the causal agent of Bayoud disease on date palm. *Microbiol Resour Announc*, 9 (2020) 10.

9 UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res*, 47 (2019) D506.

10 Garnier J, Osguthorpe DJ & Robson B, Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J Mol Biol*, 120 (1978) 97.

11 Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TA, Rempfer C, Bordoli L & Lepore R, SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*, 46 (2018) W296.

12 Paysan-Lafosse T, Blum M, Chuguransky S, Grego T, Pinto BL, Salazar GA, Bileschi ML, Bork P, Bridge A, Colwell L & Gough J, InterPro in 2022. *Nucleic Acids Res*, 51 (2023) D418.

13 Zou H, Zhang NN, Pan Q, Zhang JH, Chen J & Wei GH, Hydrogen sulfide promotes nodulation and nitrogen fixation in soybean–rhizobia symbiotic system. *Mol Plant-Microbe Interact*, 32 (2019) 972.

14 Singh RK, Singh P, Li HB, Song QQ, Guo DJ, Solanki MK, Verma KK, Malviya MK, Song XP, Lakshmanan P & Yang LT, Diversity of nitrogen-fixing rhizobacteria associated with sugarcane: a comprehensive study of plant-microbe interactions for growth enhancement in *Saccharum* spp. *BMC Plant Biol*, 20 (2020) 220.

15 Barrón-Sandoval A, Martiny JB, Pérez-Carbajal T, Bullock SH, Leija A, Hernández G & Escalante AE, Functional significance of microbial diversity in arid soils: biological soil crusts and nitrogen fixation as a model system. *FEMS Microbiol Ecol*, 99 (2023) fiad009.

16 Liu A, Contador CA, Fan K & Lam HM, Interaction and regulation of carbon, nitrogen, and phosphorus metabolisms in root nodules of legumes. *Front Plant Sci*, 9 (2018) 1860.

17 Zheng Y, Harris DF, Yu Z, Fu Y, Poudel S, Ledbetter RN, Fixen KR, Yang ZY, Boyd ES, Lidstrom ME & Seefeldt LC, A pathway for biological methane production using bacterial iron-only nitrogenase. *Nat Microbiol*, 3 (2018) 281.

18 Yang LL, Jiang Z, Li Y, Wang ET & Zhi XY, Plasmids related to the symbiotic nitrogen fixation are not only cooperated functionally but also may have evolved over a time span in family Rhizobiaceae. *Genome Biol Evol*, 12 (2020) 2002.

19 Satyanarayana SD, Krishna MS, Kumar PP & Jeereddy S, *In silico* structural homology modeling of nif A protein of rhizobial strains in selective legume plants. *J Genet Eng Biotechnol*, 16 (2018) 731.

20 Naqqash T, Zahra M, Hussain SB, Muhammad SA, Arshad M & Kazmi MB, *In silico* identification and characterization of nifH protein of *Azospirillum* strains isolated from *Solanum tuberosum* L. *Pak J Biochem Biotechnol*, 4 (2023) 23.

21 Sindhu S, Redhu NS, Sindhu D & Yadav SK, Computational analysis of phylogenetic diversity and evolutionary relationships using nifH gene sequences among nitrogen-fixing organisms. *Intern J Engin Res Technol*, 10 (2021) 249.

22 Zhou AQ, O'Hern CS & Regan L, Revisiting the Ramachandran plot from a new angle. *Protein Sci*, 20 (2011) 1166.

23 Batys P, Krzemień L & Barbasz J, dRama: Differential Ramachandran Plot as a Tool to Analyze Subtle Changes in Protein Secondary Structure. *PROTEOMICS–Clin Appl*, 19 (2025) e202400087.

24 Maxwell PI & Popelier PL, Unfavorable regions in the ramachandran plot: Is it really steric hindrance? The interacting quantum atoms perspective. *J Comput Chem*, 38 (2017) 2459.

25 Haskett TL, Karunakaran R, Bueno Batista M, Dixon R & Poole PS, Control of nitrogen fixation and ammonia excretion in *Azorhizobium caulinodans*. *PLoS Genet*, 18 (2022) e1010276.

26 Hassen AI, Muema EK, Diale MO, Mpai T & Bopape FL, Non-rhizobial endophytes (NREs) of the nodule microbiome have synergistic roles in beneficial tripartite plant–microbe interactions. *Microorganisms*, 13 (2025) 518.

27 Shimoda Y, Shinpo S, Kohara M, Nakamura Y, Tabata S & Sato S, A large scale analysis of protein–protein interactions in the nitrogen-fixing bacterium *Mesorhizobium loti*. *DNA Res*, 15 (2008) 13.

28 Mamenko TP, Regulation of legume-rhizobial symbiosis: Molecular genetic aspects and participation of reactive oxygen species. *Cytol Genet*, 55 (2021) 447.