ISSN NO. 2320-5407

Journal homepage: http://www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH

RESEARCH ARTICLE

HOMOLOGY MODELING OF NITROGENASE ENZYME ENCODED BY NIFHDK GENE CLUSTER OF PSEUDOMONAS PUTIDA

Sakshi Issar and R.K. Gaur*

Department of Science, Faculty of Arts, Science & Commerce (FASC), Mody Institute of Technology & Science (MITS), Lakshmangarh-332311, Sikar (Rajasthan), India.

Manuscript Info

Manuscript History:

Received: 15 September 2013
Final Accepted: 24 September 2013
Published Online: October 2013

Key words:
Homology modeling,
Pseudomonas putida,
Nitrogenase, Nitrogen fixation

Corresponding Author
gaurrajarshi@hotmail.com

Abstract

Nitrogenase enzyme plays a key role in the fixation of atmospheric nitrogen during Biological Nitrogen Fixation (BNF) mechanism used by many diazotrophs. This enzyme is made-up of 2 components of proteins in which Component-I is encoded by *nif*DK genes cluster (Mo-Fe protein) and Component-II is encoded by *nif*H gene (Fe protein). These components together make *nitrogenase* enzyme to work properly. In the present study, Homology modeling approach is used to model the sequence of *nif*HDK gene cluster and results show that the generated model of *nitrogenase* enzyme of *Pseudomonas putida* is showing similarity with the sequence and structure of Nitrogenase enzyme of *Azotobacter vinelandii.*

Introduction

The human genome project has already discovered millions of proteins (http://www.swissprot.com). The potential of the genome project can only be fully realized once we can assign, understand, manipulate, and predict the function of these new proteins (Sanchez and Sali, 1997; Frishman et. al., 2000; Domingues et. al., 2000). Predicting protein function generally requires knowledge of protein three-dimensional structure (Blundell et. al., 1978; Weber, 1990), which is ultimately determined by protein sequence (Anfinsen, 1973). Protein structure determination using experimental methods such as X-ray crystallography or NMR spectroscopy is very time consuming (Johnson et. al., 1994). To date, fewer than 2% of the known proteins have had their structures solved experimentally. Homology modeling is often an efficient way to obtain information about proteins of interest. Compared with *ab initio* protein folding, homology modeling is more accurate and reliable. The quality of a homology model is directly correlated with the sequence similarity between target and template. Though a homology model is not perfect, it is still very useful in a wide spectrum of applications where information about 3D conformation of a protein is required.

The nitrogen fixation process is believed to have played a crucial role in early cellular evolution, in particular when the geochemical reserves of fixed nitrogen in the biosphere became depleted. Because nitrogen fixation is maintained in *Bacteria* and *Archaea*, it is often hypothesized that *nif* genes may originate from the last common ancestor, even though other scenarios are proposed (Fani et. al., 2000; Raymond et. al., 2004). Yan et. al.(2008**)** mentioned that in *P. stutzeri* strain A1501, the nitrogenase complex comprised the Mo-Fe protein encoded by *nifDK* and the Fe protein encoded by *nifH*. Full assembly of the complex required the products of a dozen other *nif* genes extremely conserved in free-living and symbiotic diazotrophs, in particular for processing of nitrogenase metallo-clusters and catalytic stability (*nifMZ*, *nifW*, and *nifUS*) and for synthesis of a specific molybdenum cofactor (FeMo-co) bound to the Mo-Fe protein (*nifB*, *nifQ*, *nifENX*, *nifV*, and *nifH*) (Klipp et. al., 2004).

Rhizobacteria from the genus *Pseudomonas* provide an excellent example of a combination of multiple mechanisms for effective bio-control including direct antagonism and induction of plant resistance. *Pseudomonas* spp. produces several metabolites with antimicrobial activity towards other bacteria and fungi (Hass and Keel, 2003). Bacteria which are aerobic, polarly flagellated, Gram-negative, non-spore-forming, rod-shaped, and which metabolize sugars oxidatively, but do not readily fit elsewhere, are often broadly classified as *Pseudomonas* species.

Pseudomonas species are usually differentiated by their phenotypic characteristics (Stanier et. al., 1966) in conjunction with numerical data analysis (Sneath and Sokal, 1973).

The 3-D structure of *nitrogenase* enzyme of *nif* gene cluster of *P. putida* is unknown so far. Homology modelling is a reliable technique that can consistently predict the 3D structure of a protein with precision akin to one obtained at low resolution by experimental means (Renom et. al., 2000). This technique depends upon the alignment of a protein sequence of unknown structure (target) with that of a homologue of known structure (template) (Arnab et. al., 2010). Since the study of 3D structure of a protein is helpful in recognizing the details of a protein, this method is increasingly becoming of widespread use in the field of bioinformatics (Paramsivasan et. al., 2006). This work contributes to understanding of conformational and structure-function properties of MoFe protein of *nitrogenase* enzyme of *nif*HDK genes of *P. putida.*

MATERIAL AND METHODS

Isolation and Molecular Characterization of Bacterial Isolate:

Isolation of bacteria and its molecular characterization was conducted in Dept. of Biotechnology, MITS University, as previously described by Issar et. al. (2012**)**. As *nif* genes are the carrier of nitrogenase enzyme so, *nif*HDK gene was sequenced for homology modelling of MoFe protein of *nitrogenase* enzyme. Different Databases, software's, online tools and web servers' viz. PDB, NCBI, BLAST, UCEF Chimera, 3D-JIGSAW etc. were used for the homology modelling.

Homology Modelling:

For homology modelling, *nif*HDK gene sequence of *P. putida* was downloaded fromGenBank-NCBI database in FASTA format. The primary structure analysis of TARGET protein was obtained by use of ProtParam tool of ExPASy Proteomics Server (http://cn.expasy.org/cgi-bin/protparam). For this, TARGET protein sequence (*nif*HDKgene sequence of *P. putida*) was submitted to the ProtParam tool, which computes various physico-chemical properties that can be deduced from a protein sequence (Gasteiger et. al., 2005).

To be better evaluated, the possible secondary structures of TARGET protein were evaluated using the different web programs viz., PDBsum (http://www.ebi.ac.uk/thorntonsrv/databases/pdbsum/) and the Hierarchical Neural Network (HNN) (http://npsa-pbil.ibcp.fr/cgibin/npsa_automat.pl?page=npsa_nn.html) (Arnab et. al., 2010). In PDBsum Database, the GENERATE option was selected to produce possible secondary structure of TARGET protein by submitting the TARGET protein sequence in pdb file format (Laskowski, 2007).

The first step in homology modeling technique requires recognition of the protein structures linked to the target sequence and the subsequently selection of templates (Centeno et. al., 2005). A homologous template was selected on the protein database PDB (Protein Data Bank, http://www.rcsb.org/pdb/home/home.do) and then analysis of amino acid sequence showed that the optimal template protein code named 3MIN (chain A and B) in PDB was highly homologous (99%) with *nitrogenase* enzyme encoded by *nif*HDK genes of *P.putida*. The three-dimensional structures model of the TARGET protein was constructed by use of the program 3D-JIGSAW(http://bmm.cancerresearchuk.org/~3djigsaw/). 3D- JIGSAWserver looks for homologous templates in databases (PFAM+PDB+nr) and splits the query sequence into domains. If good templates are found, the best covered domain is then modelled using a maximum of 2. After alignment between target & best templates, the server sent constructed file in PDB format to the assigned e-mail (Bates et. al., 2001).

Model Building, Evaluation and Validation

The three-dimensional structure of the TARGET protein was constructed by use of the program UCSF Chimera, which utilizes PDB files of both target and template proteins(Pettersen et. al., 2004)**.** Verify3D server (http://nihserver.mbi.ucla.edu/Verify_3D/) was used forstructure evaluation of the best model generated by Chimera software (Eisenberg et. al., 1997). The constructed model was evaluated by Ramachandran plot (http://nihserver.mbi.ucla.edu/SAVES/) (Ramachandran et. al., 1963).

Function prediction:

For the prediction of functions of the modeled protein sequence, the 3d2Go server (http://www.sbg.bio.ic.ac.uk/phyre/pfd/index.html) was used by using sequence and structure in the reference of

Gene Ontology (GO). By using some information like- overall topological similarity to structures with known function, geometry and residue similarity of predicted functional sites to regions of known structure and sequence homology to functionally annotated sequences etc., the 3d2Go server predicts the function of protein. The MAMOTH structural alignment program was used for full topology search of the model (Ortiz et. al., 2002).

Target Sequence/Query Sequence (FASTA format)

>gi|369938359|gb|AEX25777.1| NifH [*P. putida*]
MAMRQCAIYGKGGIGKSTTTQNLVAALAELGKKVMIVGCDPKVDSTRLILHSKVQNTIMEMAAEAGTVE
DLELEDVLKTSYGDIKCVESGGPEPGVGCAGRGVITAINFLEEEGAYEDDLDFVFYDVLGDVVCGGFAMPI
RENKAQEIYVVCSGEMME

>gi|369938411|gb|AEX25782.1| NifD [*P. putida*]
MTGMSREEVESLIQEVLGVYPEKARKDRNKHLAVNDPAVTQSKKCIISNKKSQPGLMTIRGCAYAGSKGV
VWGPIKDMIHISHGPVGCGQYSRAGRRNYYIGTTGVNAFVTMNFTSDFQEKDIVFGGDKKLAKLIDEVETL
FPLNKGISVQSECPIGLIGDDIEPVSKVKGAELSKTIVPVRCEGFRGVSQSLGHHIANDAGRDWVLGKRDED
TTFASTPYDVAVIGDYNIGGDAWSSRILLEEMGLRCVAQWSGDGSISEIELTPKVKLNLVHCYRSMNYISRH
MEEKYGIPWMEYNFFGPTKTIESLRAIAAKFDESIQKKCEEVIAKYKPEWEAVVAKYRPRLEGKRVMLYIG
GLRPRHVIGAYEDLGMEVVGTGYEFAHNDDYDRTMKEMGDSTLLYDDVTGYEFEEFVKRIKPDLIGSGIK
EKFIFQKMGIPFRQMHSWDYSGPYHGFDGFAIFARDMDMTLNNPCWKKLQAPWEASEGAEKGAASA

>gi|369938426|gb|AEX25783.1| NifK [*P. putida*]
DKIKATYPLFLDQDYKDMLAKKRDGFEEKYPQDKIDEVFQWTTTKEYQELNFHREALTVNPAKACQPLGA
VLCALGFEKTMPYVHGSQGCVAYFRSYFNRHFREPVSCVSDSMTEDAAVFGGQQNMKDGLQNCKATYKP
DMIAVSTTCMAEVIGDDLNAFINNSKKEGFIPDGFPVPFAHTPSFVGSHVTGWDNMFEGIARYFTLKSMDD
KVVGSNKKINIVPGFETYLGNFRVIKRMLSEMGVGYSLLSDPEEVLDTPADGQFRMYAGGTTQEEMKDAP
NALNTVLLQPWHLEKTKKFVEGTWKHEVPKLNIPMGLDWTDEFLMKVSEISGQPIPASLTKERGRLVDMM
TDSHTWLHGKRFALWGDPDFVMGLVKFLLELGCEPVHILCHNGNKRWKKAVDAILAASPPGKNATVYIG
KDLWHLRSLVFTDKPDFMIGNSYGKFIQRDTLHKGKEFEVPLIRIGFPIFDRHHLHRSTTLGYEGAMQILTTL
VNSILERLDEETRGMQATDYPHDL

Template Sequence (FASTA FORMAT)
>gi|2098376|pdb|3MIN|B Chain B, Nitrogenase Mofe Protein from *A. vinelandii*, Oxidized State
SQQVDKIKASYPLFLDQDYKDMLAKKRDGFEEKYPQDKIDEVFQWTTTKEYQELNFQREALTVNPAKAC
QPLGAVLCALGFEKTMPYVHGSQGCVAYFRSYFNRHFREPVSCVSDSMTEDAAVFGGQQNMKDGLQNCK
ATYKPDMIAVSTTCMAEVIGDDLNAFINNSKKEGFIPDEFPVPFAHTPSFVGSHVTGWDNMFEGIARYFTLK
SMDDKVVGSNKKINIVPGFETYLGNFRVIKRMLSEMGVGYSLLSDPEEVLDTPADGQFRMYAGGTTQEEM
KDAPNALNTVLLQPWHLEKTKKFVEGTWKHEVPKLNIPMGLDWTDEFLMKVSEISGQPIPASLTKERGRL
VDMMTDSHTWLHGKRFALWGDPDFVMGLVKFLLELGCEPVHILCHNGNKRWKKAVDAILAASPYGKNA
TVYIGKDLWHLRSLVFTDKPDFMIGNSYGKFIQRDTLHKGKEFEVPLIRIGFPIFDRHHLHRSTTLGYEGAM
QILTTLVNSILERLDEETRGMQATDYNHDLVR

>gi|2098375|pdb|3MIN|A Chain A, Nitrogenase Mofe Protein From *A. vinelandii*, Oxidized State
TGMSREEVESLIQEVLEVYPEKARKDRNKHLAVNDPAVTQSKKCIISNKKSQPGLMTIRGCAYAGSKGVV
WGPIKDMIHISHGPVGCGQYSRAGRRNYYIGTTGVNAFVTMNFTSDFQEKDIVFGGDKKLAKLIDEVETLF
PLNKGISVQSECPIGLIGDDIESVSKVKGAELSKTIVPVRCEGFRGVSQSLGHHIANDAVRDWVLGKRDEDT
TFASTPYDVAIIGDYNIGGDAWSSRILLEEMGLRCVAQWSGDGSISEIELTPKVKLNLVHCYRSMNYISRHM
EEKYGIPWMEYNFFGPTKTIESLRAIAAKFDESIQKKCEEVIAKYKPEWEAVVAKYRPRLEGKRVMLYIGG
LRPRHVIGAYEDLGMEVVGTGYEFAHNDDYDRTMKEMGDSTLLYDDVTGYEFEEFVKRIKPDLIGSGIKE
KFIFQKMGIPFREMHSWDYSGPYHGFDGFAIFARDMDMTLNNPCWKKLQAPWEASEGAEKVAAS

RESULT AND DISCUSSION

Homology modeling estimates the 3D structure of a target protein sequence by using its alignment to one or more protein template of known structure (Renom et. al., 2000). For structure based protein molecule design and function investigation, homology modeling is most suitable method (Melo and Feytmans, 1998). The modeling process involves of target-template selection and alignment, model building and model evaluation. (Renom et. al., 2000) As

the numbers of known protein structures are increasing and protein model software's are improving, the accuracy of the models is increasing (Renom et. al., 2000).

Most microorganisms that perform biological N_2 fixation do so with an evolutionarily conserved nitrogenase protein complex (Howard and Rees, 1996). The high degree of similarity of protein sequence of nitrogenases among microorganisms suggests an early origin or lateral gene transfer among prokaryotic lineages (Postgate and Eady, 1988). As nitrogenase and 16S rRNA sequences accumulate in the databases, the high degree of sequence similarity supports early evolution of nitrogenase in an early ancestor (Postgate and Eady, 1988; Young, 1992).

Nitrogenase is a versatile enzyme which is capable of reducing N_2, nitrous oxide, acetylene, azide, cyanide, methyl isocyanide, protons, and analogs of some of these compounds. Nitrogenase consists of two proteins, a molybdenum- iron protein (MoFe protein) and an iron protein (Fe protein); neither has been demonstrated to have catalytic activity by itself (Ortiz and Burris, 1975). The nitrogenase enzyme is composed of two multi sub unit metallo-proteins. Component I contains the active site for N_2 reduction, has a molecular weight of about 250 kDa and is composed of two heterodimers, encoded by the *nifD* and *nifK* genes. Component II (about 70 kDa) couples ATP hydrolysis to inter protein electron transfer and is composed of two identical subunits encoded by the *nifH* gene. One might suppose that the presence of the *nif* genes would depend on selection for N_2 fixing capability, because the multiple genes involved in regulation and assembly occupy a fairly large piece of genomic real estate. Nitrogenase gene expression is highly regulated (Hoover, 2000), at levels ranging from transcription (Chen *et. al*., 1998) to post-translational protein modification (Kim *et. al*.,**1999**). Transcription of the *nifHDK* operon is a good marker for N_2 fixing conditions, as it is not constitutively expressed and is regulated in response to factors that control N_2 fixation.

The analysis result of the primary structure analysis of TARGET protein of *P. putida* indicated that sequence contained 1166 amino acids, the total molecular weight was130473.2kD and theoretical isoelectric point (pI) was 5.64. The total number of negatively charged residues (Asp + Glu) and positively charged residues (Arg + Lys) were 162 and 135 respectively. The formula was $C_{5822}H_{9040}N_{1546}O_{1725}S_{68}$ and total number of atoms was 18201. The instability index (II) was computed to be 35.83. The program PDBsum took the automatic algorithm: STRIDE (Frishman et. al., 1995), which uses hydrogen bond energy and main chain dihedral angles to recognize helix, coils and strands. This arithmetic was used to predict the secondary structure of TARGET protein of *P. putida*. The wiring diagram of secondary structures of TARGET protein is shown in Figure 1. The result of the Hierarchical Neural Network method showed that the percentage of alpha helix, random coil, extended strand were 31.65, 49.31 and 19.04%, respectively.

Protein-BLAST was performed against Protein Data Bank for analysis of Query/Target protein seq. Our results show that the TARGET protein of *P. putida* contained 517 amino acid residues, and results obtained with the protein data bank shown 99.0% identity with Chain A and chain B of *Nitrogenase* MoFe protein from *A. vinelandii* (3MIN).

The 3D structure of TARGET protein of *P.putida* was built by using UCSF Chimera software, which used PDB file constructed by 3D-JIGSAW (Figure 2). The 3D structure model had -23425.50 Kcal/mol energy. The stereochemical quality and geometric analysis of a protein structure was performed by analyzing the residue-by-residue geometry and overall structure geometry with the help of Verify3D and PROCHECK (Laskowski et. al., 1993).

Figure – 1: Predicted Secondary structure assignment of modeled *nitrogenase* enzyme of *P.putida* by PDBsum Database

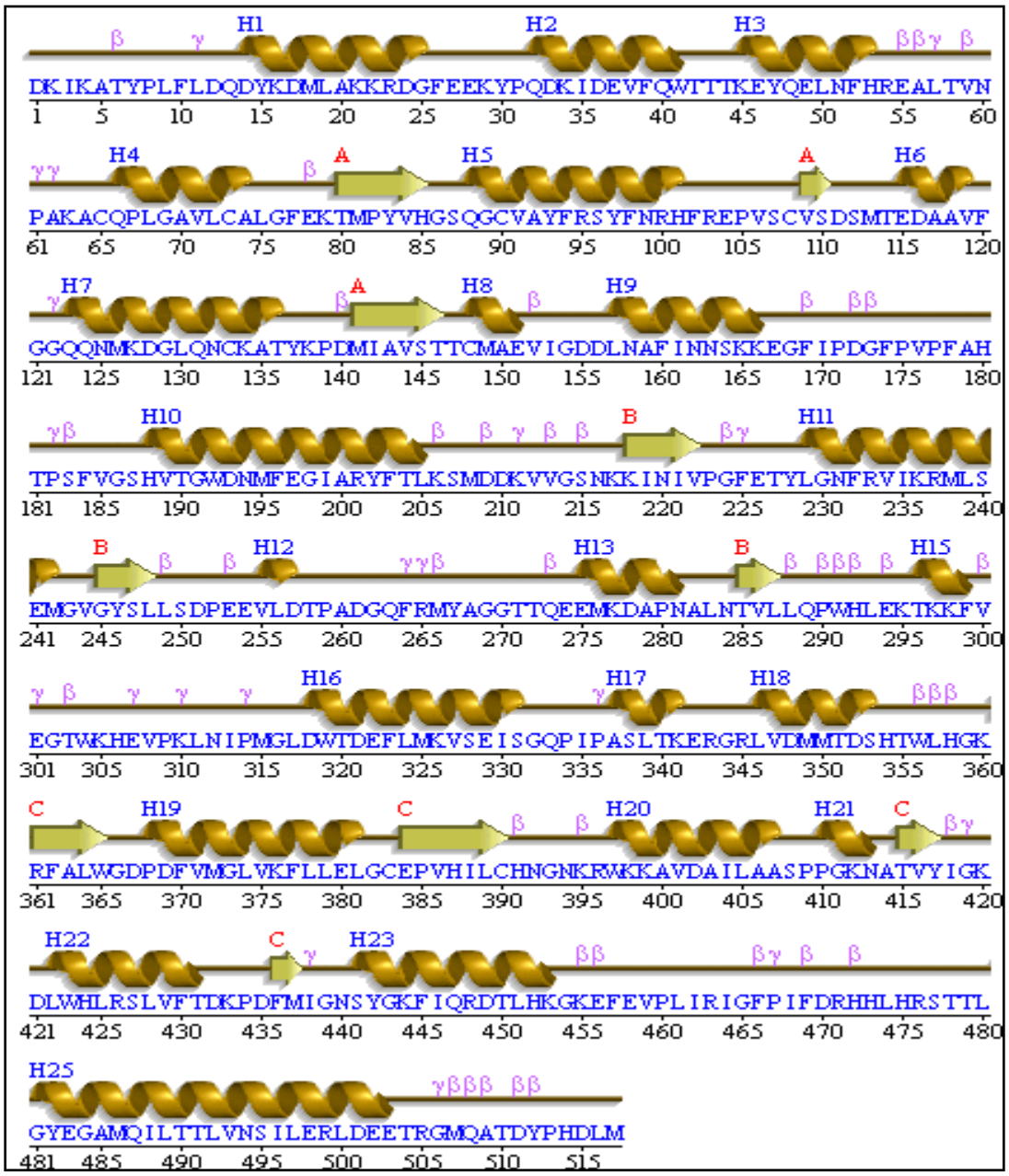
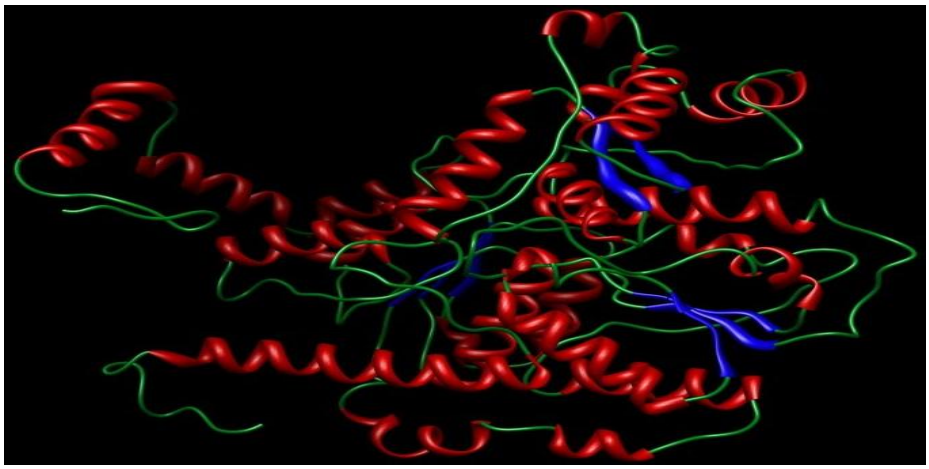


Figure 2: Ribbon diagram of 3D structures of the *nitrogenase* enzyme of *P. putida* where Red, Blue and Green color represents Helix, Strand and Coil respectively.



To validate the homology modeled MoFe protein structure encoded by *nif*HDK gene cluster (TARGET model), a Ramachandran plot was drawn and the structure was analyzed by PROCHECK, a well known protein structure checking program. It was found that the 84.8% of residues fell in the most favored regions, 13.9% residues fell in the additional allowed regions, and 0.9% fell in generously allowed regions and 0.4% of the residues fell in the disallowed conformations (Figure 3). A good quality Ramachandran plot has over 90% in the most favored regions (Xiao et. al., 2004; Balakrishnan et. al., 2010) but the Ramachandran plot of Mofe protein of *nitrogenase* had only 84.8% of residues in the most favored regions therefore, it is near to a good quality model. A comparison was also made between Ramachandran plots of TEMPLATE (3MIN) and TARGET protein models (Table-1) and results showed that the generated model of TARGET protein is acceptable. TEMPLATE and TARGET protein sequences had 89.0% and 84.8% residues respectively in their

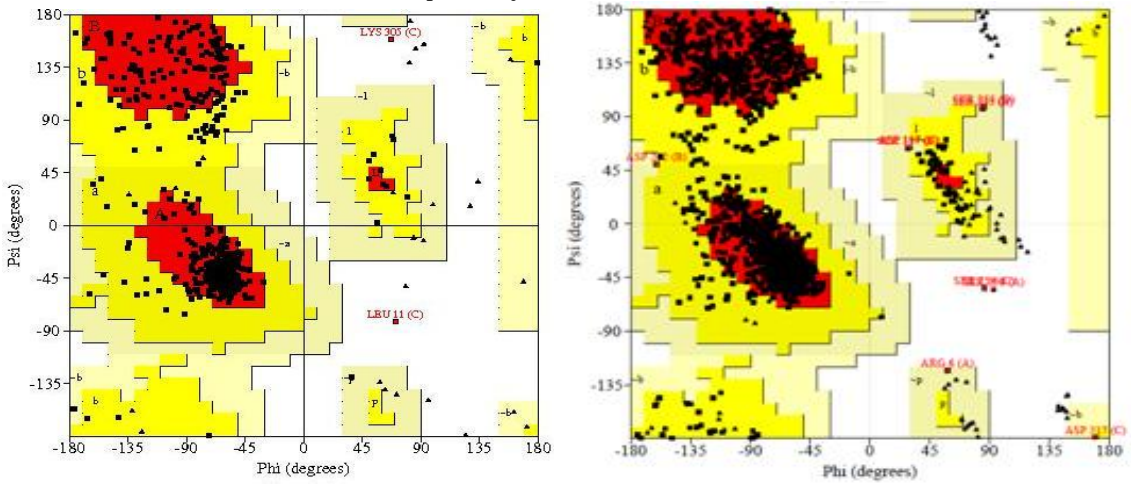


Figure 3: Ramachandran map of MoFe protein of *nitrogenase* enzyme of *P.putida*. The plot calculation on the three-dimensional (3D) model of MoFe protein was calculated with the PROCHECK program.

Table-1: Comparative analysis of Ramachandran plot statistics between TARGET and TEMPLATE protein (3MIN) models.

RAMACHANDRAN PLOT STATISTICS (RESIDUES SHOWING IN %)				
PROTEIN MODELS	**MOST FAVORED REGION (%)**	**ADDITIONAL ALLOWED REGION (%)**	**GENEROUSLY ALLOWED REGION (%)**	**DISALLOWED REGION (%)**
TARGET PROTEIN	**84.8**	**13.9**	**0.9**	**0.4**
TEMPLATE PROTEIN (3MIN)	**89.0**	**10.5**	**0.4**	**0.1**

Ramachandran plots. Therefore, both of these models can be place in near to good quality models. In Ramachandran plot, due to high percentage of residues in core region and less percentage of residues in allowed and generously allowed regions, TEMPLATE protein sequence is more stable than TARGET protein sequence.

Verify3D (Eisenberg et. al., 1997) works best on proteins with at least 100 residues. The compatibility of an atomic model (3D) with its own amino acid (1D) is analyzed by Verify3D. A collection of good structures is used as a reference to obtain a score for each of the 20 amino acids in this structural class. In Verify3D graph, Profile score above zero in the Verify3D graph (Bowieet. al., 1991; Luethy et. al., 1992) corresponds to acceptable environment of the model. The high score of 0.80 indicated that environment profile of the model is good (fig-4).

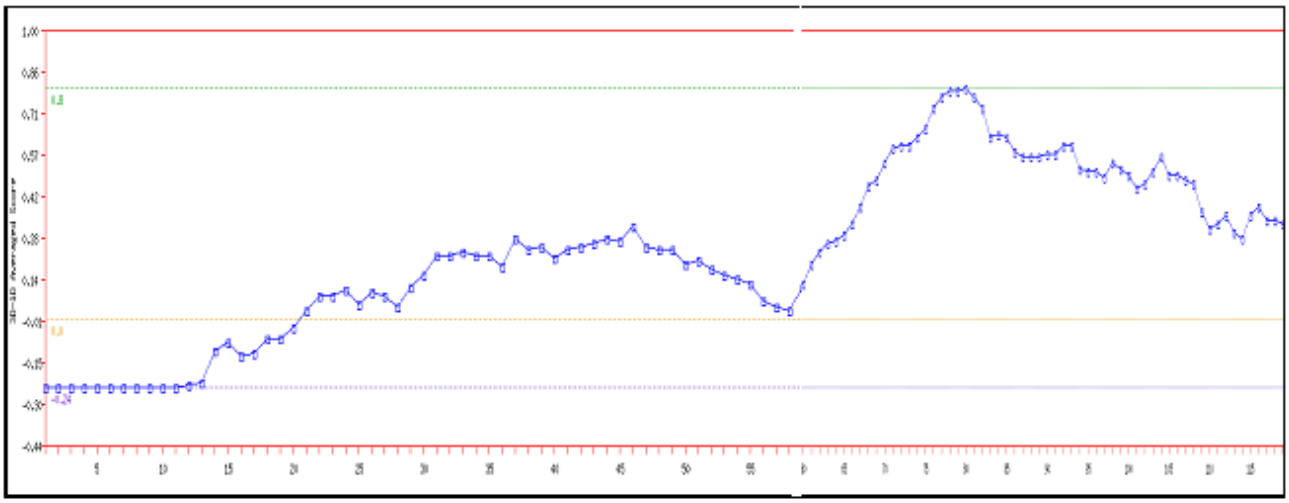


Fig-4: Verify 3D graph of *nitrogenase* enzyme of *P.putida*.

The 3d2Go server uses several method of function prediction, using sequence and structure to predict Gene Ontology (GO) terms for the protein (Shankaracharya *et. al.*, 2011). For TARGET protein sequence, 3d2Go server predicted various GO terms with their description and confidence (Table-2), where confidence ranges between 0-1 and 1 being the most confident prediction. The server predicted some functions like- oxidation reduction, oxidoreductase activity, metal-ion binding, metal cluster binding, oxidoreductase activity, acting on iron-sulfur proteins as donors, dinitrogen as acceptor, nitrogen fixation, nitrogenase activity, etc. for the TARGET protein sequence with good confidence values (Table-2), which confirms its function.

Table-2: Results showing the function prediction of modeled TARGET protein with 3d2Go server

GO TERMS	DESCRIPTION	CONFIDENCE
0005737	**cytoplasm**	**0.27**
0046872	**Metal-ion Binding**	**0.24**
0051540	**Metal Cluster Binding**	**0.11**
0016732	**oxidoreductase activity, acting on iron-sulfur proteins as donors, dinitrogen as acceptor**	**0.11**
0016612	**molybdenum-iron nitrogenase complex**	**0.10**
0016163	**Nitrogenase activity**	**0.10**
0016163	**Nitrogen fixation**	**0.10**

The tile image and superimpose image of TARGET and TEMPLATE proteins are shown in Figure 5 and 6. The 3D image alignment (Figure 6) showed that the target protein is showing similarity with chain A and B of Mofe protein of *nitrogenase* enzyme of TEMPLATE protein (3MIN) which belongs to *A. vinelendii*.

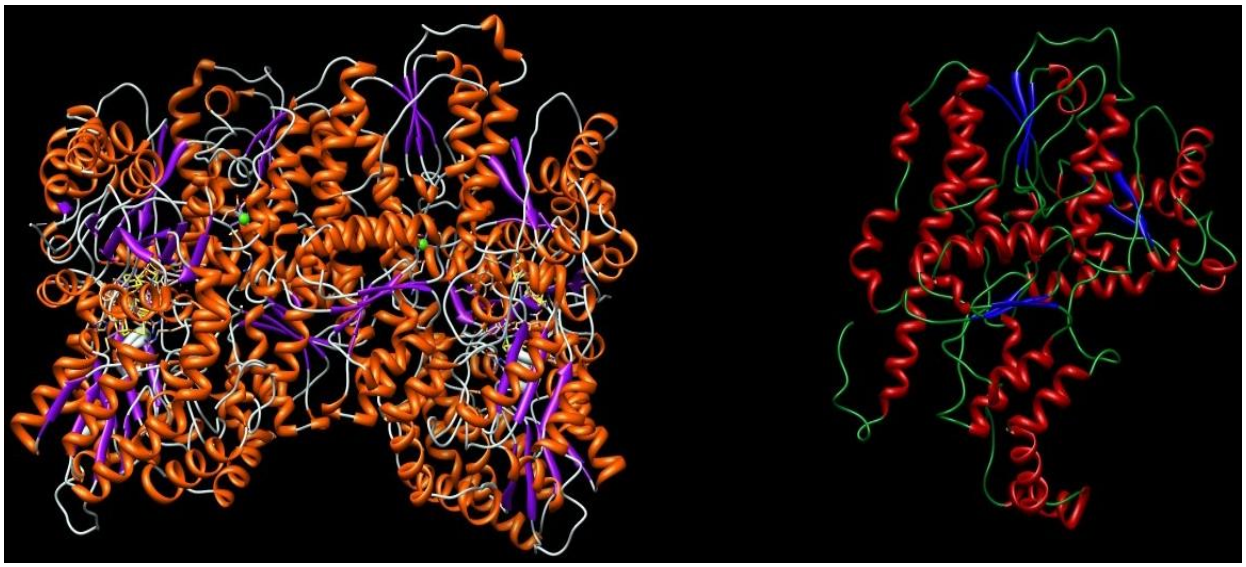


Figure 5: 3D representation of MoFe protein [A] *Azotobacter vinelandii* (3MINB) and [B] *Pseudomonas putida* (Target). Color represents purple→ helix, fusia→strands and dark red→coil for TEMPLATE protein (3MIN) & Red, Blue and Green colors represents helix, strand and coil respectively for Target protein

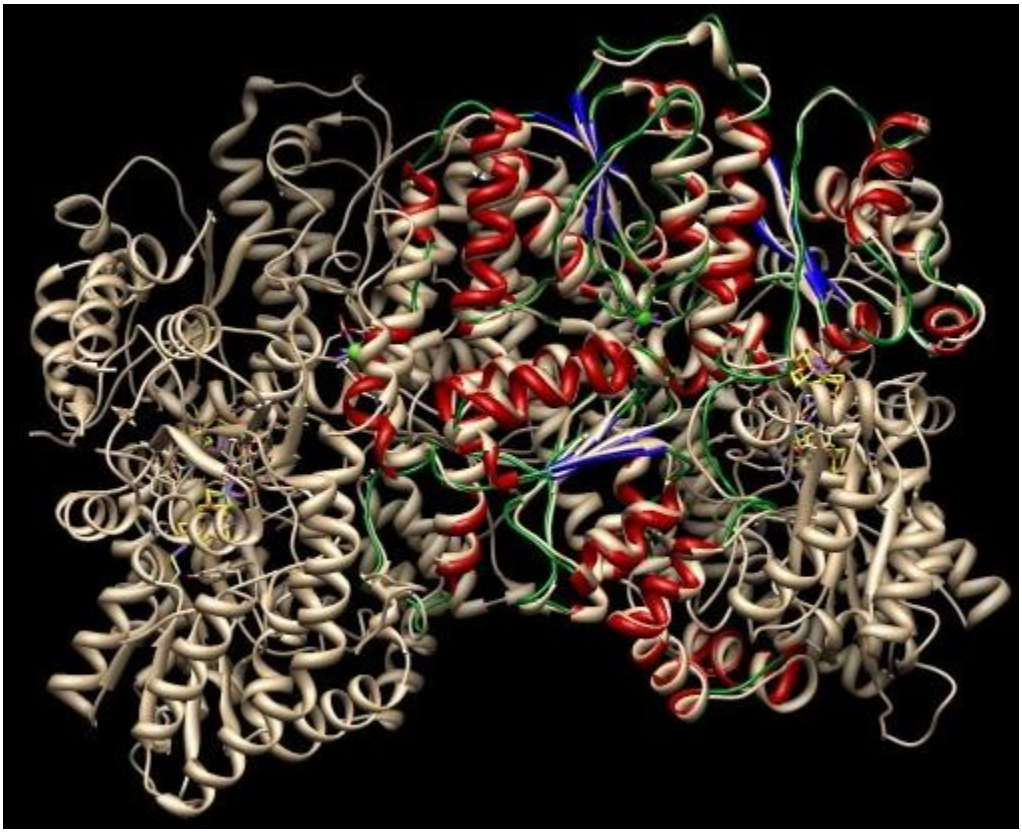


Figure 6: Superimposition image showing alignment between Template (3MIN-without color ribbon image) and Target protein (with color ribbon image)

3D protein structure prediction helps us to find out their active sites, binding sites etc. The 3D of model of TARGET protein, which was developed through Homology Modeling and subsequently prediction of functional characteristics of MoFe protein of *nitrogenase* enzyme therefore, provide better understanding of mode of action of

*nif*HDK genes encodes for Mo-Fe protein and Fe protein of *nitrogenase* enzyme which take combined active part in N_2-fixation mechanism in *P. putida* isolated from shekhawati region of Rajasthan.

CONCLUSION

Knowledge of 3D structure is very important for the understanding of the protein function. But, experimental structure determination methods are highly expensive and time consuming.So,homology or comparative modeling is very much useful in designing the three dimensional structure of proteins. These theoretical 3D structures are highly helpful in drug designing and in other studies.

The homology modeling method has not been used to generate the 3-D structure of MoFe protein from *P.putida* encoded by *nif*HDK genes. In this study, the 3-D structure of *nitrogenase* enzyme satisfied on the basis of energy and mostly the amino acid resides within the most favored region in Ramachandran plot. The analysis of primary structure prediction results obtained from ExPASy Proteomics Server and secondary structure prediction results obtained from PDBsum program shows that the constructed *nitrogenase* model of *P. putida* is mostly similar to the homologous MoFe protein structure (TEMPLATE seq.- 3MIN chain A and B) of *A. vinelandii*. The superimpose image (fig-6) and functions predicted by 3d2Go server (Table-2) confirms the activity of *nitrogenase* enzyme in *P. putida.* So, It would conclude that this structure fulfills the parameter provided for making a good model and could be used for different research prospects.

ACKNOWLEDGEMENT

Authors are thankful to all individuals associated with the present study.

REFERENCES

1. Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. Science, 181:223–230.

2. Arnab, S., Saubashya, S., Louis, S.T., Asim, K., Bothra, S.T. and Uttam, K.M. (2010). Homology modelling of the Frankia nitrogenase iron protein. Symbiosis, 50: 37-44.

3. Arnold. K., Bordoli. L., Kopp, J. and Schwede, T. (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics, 22: 195-201.

4. Balakrishnan, M., Srivastava, R.C. and Pokhriyal, M. (2010). Homology Modeling and Docking studies between HIV-I protease and carbamic acid. IJBT, 9: 96-100.

5. Bates, P.A., Kelley, L.A., MacCallum, R.M. and Sternberg, M.J. (2001). Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. Proteins, 2001:39-46.

6. Blundell, T.L., Bedarkar, S., Rinderknecht, E., and Humble, R.E. (1978). Insulin-like growth factor: A model for tertiary structure accounting for immunoreactivity and receptor binding. Proc. Natl. Acad. Sci. USA, 75:180–184.

7. Bowie, J.U., Luthy, R. and Eisenberg, D. (1991). A method to identify protein sequences that fold into known three-dimensional structures. Science, 253:164-170.

8. Centeno, N.B., Planas-Iglesias, J. and Oliva, B. (2005). Comparative modeling of protein structure and its impact on microbial cell factories. Microbial Cell Factories, 4: 20.

9. Chen, Y.B., Dominic, B., Mellon, M.T. and Zehr, J.P. (1998). *Circadian rhythm* of nitrogenase gene expression in the diazotrophic filamentous nonheterocystous cyanobacterium *Trichodesmium*sp. Strain IMS 101. J Bacteriol, 180: 3598–3605.

10. Domingues, F.S., Koppensteiner, W.A., and Sippl, M.J. (2000). The role of protein structure in genomics. FEBS Lett, 476:98–102.

11. Eisenberg, D., Luthy, R. and Bowie, J.U. (1997). VERIFY3D: Assessment of protein models with three-dimensional profiles. Methods Enzymol., 277: 396-404.

12. Fani, R., Gallo, R. and Lio, P. (2000). Molecular evolution of nitrogen fixation: The evolutionary history of the nifD, nifK, nifE, and nifN genes. J Mol Evol, 51:1–11.

13. Frishman D. and Argos, P. (1995). Knowledge-based protein secondary structure assignment. Proteins, 23: 566-579.

14. Frishman, D., Goldstein, R.A., and Pollock, D.D. (2000). Protein evolution and structural genomics. Pac. Symp. Biocompu, 12:3–5.

15. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D. and Bairoch. A. (2005). *Protein Identification and Analysis Tools on the ExPASy Server:* (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press. pp. 571-607.

16. Hass, D. and Keel, C. (2003). Regulation of antibiotic production in root colonizing Pseudomonas spp. and relevance for biological control of plant disease. Annual Review of Phytopathology, 41:117–153.

17. Hoover, T. (2000). Control of nitrogen fixation genes in Klebsiella pneumoniae. In Prokaryotic Nitrogen Fixation. Triplett, E.W. (ed). Norfolk, England: Horizon Scientific Press, pp. 131–147.

18. Howard, J.B. and Rees, D.C. (1996). Structural basis of biological nitrogen fixation. Chem Rev, 96:2965-2982.

19. Issar, S., Sharma, S., Choudhary, D.K., Gautam, H.K. and Gaur, R.K. (2012). Molecular characterization of Pseudomonas spp. isolated from root nodules of various leguminous plants of Shekhawati region, Rajasthan, India. American Journal of Plant Science, 3: 1-4.

20. Johnson, M.S., Srinivasan, N., Sowdhamini, R., and Blundell, T.L. (1994). Knowledge based protein modeling. Crit. Rev. Biochem. Mol. Biol., 29:1–68.

21. Jose M., Ortiz, R. and Burris, R. H. (1975). Interactions among Substrates and Inhibitors of Nitrogenase. Journal of Bacteriology, 123(2), 537-545.

22. Kim, K., Zhang, Y.P., and Roberts, G.P. (1999). Correlation of activity regulation and substrate recognition of the ADPribosyltransferase that regulates nitrogenase activity in *Rhodospirillum rubrum*. J Bacteriol, 181: 1698–1702.

23. Klipp, W., Masepohl, B., Gallon, J.R. and Newton, W.E. (2004). Genetics and Regulation of Nitrogen Fixation in Free-Living Bacteria. Kluwer Academic, Dordrecht, The Netherlands.

24. Laskowski, R.A. (2007). Enhancing the functional annotation of PDB structures in PDBsum using key figures extracted from the literature. Bioinformatics, 23(14):1824-1827.

25. Laskowski, R.A., Macarthur, M.W., Moss, D.S. and Hornton, J.M. (1993). Procheck: a program to check the stereochemical quality of protein structures journal of applied crystallography. 26:283-291.

26. Luthy R., Bowie, J.U. and Eisenberg, D. (1992). Assessment of protein models with three-dimensional Profiles. Nature, 356: 83-85.

27. Martin-Renom, M.A., Stuart, A.C., Fiser, A., Sanchez, R., Melo, F. and Sali, A. (2000). Comparative protein structure modeling of genes and genomes. Annu. Rev. Biophys. Biomol. Struct, 29: 291-325.

28. Melo, F. and Feytmans, E. (1998). Assessing protein structures with a non-local atomic interaction. Journal of Molecular Biology, 277 (5): 1141–1152.

29. Ortiz, A.R., Streuss, C.E. and Olmea, O. (2002). MAMMOTH (Matching Molecular Models Obtained from Theory): an automated method for model comparison. Protein Sci., 11: 2606-2621.

30. Paramsivasan, R., Sivaperumal, R., Dhnanjeyan, K.J., Thenmozhi, V. and Tyagi, B.K. (2006). Prediction of 3-dimensional structure of salivary odorant binding protein-2 of the mosquito Culex quinquefasciatus, the vector of human lymphatic filariasis. In Silico Biol., 7: 1-6.

31. Pettersen,E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis.J Comput Chem., 25(13):1605-12.

32. Postgate, J.R., and Eady, R.R. (1988). The evolution of biological nitrogen fixation. In Nitrogen Fixation: Hundred Years After. Bothe, H., de Bruijn, F.J, and Newton, W.E. (eds). Stuttgart: Gustav Fischer, pp. 31–40.

33. Ramachandran, G.N., Ramakrishnan, C. and Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. J. Mol. Biol., 7: 95-99.

34. Raymond, J., Siefert, J.L., Staples, C.R. and Blankenship, R.E. (2004). The natural history of nitrogen fixation. Mol Biol Evol, 21:541–554.

35. Renom, M., Stuart, A.C., Fiser, A., S´anchez, R., Melo, F. and Annu, A.S. (2000). Comparative protein structure modeling of genes and genomes. Marc A. Rev. Biophys. Biomol. Struct., 29:291–325.

36. Rubio1, L.M., and Ludden, P.W. (2008). Biosynthesis of the Iron-Molybdenum Cofactor of Nitrogenase. Annu. Rev. Microbiol. 62:93–111.

37. Sanchez, R., and Sali, A. (1997). Comparative protein structure modeling as an optimization problem. J. Mol. Struct. (Theochem), 398–399:489–496.

38. Shankaracharya, Das, S. and Vidhyarthi, A.S. (2011). Structure and function prediction of Human homologue hABH5 of E. coli ALKB5 using in silico approach. Nat. Proceedings, 10.1038/npre.2011.5597.2.

39. Sneath, P.H.A. and Sokal, R.R. (1973). Numerical Taxonomy. W.H. Freeman, San Francisco, CA.

40. Stanier, R.Y., Palleron, N.J. and Doudorof, M. (1966). Aerobic Pseudomonads - a Taxonomic Study. Journal of General Microbiology, 43:159-271.

41. Weber, I.T. (1990). Evaluation of homology modeling of HIV protease. Proteins, 7:172–184.

42. Xiao, C., Rajput, Z.I. and Hu, S. (2007). Improvement of a commercial foot-and mouth disease vaccine by supplement of Quil A. Vaccine, 25: 4795- 4800.

43. Yan, Y., Yang, J., Dou, Y., Chen, M., Ping, S., Peng, J., Lu, W., Zhang, W., Yao, Z., Li, H., Liu, W., He, S., Geng, L., Zhang, X., Yang, F., Yu, H., Zhan, Y., Li, D., Lin, Z., Wang, Y., Elmerich, C., Lin, L. and Jin, Q. (2008). Nitrogen fixation island and rhizosphere competence traits in the genome of root-associated *Pseudomonas stutzeri* A1501. PNAS, 105(21):7564-7569.

44. Young, J.P.W. (1992). Phylogenetic classification of nitrogen fixing organisms. In Biological Nitrogen Fixation. Stacey, G., Evans, H.J., and Burris, R.H. (eds). New York: Chapman & Hall, pp. 43–86.