

## MOLECULAR MODELING AND COMPUTATIONAL CHARACTERIZATION OF HMGR ENZYME OF *Hevea brasiliensis*

Snehalata Sahu, Bibhuti Prasad Barik\* and Sushil Kumar Dutta

DBT-Bioinformatics Infrastructure Facility, North Orissa University,

Sriramchandra Vihar, Takatpur, Baripada, Odisha, India

\*Corresponding Author: [bibhutiprasadbarik@gmail.com](mailto:bibhutiprasadbarik@gmail.com)

### ABSTRACT

*Hevea brasiliensis* is one of the common sources of rubber production in the world. Cutting the bark of this tree results in to release of latex, then it is collected, preserved and stabilized. In this plant 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR) is a critical enzyme of mevalonic acid (MVA) pathway. Structural changes in HMG-COA may result in to disruption of the pathway ultimately declining rubber production. In this study we have modeled the 3D structure of this enzyme and characterized it in-silico. The modeled structure has been verified and evaluated by using some computational programs. The study will be valuable to understand the structural and functional aspects of the enzyme for academic and industrial purposes.

**Keywords:** *Hevea brasiliensis*, HMGR enzyme, Molecular modeling, Model evaluation

### INTRODUCTION

Rubber is an integral part of human life since time immemorial. *Hevea brasiliensis* a common rubber crop plant and has been used industrially for natural rubber production. *H. brasiliensis* is basically a perennial tropical tree species. It is a source of natural rubber (cis-1, 4-polyisoprene) and is present in colloidal form in the latex [1]. Latex is extracted by wounding the bark. Para rubber is a product of several species of the genus *Hevea*, particularly *H. brasiliensis* and *H. guianensis*. Asian countries occupy maximum area of rubber cultivation among all rubber growing countries. History of cultivation of rubber in India dates back to 1878, when it was first cultivated as a forest crop in some states [2]. Pure Indian-rubber is whitish and semi-transparent. Rubber is used in industrial factories for producing rubber sheets which are further used in seals for engine's flexibility, doctor's equipments, sports wears and toys. Large quantity of rubber is used in tire industries. Perhaps, it is the only natural raw material

available for making delicate parts of machineries and life saving equipments.

3-hydroxy-3-methylglutaryl-CoA reductase (HMGR) is a critical enzyme of mevalonic acid (MVA) pathway [3]. Structural changes in HMG-COA may result in to disruption of the pathway ultimately declining rubber production. HMG-COA plays a crucial role in the pathway, producing natural rubber as end product. To our knowledge the 3D structure of this enzyme is unavailable in protein model database (<http://www.caspar.it/PMDB>) [4]. Hence the objective of this work was to model the 3D structure of this enzyme and its computational characterization. The study will be valuable to understand the structural features and molecular interactions of HMGR enzyme and will raise the prospects of its commercial or academic use.

### MATERIALS AND METHODS

#### Retrieval of target and template sequence

The amino acid sequence of *H. brasiliensis* (UniProtKB Id: P29058) was retrieved in fasta format. Sequence homologies were obtained

with BLASTp 2.2.24 [5] setting default parameters. Two hits among the homologous sequences were selected as templates based on query coverage, lower resolution, lowest E-value and 50% above identity.

#### **Physiochemical and functional characterization**

Sequence length, molecular weight, theoretical isoelectric point (pI value), total number of negatively (Asp+Glu) and positively (Arg+Lys) charged residues, extinction coefficients [6], instability index [7], aliphatic index [8] and grand average of hydropathicity (GRAVY) [9] were computed using ExPASy's proteomics server (<http://expasy.org/cgi-bin/protparam>) [10]. Secondary structural features were predicted with Self Optimized Prediction Method from Alignment (SOPMA) [11].

#### **Model building**

Complex of the catalytic portion of human HMG-CoA reductase (PDB Id:1DQ8) and another template pyrrole-based, hepatoselective HMG-CoA reductase inhibitor (PDB Id:2Q1L) were selected as templates to generate the three dimensional model of HMGR enzyme of *H.brasiliensis*. Modeller 9v9 [12] was used to generate a total of 10 models of this target and the generated models were visualized with Chimera [13].

#### **Model refinement and evaluation**

Among the constructed models, the model with low Discrete Optimized Protein Energy (DOPE) score was subjected to energy minimization using Chimera. Using the Structural Analysis and Verification Server (SAVES) (<http://nihserver.mbi.ucla.edu/SAVES/>) stereochemical quality and accuracy of the model was evaluated with PROCHECK by Ramachandran plot analysis [14]. Model refinement was performed in Modeller 9v9 till a best refined model was obtained. The model thus selected on the basis of number of residues in core, allowed, generously allowed and disallowed regions was further put to analysis by

VERIFY 3D [15], a FORTAN based program ERRAT [16] and PROtein Volume Evaluation (PROVE) [17] program. The refined and evaluated model was ready for observation.

#### **Analysis and Functional elucidation of modeled HMGR enzyme**

The modelled three dimensional structure of HMGR was analyzed using Chimera. Physiochemical analysis of HMGR enzyme was carried out by ProtParam. Z-score was calculated using interactive ProSA web service (<https://prosa.services.came.sbg.ac.at/prosa.php>) for the recognition of errors in three-dimensional structure which indicated model quality and total energy deviation of the structure with respect to energy distribution derived from random conformations [18]. Further the functional attributes of the predicted structure was analyzed by accessing 3D2GO (<http://www.sbg.bio.ic.ac.uk/phyre/pfd/index.html>) server which used the sequence and structural information to predict the Gene Ontology terms of the protein.

#### **RESULTS**

The 210 amino acids long fasta formatted sequence of *H. brasiliensis* (UniProtKB Id: P29058) was retrieved. The sequence of this enzyme was subjected to Blastp 2.2.24 (Altschul *et al.*, 1990) analysis against the PDB database for selecting the appropriate templates. Two templates were considered. The first template was a complex of the catalytic portion of human HMG-CoA reductase with HMG and CoA reductase (PDB Id: 1DQ8) and the other was pyrrole-based hepatoselective HMG-CoA reductase inhibitor (PDB Id: 2Q1L) shared 64% residue identity with the target sequence. Two templates were selected to model the structure of HMGR, which had good query coverage of 95%, E-value of 1e-90, a lower resolution of 2.0 Å and 64% of identity (Table 1). The secondary structural features were predicted by SOPMA showing significant percentage of alpha helices 50.00%. Extended strands, beta turns and

random coils were 14.29%, 7.14% and 28.57% respectively (Fig. 1).

#### **Physiochemical characterization**

Physiochemical properties of the target were computed using ProtParam tool (Table 2). The computed isoelectric point (pI value), extinction coefficient, instability index, aliphatic index and Grand average of hydropathicity (GRAVY) value of the enzyme was 5.80, 8730-8480, 33.76, 102.67 and 0.293 respectively.

#### **Model building, refinement and evaluation**

Ten models were generated by Modeller 9v9. The model with low DOPE score value was taken in to refinement. Then modeled enzyme was subjected for energy minimization with appropriate Chimera inbuilt force field with initial energy minimization value of 31663.111536 (Fig. 2).

#### **PROCHECK analysis**

Ramachandran plot generated by PROCHECK showed that 92.5% of the residues were located in the most favored regions, 6.8% in additional allowed regions another 0.7% in generously allowed regions and there were no residues in disallowed regions (Fig. 3).

#### **Verify 3D, ERRAT and PROVE analysis**

In the Verify3D analysis, it was found that 85.12% of the residues assigned a 3D-1D score of more than 0.2. The amino acid environment was evaluated using ERRAT plots, ERRAT showed an overall quality factor of 79.511. Analysis of the entire structure using the PROVE program gave an average Z-score of 0.843, indicating reasonable packing quality.

#### **ProSA analysis**

The graph generated by ProSA showed overall model quality and location of the Z-score for the structure. Z-score indicated model quality as well as measures the deviation of the total energy of the structure with respect to an energy distribution derived from random conformations (Fig. 4 and 5). The position of HMGR model in NMR region showed Z-score value of -8.63.

#### **3D2GO analysis**

The 3d2GO server produced Gene Ontology (GO) based metabolic process and oxidoreductase activity up to the confidence level of greater than 0.80. This ensured the possible statistical accuracy of the model.

#### **DISCUSSION**

Three dimensional structure for HMGR enzyme was not available in Protein Model Portal; hence, it was chosen for modeling. HMGR is involved in rubber biosynthesis. An increase of this enzyme, leads to increase in rubber production [19]. The computed isoelectric point (pI value) for HMGR was 5.80 indicating its acidic character. Isoelectric point is a pH at which a protein carries no net charge. It is of significance in protein purification as it is the pH at which solubility is often minimal and mobility in an electrofocusing system is zero. The extinction coefficient was calculated as 8730-8480 based on the molar extinction coefficient of Tyr, Trp and Cys residues. This measure indicates how much light is absorbed by a protein at a particular wave length. Computed value of instability index of HMGR was 33.76. The value for instability index is less than 40 contributing stability to the over all 3D structure of the protein [7]. Instability index relies upon the occurrence of certain dipeptides along the length of the enzyme. Aliphatic index was computed as 102.67. Higher aliphatic index of HMGR also indicated structural stability. The aliphatic index refers to the relative volume of a protein that is occupied by aliphatic side chains. An increase in the aliphatic index increases the thermostability of enzyme. Grand average of hydropathicity (GRAVY) value of this enzyme was found to be 0.293. A positive GRAVY value for HMGR designates it to be hydrophobic in nature.

It is believed that 90% or more residues in the allowed region of the Ramachandran plot support the reliability of the model. In the current study Ramachandran plot generated by

PROCHECK showed that 92.5% of the residues were located in the most favored regions. For the quality of the model to be considered satisfactory, it is expected to have a Verify3D score greater than 80%. In the Verify3D analysis, it was found that 85.12% of the residues assigned a 3D-1D score of more than 0.2. So the predicted model was compatible with its amino acid sequence. ERRAT showed the distribution of different types of atoms with respect to one another in the protein model and is used for making decision about its reliability. ERRAT showed an overall quality factor of 79.511, a result expected for crystallographic models with resolution  $>2.5$  Å. Analysis of the entire structure using the PROVE program gave an average Z-score of 0.843, indicating reasonable packing quality. In PROVE, Z-scores above 4.0 and below -4.0 suggest the presence of significant errors in the structure in terms of packing of buried atoms. PROSA was used to check three dimensional model of protein for potential errors. The program displayed two characteristics of the input structure, its z-score and a plot of its residue energies. The energy plot showed the local model quality by plotting energies as a function of amino acid sequence position. In general it accepted that positive values correspond to problematic or erroneous parts of a model. Our analysis showed that the predicted structure quality was acceptable. The position of HMGR model in NMR region showing Z-score value of -8.63 indicated the high reliability of the structure. A comparable energy plot for both the target and template structures was given in Fig. 5.

## CONCLUSION

In the present study, an attempt has been made to construct a model of HMGR enzyme, as knowledge of the three-dimensional structure is essential for a better understanding of its functional attributes. We hope that this model will inspire new experimental efforts in this

area; specifically, the critical assessment of our computational approach can be readily tested for their biochemical relevance. The present study will be useful to understand the molecular architecture and function of HMGR enzyme and will be of importance for commercial and academic understanding of this enzyme.

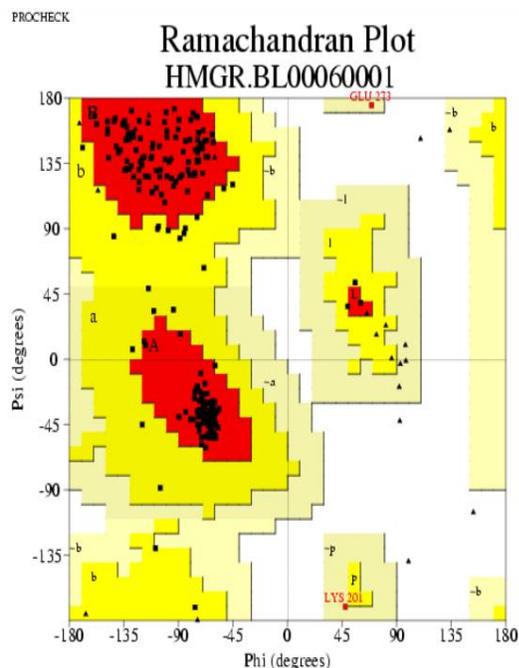
## ACKNOWLEDGEMENT

The authors are grateful to Department of Biotechnology, Ministry of Science and Technology, Government of India for funding Bioinformatics Infrastructure Facility (BIF) at North Orissa University, Odisha.

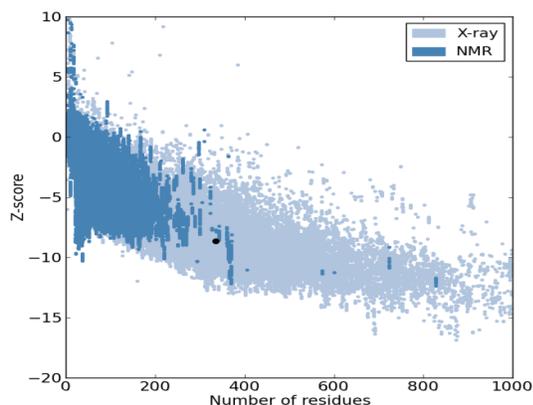
## REFERENCES

1. Uthup T K, Ravindran M, Bini K and Saha T. 2011. Divergent DNA Methylation Patterns Associated with Abiotic Stress in *Hevea brasiliensis*. *Molecular Plant* 4(6): 996-1013.
2. Varghese Y A and Abraham S T. 2007. In : *Biodiversity in Horticultural Crops*, (Ed) Peter K V and Abraham Z. Daya Publishing House, New Delhi, India, 1: 340-364.
3. Friesen J A and Rodwell V W. 2004. The 3-hydroxy-3-methylglutaryl coenzyme-A (HMG-CoA) reductases. *Genome Biol.* 5: 248-255.
4. Castrignano T, De Meo P D, Cozzetto D, Talamo I G and Tramontano A. 2005. The PMDB Protein Model Database. *Nucleic Acids Res* 34: 306-309.
5. Altschul S F, Gish W, Miller W, Myers E W and Lipman D J. 1990. Basic local alignment search tool. *J. Mol. Biol* 215: 403-410.
6. Gill S C and Hippel P H V. 1989. Calculation of Protein Extinction Coefficients from Amino Acid Sequence Data. *Analytical Biochem.* 182: 319-326.
7. Guruprasad K, Reddy BVB and M W. Pandit. 1990. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting *in vivo* stability of a protein from its primary sequence. *Protein Eng.* 4(2): 155-161.
8. Ikai A. 1980. Thermostability and Aliphatic Index of Globular Proteins. *J. Biochem.* 88 (6): 1895-1898.
9. Kyte J and Doolittle R F. 1982. A simple method for displaying the hydrophobic character of a Protein. *J. Mol. Biol.* 157(1): 105-132.
10. 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*

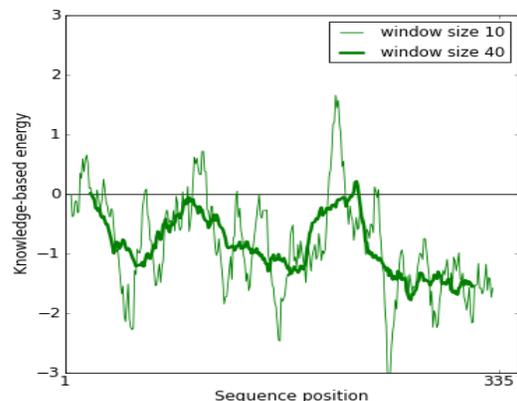




**Fig. 3:** Ramachandran plot analysis for HMGR. The plot statistics are: total number of residues 335 with 92.5% residues in the core region (red); 6.8 % residues in allowed (yellow); 0.7% in generously allowed (light yellow) ; 0.0% residue in the disallowed region (white) and Number of end residue 1. Number of glycine residues (labeled as G) is 25 and Number of Pro residues is 16.



**Fig. 4:** ProSA-web server analysis of HMGR. ProSA-web z-scores of all protein chains in PDB determined by X-ray crystallography (light blue) or NMR spectroscopy (dark blue) with respect to their length. The z-scores of HMGR is highlighted as large dot.



**Fig. 5:** Energy plot

**Table 1:** Templates selected from BLASTp result

GenPept Acc.no.	E-value	Score	Query coverage	% identity	% Gap
1DQ8 A	1e-90	273	95%	64%	0
2Q1L A	1e-90	272	95%	64%	0

**Table 2:** Properties of HMGR enzyme as predicted by ProtParam program

1	Molecular weight	21563.8
2	Theoretical pI	5.80
3	Total no. of -ve residues (Asp + Glu)	20
4	Total no. of +ve residues (Arg + Lys)	17
5	Extinction coefficients $M^{-1} cm^{-1}$ at 280 nm	8730-8480
6	Instability index	33.76
7	Aliphatic index	102.67
8	GRAVY	0.293