Indexed, Abstracted and Cited: ISRA Journal Impact Factor, International Impact Factor Services (IIFS), Directory of Research Journals Indexing (DRJI), International Institute of Organized Research and Scientific Indexing Services, Cosmos Science Foundation (South-East Asia), Einstein Institute for Scientific Information {EISI}, Directory of Open Access Scholarly Resources,

citefactor.org journals indexing

Directory Indexing of International Research Journals

World Journal of Biology and Medical Sciences

Published by Society for Advancement of Science®

ISSN 2349-0063 (Online/Electronic)

Volume 2, Issue- 4, 57-65, October to December, 2015

WJBMS 2/04/40/2015

www.sasjournals.com

A Double Blind Peer Reviewed Journal / Refereed Journal

wjbmedsc@gmail.com/wjbms.lko@gmail.com

RESEARCH PAPER

Received: 14/07/2015 Revised: 02/09/2015 Accepted: 03/09/2015

Protein Modeling Studies on RuBisCO Enzyme of *Rhododendron maddeni*

Gloria Nyenthang, *Salam Pradeep Singh, Chitta Ranjan Deb, **Lakshmi Narayan Kakati, and ***Bolin Kumar Konwar

Department of Botany, Nagaland University, Lumami-798627, Nagaland, India

***Bioinformatics Infrastructure Facility, Nagaland University, Lumami-798627, Nagaland, India**

****Department of Zoology, Nagaland University, Lumami-798627, Nagaland, India**

*****Department of Molecular Biology and Biotechnology, Tezpur University, Tezpur-784028, Assam, India**

ABSTRACT

Rhododendron maddeni is a plant species native to Bhutan and a very common in North Eastern part of India. However the in accessibility to exact habitat location has created great hindrance to the scientific community. Hence, in this article, we have retrieved the amino acid sequence of RuBisCo enzyme of Rhododendron maddeni and its three dimensional structure was generated. The generated 3D structure can be used to understand the various functions and biological pathway of this particular plant species. The 3D structure was predicted using the I-TASSER protein 3D structure modeling server. Further, the Ramachandran plot of the generated structure was plotted as well as the electrostatic surface and energy surface of the structure was mapped.

Keywords: Rhododendron, RuBisCo, I-TASSER and Modeling.

INTRODUCTION

Ribulose-1, 5-bisphosphate carboxylase / oxygenase (RuBisCO) is one of the most abundant proteins that can be found on earth. In the first major step of carbon fixation RuBisCO is involved in a process by which atmospheric carbon dioxide is converted by plants to energy-rich molecules such as glucose. It also catalyzes the carboxylation of ribulose-1, 5-bisphosphate [**Dhingra et al., 2004 and Feller et al., 2004**].

RuBisCO is a biologically important enzyme because it catalyzes the primary chemical reaction by which inorganic carbon enters the biosphere. RuBisCO is a major contributor to global carbon fixation while the reductive acetyl CoA pathway, the 3-hydroxypropionate cycle, or the reverse Krebs cycle, these pathways are relatively smaller contributors to global carbon fixation which are used by many autotrophic bacteria and archaea to fix carbon [**Curmi et al., 1992 and Dhingra et al., 2004**]. Phosphoenolpyruvate carboxylase, also only fixes carbon temporarily. As mentioned earlier, reflecting its importance, RuBisCO is the most abundant protein in leaves, accounting for 50% of soluble leaf protein in C3 plants (20–30% of total leaf nitrogen) and 30% of soluble leaf protein in C4 plants (5–9% of total leaf nitrogen) [**Feller et al., 2004**]. Given its important role in the biosphere which make the researcher's to know more of the genetic engineering of RuBisCO [**Yoon et al., 2001, Farazdaghi, 2011 and Portis, 2003**].

During carbon fixation, the substrate molecules for RuBisCO are ribulose-1, 5-bisphosphate, carbon dioxide (distinct from the "activating" carbon dioxide) [**Jin et al., 2004**]. RuBisCO also catalyses a reaction between ribulose-1, 5-bisphosphate and molecular oxygen (O2) instead of carbon dioxide (CO2). The product of the carboxylase reaction is a highly unstable six-carbon phosphorylated intermediate known as 3-keto-2-carboxyarabinitol-1, 5-bisphosphate, when carbon dioxide is the substrate, because of which it decays virtually instantaneously into two molecules of glycerate-3-phosphate. Larger molecules, such as glucose can be used to produce 3-phosphoglycerate [**Andralojc et al., 1994 and Crafts-Brandner and Salvucci, 2000**]. When molecular oxygen is the substrate, the products of the oxygenase reaction are phosphoglycolate and 3-phosphoglycerate. Phosphoglycolate is recycled through a sequence of reactions called photorespiration, and it involves enzymes and cytochromes which are located in the mitochondria and peroxisomes [**Khan et al., 1999**]. In this process, two molecules of phosphoglycolate are converted to one molecule of carbon dioxide and one molecule of 3-phosphoglycerate, which can reenter the Calvin cycle. Molecules such as glycine can be produced by plants which can retain some of the phosphoglycolate which enters the pathway. At ambient levels of carbon dioxide and oxygen, the ratio of the reactions is about 4 to 1, which results in a net carbon dioxide fixation of only 3.5 [**Salvucci et al., 2001, Zhang et al., 2002 and Marcus and Gurevitz, 2000**]. Thus, the inability of the enzyme to prevent the reaction with oxygen greatly reduces the photosynthetic capacity of many plants. Some plants, many algae, and photosynthetic bacteria have overcome this limitation by devising means to increase the concentration of carbon dioxide around the enzyme, including C4 carbon fixation, crassulacean acid

metabolism (CAM), and the use of pyrenoid [**Spreitzer and Salvucci, 2002 and Parry et al., 2003**].

RuBisCO fixes only 3-10 carbon dioxide molecules each second per molecule of enzyme which is relatively very slow as compared to some enzymes which can carry out thousands of chemical reactions each second [**Salvucci et al., 2001 and Zhang et al., 2002**]. However, the reaction catalyzed by RuBisCO is the primary rate-limiting factor of the Calvin cycle during the day. Nevertheless, under most conditions, and when light is not otherwise limiting photosynthesis, the speed of RuBisCO responds positively to increasing carbon dioxide concentration [**Whitney and Andrews, 2001**]. However a biochemical model [**Zhang et al., 2002**], is developed to represent the effects of the steps quantitatively to broaden our knowledge.Since carboxylation or fixation of CO2 is possible only after the synthesis of enediol, thus it is suggested that the role of RubisCO is to produce enediol that is carboxylase and oxygenase (EnCO) [**John Andrews, and Whitney, 2003**]. Accordingly, since RubisCO is neither carboxylase nor oxygenase, it is also called enolase-phosphglycerase (EPGase). It requires the consumption of ATP by activase to remove the inhibitory RuBP, CA1P, and the other inhibitory substrate analogs which are also inhibited by the presence of ADP [**Tcherkez et al., 2006**]. Thus, activase activity depends on the ratio of these compounds in the chloroplast stroma. Furthermore, in most plants, the sensitivity of activase to the ratio of ATP/ADP is modified by the stromal reduction/oxidation (redox) state through another small regulatory protein, thioredoxin [**Wildman, 2002**]. In this manner, the activity of activase and the activation state of RuBisCO can be modulated in response to light intensity and, thus, the rate of formation of the ribulose 1, 5-bisphosphate substrate [**Portis, and Parry, 2007**].

MATERIALS AND METHODS

The amino acid sequence of the RuBisCO of *Rhododendron maddeni* was analyzed using CLC Sequence Viewer. The amino acid sequence of RuBisCO of *Rhododendron maddeni* was retrieved from the NCBI Genbank Database. The Protein Data Bank (PDB) is a repository for the 3-D structural data of large biological molecules, such as proteins and nucleic acids. The data, typically obtained by X-ray crystallography or NMR spectroscopy and submitted by biologists and biochemists from around the world, are freely accessible on the Internet via the websites of its member organizations (PDBe, PDBj and RCSB). The PDB is overseen by an organization called the Worldwide Protein Data Bank, wwPDB. The PDB is a key resource in areas of structural biology, such as structural genomics. Most major scientific journals, and some funding agencies, such as the NIH in the USA, now require scientists to submit their structure data to the PDB. If the contents of the PDB are thought of as primary data, then there are hundreds of derived (i.e., secondary) databases that categorize the data differently. For example, both SCOP and CATH categorize structures according to type of structure and assumed evolutionary relations; GO categorize structures based on genes. A PDB BLAST search was performed against the RUBISCO of *Rhododendron maddeni* resulting with a similarity with RUBISCO of *Oryza sativa*.

I-TASSER server was chosen for predicting the 3D structure. The I-TASSER server is an on-line platform for protein structure predictions as well as its function. The 3D models are built based on multiple-threading alignments by LOMETS and iterative template fragment assembly

simulations. The functional insights are derived by matching the 3D models with BioLiP protein function database. The reliability of the I-TASSER is that it was ranked as the No 1 server for protein structure prediction in CASP7, CASP8, CASP9, and CASP10 experiments. The server provides the most accurate structural and function predictions using state-of-the-art algorithms.

RESULTS AND DISCUSSION

The amino acid sequence was analyzed and its amino acid composition is shown in Table 1. The amino acid sequence has a molecular weight of 23.887 kDa, Isoelectric point of 7.43 and Aliphatic index of 75.896. The B-factor is a value to indicate the extent of the inherent thermal mobility of residues/atoms in proteins. In I-TASSER, this value is deduced from threading template proteins from the PDB in combination with the sequence profiles derived from sequence databases. The reported B-factor profile in the figure below corresponds to the normalized B-factor of the target protein, defined by B=(B'-u)/s, where B' is the raw B-factor value, u and s are respectively the mean and standard deviation of the raw B-factors along the sequence.

Table 1. Amino acid composition of RUBISCO enzyme.

Amino acid	Count	Frequency	Freq. in Plants
Alanine (A)	16	0.075	0.072
Cysteine (C)	4	0.019	0.016
Aspartic Acid (D)	13	0.061	0.05
Glutamic Acid (E)	13	0.061	0.061
Phenylalanine (F)	10	0.047	0.047
Glycine (G)	18	0.085	0.072
Histidine (H)	2	0.009	0.023
Isoleucine (I)	8	0.038	0.06
Lysine (K)	13	0.061	0.06
Leucine (L)	18	0.085	0.095
Methionine (M)	2	0.009	0.024
Asparagine (N)	8	0.038	0.043
Proline (P)	14	0.066	0.048
Glutamine (Q)	5	0.024	0.036
Arginine (R)	13	0.061	0.053
Serine (S)	7	0.033	0.075
Threonine (T)	16	0.075	0.052
Valine (V)	15	0.071	0.067
Tryptophan (W)	3	0.014	0.013
Tyrosine (Y)	14	0.066	0.033

The three dimensional structure of RUBISCO of *Rhododendron maddeni* was predicted. The predicted secondary structure of *R. maddeni* is shown in Fig 1 and that of solvent accessibility of *R. maddeni* is shown in Fig 2.

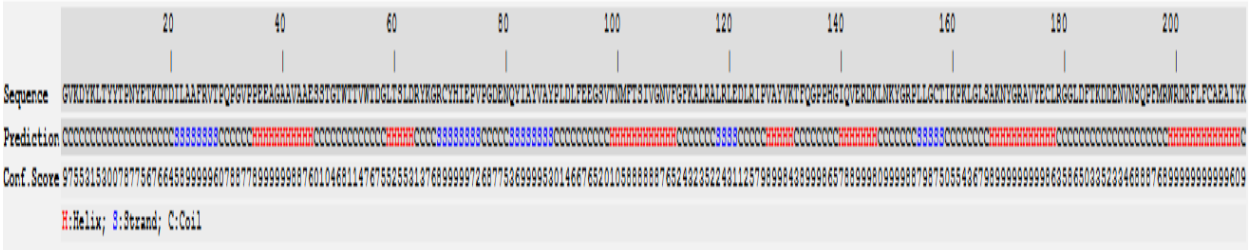

Fig. 1. Predicted secondary structure of of *R. maddeni*.

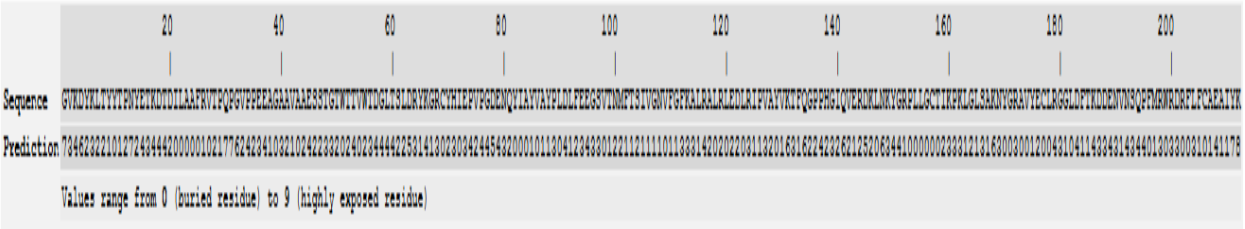

Fig. 2. Predicted Solvent Accessibility of *R. maddeni.*

The local accuracy of the predicted protein which is defined as the distance deviation (in Angstrom) between residue positions in the model and the native structure was estimated using support vector regression that makes use of the coverage of threading alignment, divergence of I-TASSER simulation decoys, and sequence-based secondary structure and solvent accessibility predictions. Large-scale benchmark tests show that the estimated local accuracy has an average error of 2.21 Angstrom and the Pearson's correlation coefficient between estimated and actual error is 0.7. The Normalized B-factor which is defined as z-score-based normalization of the raw B-factor values is shown in Fig. 3. The normalized B-factor (called B-factor profile, BFP) is predicted using a combination of both template-based assignment and profile-based prediction. Based on the distributions and predictions of the BFP, residues with BFP values higher than 0 are less stable in experimental structures.

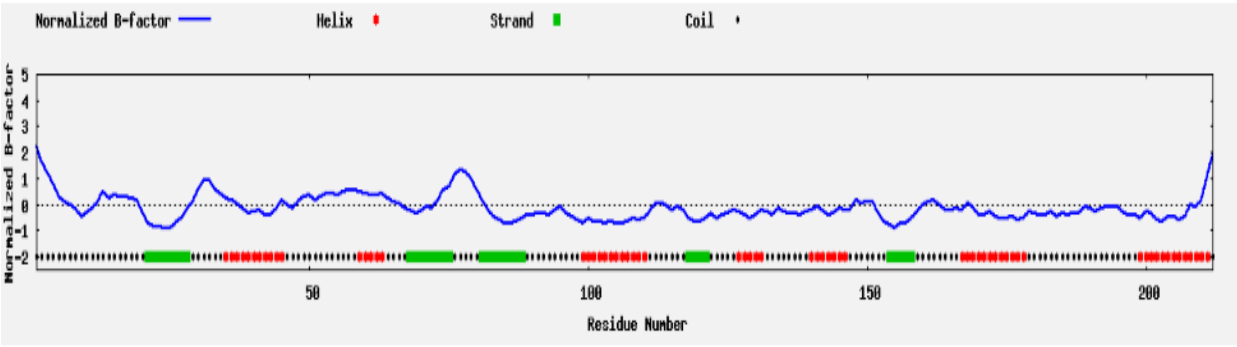

Fig. 3. Predicted normalised B-factor of *R. maddeni.*

A three dimensional model were generated using the template shown in Table 2 and alignment is shown in Fig 4. I-TASSER generates full length model of proteins by excising continuous fragments from threading alignments and then

reassembling them using replica-exchanged Monte Carlo simulations. Low temperature replicas (decoys)generated during the simulation are clustered by SPICKER and top five cluster centroids are selected for generating full atomic models. The cluster density is defined as the number of structure decoys at a unit of space in the SPICKER cluster. A higher cluster density means the structure occurs more often in the simulation trajectory and therefore signifies a better quality model.

The I-TASSER modeling starts from the structure templates identified by LOMETS from the PDB library. LOMETS is a meta-server threading approach containing multiple threading programs, where each threading program can generate tens of thousands of template alignments. I-TASSER only uses the templates of the highest significance in the threading alignments, the significance of which are measured by the Z-score, i.e. the difference between the raw and average scores in the unit of standard deviation. The templates in this section are the 10 best templates selected from the LOMETS threading programs. Usually, one template of the highest Z-score is selected from each threading program, where the threading programs are sorted by the average performance in the large-scale benchmark test experiments.

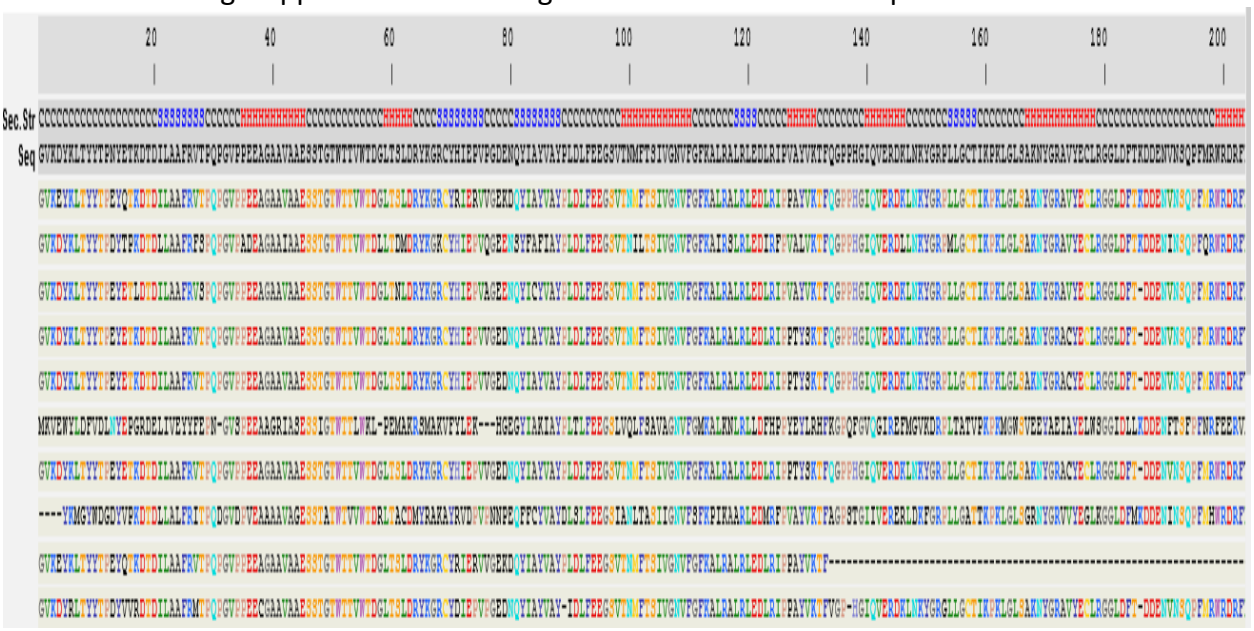

Fig. 4 Sequence alignment of the top templates used by I-TASSER

Table 2. Templates used in modeling *R. maddeni RUBISCO.*

Rank[a]	PDB Hit	Iden1[b]	Iden2[c]	Cov[d]	Norm. Z score[e]
1	4rubA	0.95	0.95	1.00	2.51
2	IrscA	0.84	0.84	1.00	3.21
3	8rucA	0.96	0.96	1.00	2.94
4	1wddA	0.96	0.95	1.00	3.86
5	1wddA	0.96	0.95	1.00	2.76
6	2d69A	0.39	0.41	0.98	2.17
7	1wddA	0.96	0.95	0.99	4.19
8	1bxnA	0.64	0.63	0.98	4.44
9	4rubA	0.92	0.58	0.63	1.88
10	1gk8A	0.93	0.92	0.99	5.08

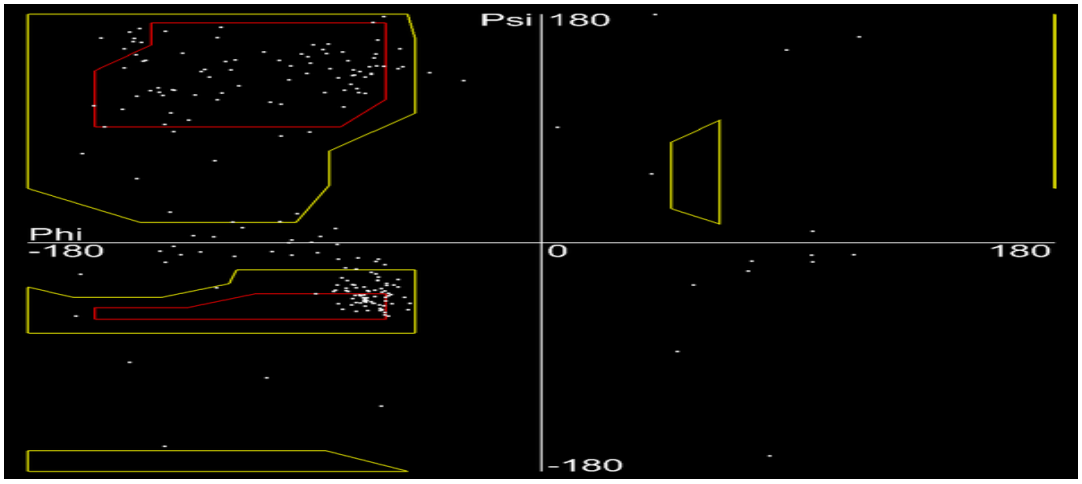

Fig. 5 Ramachandran plot of *R. maddeni RUBISCO*

The snapshots of the backbone structure of the generated model of *R. maddeni* RuBisCO is shown in Fig. 6 A. While the electrostatic surface view of the generated model is shown in Fig 6 B. Also the energy map depicting the possible Steric favorable, Hydrogen acceptor and donor favorable and electrostatic favorable site is shown in Fig. 7.

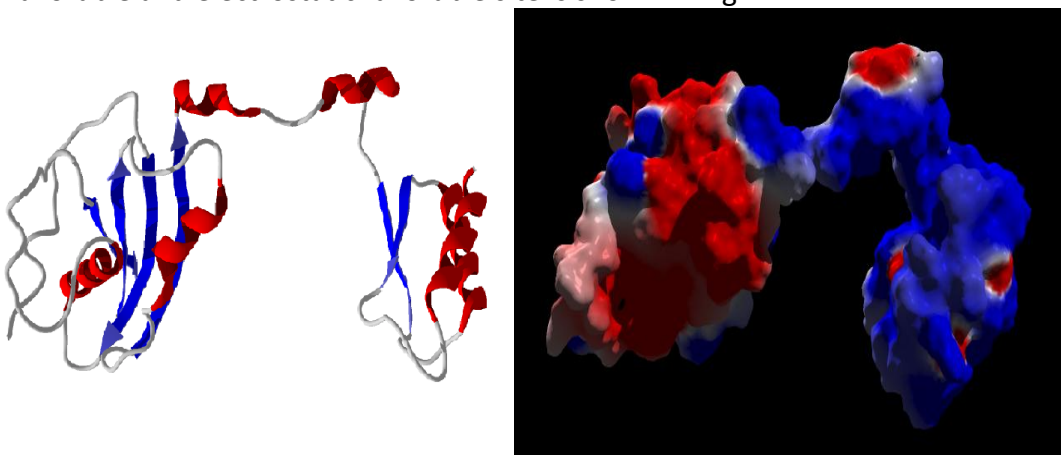

Fig. 6 (A) Secondary structure of RuBisCO of *R. maddeni*. (B) Electrostatic surface structure of RuBisCO of *R. maddeni*

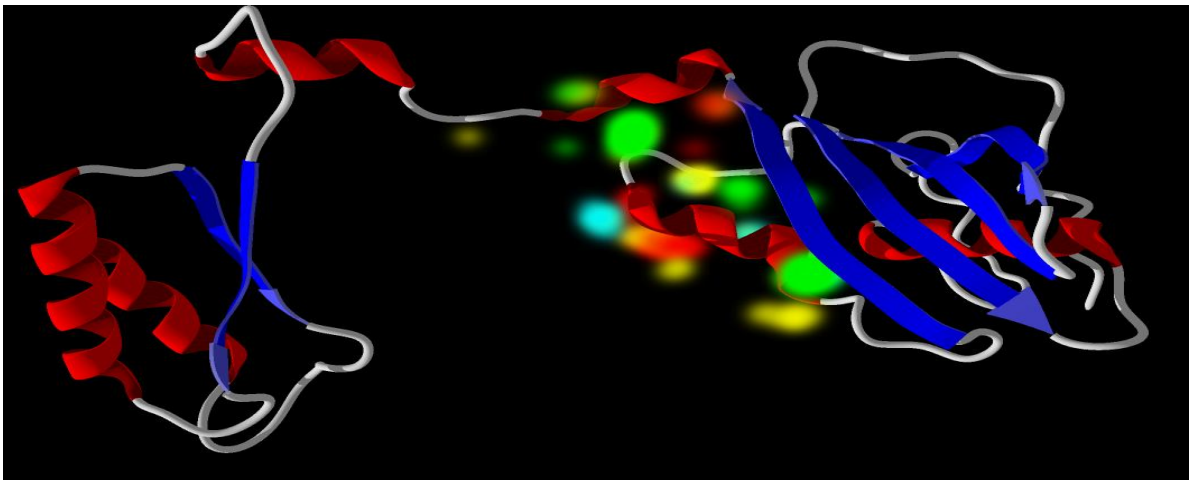

Fig. 7. Steric favorable, Hydrogen acceptor and donor favorable and electrostatic favorable site of Rubisco enzyme.

(a) Rank of templates represents the top ten threading templates used by I-TASSER.
(b) Ident1 is the percentage sequence identity of the templates in the threading aligned region with the query sequence.
(d) Ident2 is the percentage sequence identity of the whole template chains with query sequence.
(d) Cov represents the coverage of the threading alignment and is equal to the number of aligned residues divided by the length of query protein.
(e) Norm. Z-score is the normalized Z-score of the threading alignments. Alignment with a Normalized Z-score >1 mean a good alignment and vice versa.
Ramachandran plot analysis was carried out using VEGA for the predicted model. Ramachandran plot is a way to visualize backbone dihedral angles ψ against ϕ of amino acid residues in protein structure. A Ramachandran plot can be used in two different ways. One is to show in theory which values, or conformations, of the ψ and ϕ angles are possible for an amino-acid residue in a protein. A second is to show the empirical distribution of data points observed in a single structure in usage for structure validation. The Ramachandran plot of the final refined model is shown in Fig. 5.

CONCLUSION

The investigation comes to a conclusion with the following points:-
1. The amino acid sequence of RuBisCO enzyme of *R. maddeni* was retrieved from the NCBI Gen Bank database and its sequence was analyzed.
2. The 3D dimensional structure of RuBisCO enzyme of *R. maddeni* was predicted.
3. The secondary structure alpha helix and beta sheets were predicted for the said enzyme.
4. The electrostatic surface and energy map was predicted too.
5. This information on the structure prediction of RuBisCO enzyme of *R. maddeni* would aid in understanding its functions such as the step of carbon fixation, carboxylation of RuBP in various *Rhodendron* sps.

ACKNOWLEDGEMENT

The authors would like to acknowledge the Department of Biotechnology, Ministry of Science and Technology, Government of India, New Delhi for providing Bioinformatics Infrastructure Facility at Nagaland University, Lumami, Nagaland, India

REFERENCES

Dhingra, A., Portis, A.R. and Daniell, H. (2004). Enhanced translation of a chloroplast-expressed RbcS gene restores small subunit levels and photosynthesis in nuclear RbcS antisense plants. *Proc. Natl. Acad. Sci*. U.S.A. 101 (16): 6315–20.

Feller, U., Anders, I. and Mae, T. (2008). Rubiscolytics: fate of Rubisco after its enzymatic function in a cell is terminated. *J. Exp. Bot.* 59 (7): 1615–24.

Curmi, P.M., Cascio, D., Sweet, R.M., Eisenberg, D. and Schreuder, H. (1992). Crystal structure of the un activated form of ribulose-1, 5-bisphosphate carboxylase/oxygenase from tobacco refined at 2.0-A resolution. *J. Biol. Chem*. 267 (24): 16980–9.

Dhingra, A., Portis, A.R. and Daniell, H. (2004). Enhanced translation of a chloroplast-expressed RbcS gene restores small subunit levels and photosynthesis in nuclear RbcS antisense plants. *Proc. Natl. Acad. Sci.* U.S.A. 101 (16): 6315–20.

Yoon, M., Putterill, J.J., Ross, G.S. and Laing, W.A. (2001). Determination of the relative expression levels of rubisco small subunit genes in Arabidopsis by rapid amplification of cDNA ends. *Anal. Biochem*. 291 (2): 237–44.

Farazdaghi, H. (2011). The single-process biochemical reaction of Rubisco: A unified theory and model with the effects of

irradiance, CO2 and rate-limiting step on the kinetics of C3 and C4 photosynthesis from gas exchange. *Bio Systems,* 103 (2): 265–284.

Portis, A.R. (2003). Rubisco activase — Rubisco's catalytic chaperone. *Photosyn. Res*. 75 (1): 11–27.

Jin, S.H., Jiang, D.A., Li, X.Q. and Sun, J.W. (2004). Characteristics of photosynthesis in rice plants transformed with an antisense Rubisco activase gene. *J. Zhejiang Univ. Sci*. 5 (8): 897–9.

Andralojc, P.J., Dawson, G.W., Parry, M.A. and Keys, A.J. (1994). Incorporation of carbon from photosynthetic products into 2-carboxyarabinitol-1-phosphate and 2-carboxyarabinitol. *Biochem. J*. 304 (3): 781–6.

Crafts-Brandner, S.J. and Salvucci, M.E. (2000). Rubisco activase constrains the photosynthetic potential of leaves at high temperature and CO2. *Proc. Natl. Acad. Sci*. U.S.A. 97 (24): 13430–5.

Khan, S., Andralojc, P.J., Lea, P.J. and Parry, M.A. (1999). 2'-carboxy-D-arabitinol 1-phosphate protects ribulose 1, 5-bisphosphate carboxylase/oxygenase against proteolytic breakdown. *Eur. J. Biochem*. 266 (3): 840–7.

Salvucci, M.E., Osteryoung, K.W., Crafts-Brandner, S.J. and Vierling, E. (2001). Exceptional sensitivity of Rubisco activase to thermal denaturation in vitro and in vivo. *Plant Physiol*. 127 (3): 1053–64.

Zhang, N., Kallis, R.P., Ewy, R.G. and Portis, A.R. (2002). Light modulation of Rubisco in Arabidopsis requires a capacity for redox regulation of the larger Rubisco activase isoform. *Proc. Natl. Acad. Sci*. U.S.A. 99 (5): 3330–4.

Marcus, Y. and Gurevitz, M. (2000). Activation of cyanobacterial RuBP-carboxylase/oxygenase is facilitated by inorganic phosphate via two independent mechanisms. *Eur. J. Biochem*. 267 (19): 5995–6003.

Spreitzer, R.J. and Salvucci, M.E. (2002). Rubisco: structure, regulatory interactions, and possibilities for a better enzyme. *Annu. Rev. Plant Biol*. 53: 449–75.

Parry, M.A., Andralojc, P.J., Mitchell, R.A., Madgwick, P.J. and Keys, A.J. (2003). Manipulation of Rubisco: the amount, activity, function and regulation. *J. Exp. Bot.* 54 (386): 1321–33.

Whitney, S.M. and Andrews, T.J. (2001). Plastome-encoded bacterial ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO) supports photosynthesis and growth in tobacco. *Proc. Natl. Acad. Sci.* U.S.A. 98 (25): 14738–43.

John Andrews, T. and Whitney, S.M. (2003). Manipulating ribulose bisphosphate carboxylase/oxygenase in the chloroplasts of higher plants". *Arch. Biochem. Biophys*. 414 (2): 159–69

Tcherkez, G.G., Farquhar, G.D. and Andrews, T.J. (2006). Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized". *Proc. Natl. Acad. Sci.* U.S.A. 103 (19): 7246–51.

Wildman, S.G. (2002). Along the trail from Fraction I protein to Rubisco (ribulose bisphosphate carboxylase-oxygenase). *Photosyn. Res*. 73 (1-3): 243–50.

Portis, A.R. and Parry, M.A. (2007). Discoveries in Rubisco (Ribulose 1,5-bisphosphate carboxylase/oxygenase): a historical perspective. *Photosyn. Res*. 94 (1):,121–43.

Corresponding author: Dr. Salam Pradeep Singh, Bioinformatics Infrastructure Facility, Nagaland University, Lumami-798627, Nagaland, India
salampradeep@gmail.com