Original Research Article **Bioinformatics**

International Journal of Pharma and Bio Sciences **ISSN 0975-6299**

MODELLING & DOCKING STUDIES OF WILD AND MUTANT RUBISCO ENZYME IN DUNALEILA SALINA

INAMUL HASAN MADAR[1], PRIYADARSHAN KATHIRVEL[1], SUDHA SUBRAMANIAN[2], HUSSAIN AL SSADH[3], ASHVINI DESAI[4], IFTIKHAR ASLAM TAYUBI[5], AMJAD HUSSAIN.A[6], AND S.K.M. HABEEB[7]*

[1]Department of Biotechnology, Bharathidasan University, Tiruchirappalli, 620024, India.
[2]Department of Animal Biotechnology, Madras Veterinary College, Chennai, 600007 India.
[3]School of Biological Sciences, University of Essex, Colchester, CO43SQ, United Kingdom.
[4]Department of Bioinformatics, School of Biosciences and Technology, Vellore Institute of Technology, Vellore, 632014, India.
[5]Faculty of Computer Science and Information Technology, Rabigh, King AbdulAziz University, Rabigh, 21911. Saudi Arabia.
[6]Department of Biochemistry, Islamiah College, Vaniyambadi, 635 752, Vellore Dist, India.
[7] Department of Bioinformatics, School of Bioengineering, SRM University, Kattankulathur, Chennai 603203, India*

ABSTRACT

Rubisco is an important enzyme in photosynthesis as well as photorespiration. Rubisco plays dual role of carboxylation and oxygenation and the type of reaction is determined by specific amino acids at the active site. The carbon fixation has been shown to be favored by the involvement of Lys-334 of Rubisco with enediol of RuBP. The role of lysine at position 334 of loop 6 of Rubisco has been experimentally studied in many species. In the present study, we attempted to predict the change encountered during modification of the amino acid Lys-334 in the active site of *D. salina* Rubisco. For this, we had modelled the wild type and mutant structure of Rubisco enzyme from *Dunaliella salina with* MODELLER 9v3 and evaluated the model using Procheck3.5.4. The modelled structure was then docked using Autodock with CABP, the substrate analogue of RuBP as Lys-334 at the apex of the loop has iconic interaction with it. The dock score of wild type and mutant type were -9.26 and -7.91 respectively. The result of algorithmic predictions of observed differences in the binding affinities of wild type and mutant structure of Rubisco. As our results also predicted the importance of Lys-334 in Rubisco of *D. salina.*

KEYWORDS: ***RubisCO, Dunaliellasalina, Photosynthesis, K334, CABP, Homology modeling, Molecular Docking, modeller 9v3.***

S.K.M. HABEEB*

Department of Bioinformatics, School of Bioengineering, SRM University, Kattankulathur, Chennai 603203, India

Received on: 03-04-2017
Revised and Accepted on : 23-05-2017
DOI: http://dx.doi.org/10.22376/ijpbs.2017.8.3.b210-216

INTRODUCTION

Ribulose-l,5-bisphosphate carboxylase/ oxygenase (Rubisco) catalyses the initial step of the reduction of atmospheric CO_2 in photosynthesis as well as photo respiratory carbon oxidation[1].Though Rubisco catalyses both reactions, the reaction with CO_2 than with O_2 is of agronomic importance. Rubisco is often rate limiting for photosynthesis in plants, it may be possible to improve photosynthetic efficiency by modifying Rubisco genes in plants to increase its catalytic activity and decrease the rate of the oxygenation activity[2]. The Rubisco manipulated such that to decrease the oxygen reactivity and enhanced reaction with CO_2 would increase the agricultural productivity.The carboxylation and oxygenation reactions are catalyzed at the same active site on the enzyme and the neither gas binds directly to the active site but reacts directly to the enediol of Ribulose Bisphosphate (RuBP)**3** [3]. Approaches that have begun to be investigated include expressing Rubisco genes from one organism in another organism, increasing the level of expression of Rubisco subunits, expressing Rubisco small chains from the chloroplast DNA and altering Rubisco genes try to increase specificity for carbon dioxide or otherwise increase the rate of carbon fixation[3]. Among these strategies, the identification of an amino acid or residue promoting catalytic activity in favor of CO_2 would be of great importance in manipulation. Many of the loop 6 residues of Rubisco are conserved, but only lysine 334 at the apex of the loop has ionic interactions with the transition state analogue 2-carboxyarabinitol 1, 5-bisphosphate (CABP) [3.] During carboxylation, lysine 334 polarizes both oxygen and carbon -dioxide, thereby enhancing the electrophilic status of the carbon atom of carbon -dioxide, promoting electrophilic attack by carbon-dioxide on C_2 of the enediol to form the 2-carboxy, 3-keto intermediate [4]. Although site-directed mutants of lysine 334 catalyzed enediolate formation, they were unable to catalyze the reaction of the enediolate with carbon-dioxide or form a stable complex with CABP [5].Thus lysine 334 is thought to play a specific role in stabilizing the transition state intermediates of both the carboxylation and oxygenation reactions, thereby facilitating the reaction between the gaseous substrate and the enediolate[3].*Dunaliellasalina* is a microalgae of economic interest because it is intensively cultured for the commercial production of both glycerol and the photo protective pigment ß -carotene[6,7]. However, the relative slow growth rate of *D. salina*made it non-convenient, one possible reason could be the rate of photosynthesis[8]. As a consequence of this, substantial reductions in cell volume and division rate in low CO_2 conditions, cells gave lower productivity of glycerol and ß-carotene[9].Hence identification of genes associated with photosynthesis in D.salina would increase production of its glycerol and pigment. In the present study, we attempt to computationally predict in *D. salina,* the effect of mutation in an amino acid (Lys-334) that has been shown to favour carboxylation of Rubisco. Modification of Rubisco genes to promote increased growth rate and carbon fixation could improve the cultivation of *D.salina* having economic importance.Prediction of protein function using computational support becomes an important tool as the gap between the increasing amount of sequences and the experimental characterization of the respective protein widens. It is essential in almost every area of protein research such as enzyme kinetics, ligand – protein binding studies, gene characterization and construction, structure based designing of drugs and rational designing of proteins. Homology modelling is a computational method used in the prediction of protein, which relies on the identification of one or more known protein structures likely to resemble the structure of the query sequence. Modeling of the protein is classified into four sequential steps such as template selection, target-template alignment, model construction and model assessment.

MATERIAL AND METHODS

Protein Knowledge Using Swissprot

The UniProtKB/Swiss-Prot is a protein sequence database that provides a high level of annotation, a minimal level of redundancy. The Rubisco protein sequence and details regarding the protein were obtained from this database. *Domain and Pattern Prediction* Pfam[10], contains information about protein domains and families. Pfam-A is the manually curated portion of the database that contains over 9,000 entries[11]. *Prosite* is a database of protein families and domains. It is based on the observation that, while there is a huge number of different proteins, most of them can be grouped on the basis of similarities in their sequences into a limited number of families. By analyzing the constant and variable properties of such groups of similar sequences, it is possible to derive a signature for a protein family or domain.ProRule builds on the domain descriptions of prosite, provides information about functionally and/or structurally critical amino acids, like active sites, binding sites, post-translational modification sites or disulfide bonds, to help function determination[12]. *PSI-BLAST* (Position Specific Iterative BLAST) and PHI-BLAST (Pattern Hit Initiated BLAST) is constructed (automatically) from a multiple alignment of the highest scoring hits in an initial BLAST search. The PSSM is generated by calculating position-specific scores for each position in the alignment. Highly conserved positions receive high scores and weakly conserved positions receive scores near zero

Choosing the template

PSI-BLAST iteratively updates their position-specific scoring matrixto successively identify more distantly related homolog. This family of methods has been shown to produce a larger number of potential templates and to identify better templates for sequences that have only distant relationships to any solved structure. When performing a blast search, a reliable first approach is to identify hits with a sufficiently low *E*-value.

Steps carried out in our study

The structure of wild type and mutant protein of Rubisco of *D. salina* was modelled using Modeller9v3 and the structure then was validated using Ramachandran plot.

Modelling structure of wild and mutant type Rubisco Template selection

The target protein was submitted to BLASTp and then template was chosen based on the percentage identity

with the target protein. The protein with the percentage identity greater than 35% was taken as template.

Domain and Pattern Prediction

The template and target protein sequences were submitted to Pfam[10], PROSITE Data Bases[13], to analyze the domains, pattern profile and atom files.

Homology modeling using Modeller

The Rubisco was modelled using 2V69 as template. The amino acid sequence of 2V69 was obtained from the PDB database and inserted in the script file of modeller.

Target template alignment

The target and template sequences was aligned in MODELLER 9V3, which runs with python.

Model Construction

The refined sequence-structure alignment obtained from MODELLER 9v3 was used to construct 3D models of the target with the help of the known structures of the template using MODELLER 9v3.

Model assessment

Model evaluation and refinement is highly necessary as model may contain some errors. The numerical values representing the residue specific environment in a given structure are computationally evaluated. These numerical values are called 3D profile scores. The errors found are corrected by remodeling regions by correcting the alignment. The tool used for model evaluation was Procheck.

Procheckv.3.5.4

Procheck checks the stereo chemical quality of a protein structure, producing a number of postscripts plots analyzing its overall residue-by-residue geometry.

Ramachandran Plot

Stereo chemical quality of the model was analysed by plotting the Ramachandran plot in the Rampage from the FUGUE server[14].

Docking

Docking was performed to predict the strength of association or binding affinity of the substrate CABP (an analogue of RuBP) to the wild / mutant type Rubisco of *D.salina*.The docking was done using Autodock, which consists of two main programs: AutoDock performs the docking of the ligand to a set of grids describing the target protein and AutoGrid pre-calculates these grids.

RESULTS

Target protein was modeled by submitting to BLASTp[15]. The template was chosen based on the identity score, which was about 93% for the target protein. The amino acid sequence of Rubisco from the organism *D. salina* and the template ID 2V69 were obtained from PDB.The template sequence was processed in different databases such as Pfam, PROSITE Data Bases and the pattern; profiles and the domains obtained were compared with the target sequence. The template sequence was found to have two domains Figure 1.

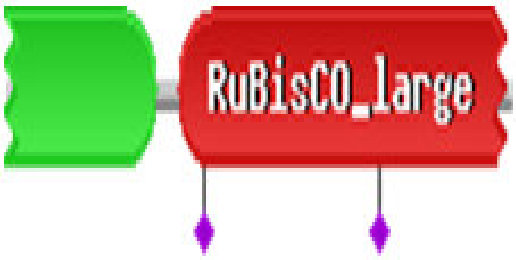


Figure 1
Pfam result of the target.

The template (PDB id 2V69) and the target similarity shown in BLAST result were as follows, Length=475 Score = 677 bits (1746), Expect = 0.0, Method: Compositional matrix adjust. Identities = 325/346 (93%), Positives = 335/346 (96%), Gaps = 0/346 (0%).

Homology modelling using Modeller

The target and the template sequences were aligned using MODELLER9v3 and the alignment was used to generate the 3D structure of Rubisco of *D.salina*, using the same tool. Each chain of the enzyme is modelled separately in MODELLER9v3 and all the PDB files were merged to get a hexadecameric structure of Rubisco Fig. 2.

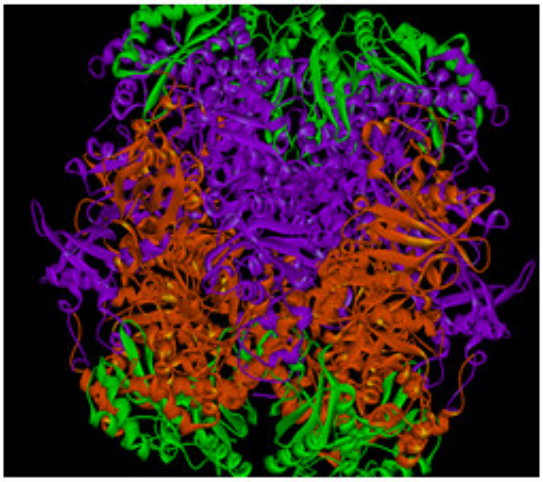
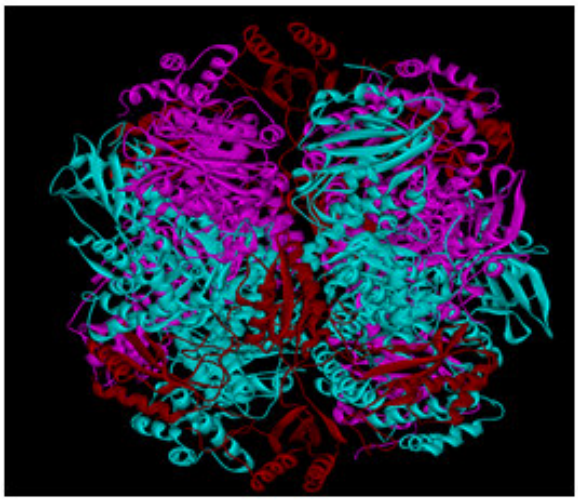


Figure 2
Structure of wild-type rubisco generated by modeller9v3 viewed in ds visualizer 2.0.



The mutant structure of Rubisco was modelled in almost similar way as the wild type with same template except for a change of single amino acid R for K at 334 Fig. 3.

Figure 3
Structure of mutant k334r protein rubisco generated by modeller9v3 viewed in ds visualizer 2.0.

Ramachandran plot using rampage

The stereo-chemical quality of the protein was assessed by plotting the Ramachandran plot using RAMPAGE[16]. This program is used for visualizing and assessing the Ramachandran plot of a protein structure. On the basis of a manually curated set of high-quality protein structures (from the Richardson's Group at Duke University) and a number of filters (such as B-factor cutoff and van der Waals clashes), reference phi/psi plots were derived for Gly, Pro, pre-Pro and general (other) residue types, and subdivided into "favored", "allowed" and "outlier" regions. Residues in the uploaded PDB file that fall into these subdivided regions are as follows for Wild type and mutant Rubisco.

Table 1
Ramachandran plot results in Wild type and the Mutant (K334R)

	Wild type Rubisco	Mutant K334R Rubisco
Number of residues in favoured region (~98.0% expected)	437(95.4%)	441 (96.3%)
Number of residues in allowed region (~2.0% expected)	17 (3.7%)	13 (2.8%)
Number of residues in outlier region	4 (0.9%)	4 (0.9%)

Docking Study

Ligand CABP(2-carboxyarabinitol 1,5-bisphosphate) was obtained from Pubchem and its 3D structure was generated using CORINA Fig. 4. Active site of the protein was predicted using PDB, sum of the residues of active site are THR135, LYS137, LYS139, ASP165, GLU166, HIS256, ALA258, HIS289, LYS296, LEU297, SER341, GLY342, GLY343. Then undergoes docking study with the help of autodock tools. After docking, both files are viewed in Accelrys DS visualize (Fig. 5 a & b)

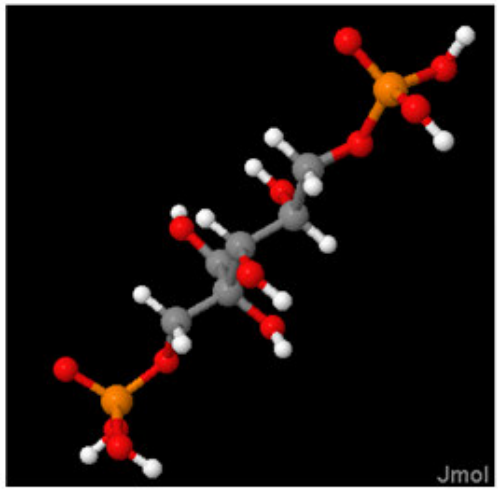


Figure 4
structure of ligand (CABP) generated by corina

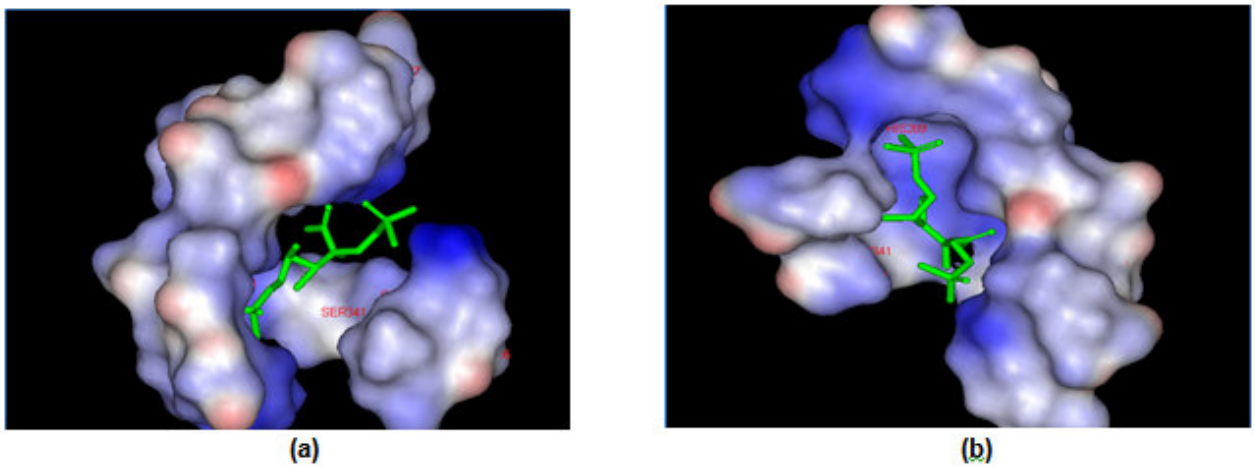


Figure 5
Docking of wild type (a) and mutant type
(b) of Rubisco with the ligand, viewed using Accelrys DS Visualizer.

The dock score with wild type was found to be -9.26 showing a good binding between enzyme and the substrate whereas with the mutant type, the dock score was only -7.91 showing a decrease in binding affinity between enzyme and substrate than the wild type that the change in K334 to R334 brings change in binding of the substrate. This shows the importance of Lys-334, mutation at this site decreases the affinity, specificity and catalytic activity of Rubisco for bisphosphate substrate.

DISCUSSION

Plant productivity can be reduced by up to 50% because the enzyme that initiates photosynthesis, ribulose 1, 5-bisphosphate carboxylase / oxygenase (Rubisco; EC 4.1.1.39), is not specific[17].Rubiscocatalyses a variety of side-reactions which involve the enediol of ribulose 1,5-bisphosphate (RuBP), formed during the initial step of catalysis[18]. This enediol of RuBP is a potent nucleophile, which can react with a range of electrophiles in addition to CO_2[19].Discrimination between the carboxylase and oxygenase reactions by Rubisco is determined by the stabilization of the carboxy or hydroperoxy intermediates by specific residues at the active site. Genetic manipulation of Rubisco to double its specificity for CO_2 would theoretically increase A (max) by perhaps 20%, and photosynthesis at sub-saturating light intensities would also be improved[20]. Since the Rubisco large subunit carries all active-site residues, the fate of the enediolate intermediate depends primarily on the identity and disposition of the amino acid residues within the large subunit[21]. High-resolution crystal structures of Rubisco from *Anacystis nidulans*[22] and other species have shown that, in the reaction of CO_2 with RuBP, Lys-334 (positioned at the apex of loop 6) has electrostatic interactions with the phosphate group on C-1 and the 2"-carboxy group of the transition-state intermediate. In our study, we predicted the change encountered during modification of an amino acid in the active site. The template chosen for Rubisco had 93% identity. With the target and template sequences, we modelled the Rubisco structure with Modeller 9V3. The mutant type of Rubisco was modelled with a change of arginine in place of Lysine at position 334 of loop 6 of Rubisco.The modelled structure was then evaluated using RAMPAGE. RAMPAGE is based on an analysis of (φ, ψ) angles using Ramachandran map[23], peptide bond planarity, bond lengths, bond angles, hydrogen-bond geometry, and side-chain conformations of known protein structures as a function of atomic resolution[24].

For example, phi and psi angles are plotted onto Ramachandran map from which the percentage of phi, psi in the allowed and disallowed regions was found out. For a reasonably good structure, it is expected to find a large majority of the residues (more than 90%) in the core region (most favoured region), about 3-5% residues in the additional allowed region, about 1% residues in the generally allowed region and 0% in the disallowed region, indicates that the model is stereo-chemically favoured. If the model is not convincing in terms of profile scores and stereo-chemical quality, the modelling is either repeated by making local changes in the sequence alignment or simply by rebuilding the erred regions. The results of our study was convincing since the "favoured", "allowed" and "outlier" regions are within the ranges of expected percentage.The docking studies were analysed using Autodock. The dock score value gives the binding affinity of the substrate to the protein. The dock value of the substrate with the wild type (-9.26) shows that binding of the substrate with the wild type is best than that of the mutant (-7.91).This shows the importance of lysine residue at 334 in the enzyme activity which is in correlation with the experimental results[22]. It is shown experimentally that mutants containing substitutions of Lys-334 catalysed enediolate formation, but were unable to catalyse the reaction of the enediolate with CO_2 or to form a stable complex with CABP[25-27].

CONCLUSION

In the present study, wild type and mutant structure of Rubisco enzyme from *Dunaliellasalina* were modelled using Modeller9v3 and docked with CABP to find the difference in binding, when a change is induced in the amino acid at active site, which may lead to reduced activity of the enzyme. The dock score with wild type showed binding affinity between enzyme and the substrate increased as compared with the mutant type dock score. The Lys-334 of loop6 of Rubisco is shown to determine its specificity for gaseous substrate. As many authors have showed it our results also predicted the importance of Lys-334 in Rubisco of *D. salina*. Still experimental validation is required to confirm our prediction, which is important to increase the productivity of *D. salina* having commercial value.

ACKNOWLEDGEMENT

The author is grateful to Dr. A. Anthoni Samy, Ph.D, Associate Professor in Biochemistry, College of Health and Medical Sciences, Wollega University, Nekemte, Ethiopia, Bharathidasan University, Madras Veterinary College, Islamiah College,Vaniyambadi for providing all the needed facilities to complete this work successfully.

CONFLICT OF INTEREST

Conflict of interest declared none.

REFERENCES

1. Parry MA, Madgwick P, Parmar S, Cornelius MJ, Keys AJ. Mutations in loop six of the large subunit of ribulose-1,5-bisphosphate carboxylase affect substrate specificity. Planta. 1992;187(1):109-12
2. Spreitzer RJ, Salvucci ME. Rubisco: structure, regulatory interactions, and possibilities for a better enzyme. Annual review of plant biology. 2002;53:449-75.
3. Parry MA, Andralojc PJ, Mitchell RA, Madgwick PJ, Keys AJ. Manipulation of Rubisco: the amount, activity, function and regulation. Journal of experimental botany. 2003;54(386):1321-33.
4. Cleland WW, Andrews TJ, Gutteridge S, Hartman FC, Lorimer GH. Mechanism of Rubisco: The Carbamate as General Base✗. Chemical reviews. 1998;98(2):549-62.
5. Soper T, Larimer F, Mural R, Lee E, Hartman F. Role of asparagine-111 at the active site of ribulose-1, 5-bisphosphate carboxylase/oxygenase from Rhodospirillum rubrum as explored by site-directed mutagenesis. Journal of Biological Chemistry. 1992;267(12):8452-7.
6. Ginzburg M. Dunaliella: a green alga adapted to salt. Advances in Botanical research. 1988;14:93-183.
7. Borowitzka MA, Borowitzka LJ, Microalgal Biotechnology. Cambridge, UK. Cambridge University Press, 1988 Jan 27-58.
8. Chai Y, Sun Y, Pan W, Jia Y, Xue L. Isolation and characterization of rbcL gene from Dunaliella salina. Life Science Journal. 2007;4(1):6-12.
9. Giordano M, Bowes G. Gas exchange and C allocation in Dunaliella salina cells in response to the N source and CO2 concentration used for growth. Plant physiology. 1997;115(3):1049-56.
10. Europeon Bioinformatics Institue. Pfam 31.0[Internet]. Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, UK: Europeon Bioinformatics Institute; 2004 January 1[updated 2017 March; cited yr month date]. Available URL: http://pfam.xfam.org/
11. Sonnhammer EL, Durbin R. A workbench for large-scale sequence homology analysis. Computer applications in the biosciences: CABIOS. 1994;10(3):301-7.
12. Falquet L, Pagni M, Bucher P, Hulo N. The PROSITE database. Nucleic Acids Res. 2006 Jan 1; 34 (Database issue):D227-30.
13. Swiss Institute of Bioinformatics. Database of protein domains, families and functional sites[Internet] University of Geneva, Switzerland: Swiss Institute of Bioinformatics ; 2006 February[2016 September 8 ; cited year month date]. Available URL: http://prosite.expasy.org/
14. J.Shi, T.L.Blundell, and K.Mizuguchi. FUGUE:sequence-structure homology recognition [Internet]. Universsity of Cambridge, Cambridge, UK: J Mol Biol; 2005 June 30[updated 2015 April 15; cited yr month date]. Available URL: http://mizuguchilab.org/fugue/
15. National Center for Biotechnology Information. Basic Local Alignment Search Tool[Internet]. United States National Library of Medicine(NLM), Maryland : NCBI; 1990 May 15[updated 2017

February 27; cited year month date]. Available URL: https://blast.ncbi.nlm.nih.gov/Blast.cgi
16. Paul de Bakker and Simon Lovell. RAMPAGE Ramachandran Plot Analysis[Internet]. University of Cambridge, Cambridge, UK: Crystallography and Bioinformatics Group; Available URL: http://mordred.bioc.cam.ac.uk/~rapper/rampage.php
17. Keys AJ. Prospects for increasing photosynthesis by control of photorespiration. Pest Management Science. 1983;14(3):313-6.
18. Pierce J, Lorimer GH, Reddy GS. Kinetic mechanism of ribulosebisphosphate carboxylase: Evidence for an ordered, sequential reaction. Biochemistry. 1986;25(7):1636-44.
19. BAINBRIDGE G, ANRALOJC PJ, MADGWICK PJ, PITTS JE, PARRY MA. Effect of mutation of lysine-128 of the large subunit of ribulose bisphosphate carboxylase/oxygenase from Anacystis nidulans. Biochemical Journal. 1998;336(2):387-93.
20. Reynolds MP, Van Ginkel M, Ribaut JM. Avenues for genetic modification of radiation use efficiency in wheat. Journal of experimental botany. 2000;51(suppl 1):459-73
21. Gutteridge S. The relative catalytic specificities of the large subunit core of Synechococcus ribulose bisphosphate carboxylase/oxygenase. Journal of Biological Chemistry. 1991;266(12):7359-62.
22. Newman J, Gutteridge S. The X-ray structure of Synechococcus ribulose- bisphosphate carboxylase/ oxygenase -activated quaternary complex at 2.2-A resolution. Journal of Biological Chemistry. 1993;268(34):25876-86.
23. Ramachandran Gt, Sasisekharan V. Conformation of polypeptides and proteins. Advances in protein chemistry. 1968;23:283-437.
24. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. Journal of applied crystallography. 1993;26(2):283-91.
25. Gutteridge S, Rhoades DF, Herrmann C. Site-specific mutations in a loop region of the C-terminal domain of the large subunit of ribulose bisphosphate carboxylase/oxygenase that influence substrate partitioning. Journal of Biological Chemistry. 1993;268(11):7818-24.
26. Hartman FC, Lee E. Examination of the function of active site lysine 329 of ribulose-bisphosphate carboxylase/oxygenase as revealed by the proton exchange reaction. Journal of Biological Chemistry. 1989;264(20):11784-9.
27. Soper TS, Mural RJ, Larimer FW, Lee EH, Machanoff R, Hartman FC. Essentiality of Lys-329 of ribulose-1, 5-bisphosphate carboxylase/oxygenase from Rhodospirillum rubrum as demonstrated by site-directed mutagenesis. Protein engineering. 1988;2(1):39-44.

Reviewers of this article

Dr S.K.M.Habeeb Ph.D

Assistant Prof,Dept of Bioinformatics,School of Bioengineering ,Faculty of Engineering and Technology ,SRM University,Kattankulathur,India

Prof. Srawan Kumar G.Y

Associate Professor,Nalanda Institute of Pharmaceutical Sciences, Sattenapalli, Guntur, Andrapradesh, India

Prof.Dr.K.Suriaprabha

Asst. Editor , International Journal of Pharma and Bio sciences.

Prof.P.Muthuprasanna

Managing Editor , International Journal of Pharma and Bio sciences.

We sincerely thank the above reviewers for peer reviewing the manuscript