JOURNAL OF THE ROYAL LAUREATES ACADEMY

www.rlaindia.org

RESEARCH ARTICLE

COMPUTATIONAL SEQUENCE ANALYSIS AND IN SILICO MODELING OF LEAF PART PROTEIN OF *Pouzolzia zeylanica* (L) BENN.

Deepsha Das*

Department of Microbiology, A.P.C College, New Barrack pore, North 24 Parganas, Kolkata - 700131, West Bengal, India.

* Corresponding Author.

ABSTRACT:

Pouzolzia zeylanica (L) Benn is a unisexual flowering plant in the family of Urticaceae. Ribulose bis phosphate carboxylase chloroplast protein of Pouzolzia zeylanica (L) Benn played a significant role in plant metabolism and as well as this leave part protein has great medicinal activity. Hence, understanding molecular structure and function of the protein is important for biologist. The present study was aimed at sequence and In silico analysis of Ribulose bisphosphate carboxylase protein coded by rbcL gene, through comparative modeling. Validation of the structure, stereo-chemical parameter was carried out using different computational tool and servers. This structural analysis will provide an idea about ligand design, drug design and transcriptional study. Hence this work will also help for detecting protein in vivo.

KEYWORDS: In Silico, Homology modeling, Ramchandran plot, VDW repulsion energy, Zeta angle.

INTRODUCTION:

Pouzolzia zeylanica (L) Benn is a unisexual flowering plant in the family of Urticaceae. Common name of the Pouzolzia is Graceful Pouzolzia. This plant exclusively grows in Asia. It also occurs in Australia, the northern part of Cape York Peninsula, North east Queensland and Malaysia [1]. Urticaceae is angiosperms plant normally its height 12 cm to 40cm. leaves are elevated to lanceolate and deeply viand (1.2-9 cm long). They are covered with long white hair, especially along the leaf margin. Reddish stems are densely lined with white hairs; tiny unisexual flowers are clustered together in the leaf axis. Leaf part of the Pouzolzia *zeylanica* (L) Benn. has great medicinal value. It has anti – oxidant capacity, free radical-scavenging capacity, analgesic activity and anti inflammatory activity [2]. Leaves of this plant are anthelmintic and vulnerary used as cicatrizant for gangrenous ulcers, in syphilis and gonorrhea. Leaf pest is used as a mask for recovery of Pustule disease [3]. *Pouzolzia zeylanica* (L) Benn. is selected as an herbal drug. The effective part of *Pouzolzia zeylanica* is leaves. So, Ribulose bisphosphate carboxylase, chloroplast protein has been considered in this study. Ribulose bisphosphate carboxylase (RuBisCO), catalyses the initial step in Calvin's cycle of plants. It catalyzes the primary CO_2 fixation step. RuBisCO is activated by carbamylation of an active site lysine and stabilized by the divalent cation. Then the activated RuBisCO which catalyzes the substrate ribulose 1, 5 bisphosphate (RuBP) and leads to the formation of two molecules of 3-phosphoglycerate. Magnesium ions (Mg^{2+}) are required for enzymatic activity. Correct positioning of (Mg^{2+}) in the active site of the enzyme involves addition of an "activating" carbon dioxide molecule (CO_2) to a lysine in the active site and forming a carbamate [4]. In the light, RuBisCO promotes the release of the inhibitory RuBP from the catalytic sites. In darkness, RuBisCO is inhibited a substrate analog 2-Carboxy-D-arabitinol 1-phosphate (CA1P) [5]. CA1P binds to the active site of activated RuBisCO and inhibits catalytic activity. In the light, RuBisCO promotes the release of CA1P from the catalytic sites; it is rapidly converted to a non-inhibitory form by a light-activated CA1P-phosphatase. This protein enzyme plays an important role in plant metabolism and as well as stabilized the function of chloroplast. Thus, identification of this protein enzyme here use In silico method. In silico approach that has been highlighted on physico-chemical character and structure of protein which have no PDB identification code, apprehends the efficacy of various tool of bioinformatics. This study and utilized further for molecular research,

drug design and therapeutic purpose. Computational tools provide researches to understand the overview of protein sequence and as well as protein structure. A large number of computation tools are available from different sources for making prediction regarding the identification and structure prediction of proteins. The major drawbacks of experimental methods that have been used to characterize the proteins of various organisms are time consuming, costly and fact that this methods not amendable to high throughput techniques. Computationally based characterization of the features of proteins found or predicted in completely sequenced proteomes is an important task in search for knowledge of protein function. In this paper, the in silico analysis and homology modeling studies of target proteins were reported. Three dimensional structures for these proteins were yet not available in protein database. Hence to describe it structural features and to understand molecular function, the model structures for these proteins were constructed [6].

MATERIALS AND METHODS:

The 230 amino acid containing sequence of a protein Ribulose bis phosphate carboxylase (RuBisCO) encoded by the gene rbcL with Accession AGZ90846.1, Gi No -557163227 were retrieved from the NCBI database and UniProt (Universal Protein Resource) which are freely accessible database which contains data of proteins. Table showed the protein sequences consider in this study. The target protein sequences were retrieved in FASTA format.

Table 1. Protein sequence consider for the study:

SERIAL NO	PROTEIN NAME	GI NO.	UNIPORT NO.	LENGTH	DISCRIPTION
1	RuBisCO	557163227	U5YGC0_9ROSA	230	Ribulose bisphosphate carboxylase, chloroplast from *Pouzolzia zeylanica*

					(L) Benn.

Sequence analysis:

Physiochemical properties of the protein were computed by ProtParam tool (http://web.expasy.org/protparam/). The parameters computed by ProtParam included the molecular weight, theoretical pI, instability index, aliphatic index, and grand average of hydropathicity (GRAVY) [7]. Subcellular organisation of any protein aids understanding protein function. Phosphorylation profile was analysis by NetPhos (http://www.cbs.dtu.dk/services/NetPhos/). NetPhos Server is a tool to predict phosphorylation site at threonine, serine and tyrosine residue because these are mostly phosphorylated as they contain hydroxyl group thus are capable of binding phosphate group [8]. Protein ubiquitination is one of the most vital post-translational modifications by covalent attachment of ubiquitin to lysine residues. UbiSite (http://csb.cse.yzu.edu.tw/UbiSite/) is a universal database for ubiquitination of proteins. The tool provides location of the Ubiquitination site, correspond fragment and substrate motifs [9]. Prediction of potential methylation and acetylation of protein sequence is done used In silico tool PLMLA (Prediction of potential lysine methylation and lysine acetylation) .Sequence is submitted in FASTA format and appropriate option is selected for prediction. Name of proteins and their site methylation, acetylation position predicted result is returned.

Structure analysis:

The secondary structural features prediction of the protein sequence was carried out by SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat) [10].

RaptorX (http://raptorx.uchicago.edu/StructurePrediction/predict/) modeler was used to build a protein 3D model using automated approach to comparative protein structure modeling by satisfaction of spatial restraints [11].

The constructed 3D models were energy minimized in CHIRON (http://redshift.med.unc.edu/chiron/login.php) by short discrete molecular dynamics (DMD) simulation [12].

The overall stereochemical property of the protein was assessed by Ramchandran plot analysis. The validation for structure model obtained from the software tool was performed by using PROCHECK (http://servicesn.mbi.ucla.edu/PROCHECK/) [13].

PROCHECK verify the quality of model by generating Ramchandran plot and assess the quality of the structure was computed in terms of percentage of residues in favorable regions, percentage of non Proline, Glycine residues etc. Different kinds of Pockets present on surface of protein and amino acid residues. In these pockets are powerful to generate physiochemical properties of protein. These properties are required for protein to perform its function. RaptorX binding is an online tool which was used to analyze the active site of the protein and the amino acid that are present in those sites. It provides information of amino acid residues that would be binding with ligand and details overview of the specific ligand.

RESULTS AND DISCUSSION:

Availability of abundance of quality tools and web servers has enabled computational biologists to perform reliable analysis of protein sequence and structure. The present study was aimed at sequence analysis and homology modeling of the Ribulose bisphosphate carboxylase protein to focus on its function.

Sequence analysis:

It showed targeted protein of *Pouzolzia zeylanica* herbal plant was considered in this study. Protein sequence was retrieved in FASTA format from the UniProt, a public domain protein database. Parameters computed using ExPASy's ProtParam. The calculated isoelectric point (pI) will be useful because at pI, solubility of the protein is least and mobility in an electro focusing system is 0 (zero). Isoelectric point (pI) is the pH at which the surface of protein is covered but

net charge of protein is zero. At pI proteins are stable and compact. Here the computed pI value of the protein was observed below 7 which indicate the protein was acidic. Isoelectric point (pI) will be useful for preparing buffer system for purification by using isoelectric focusing method. The computed extinction coefficient (EC) helps in the quantitative study of protein–protein and protein–ligand interactions in solution.

Table 2: Parameters computed using ExPASy's Prot Param tool:

Gi no	Protein name	Length	MW (a.a)	pI	-R	+R	EC	II	AI	GRAVY
557163227	RuBisCO	230	25775.16	6.01	27	29	39100	33.01	74.26	-0.417

Here Extinction coefficients are in units of M-1 cm-1, at 280 nm measured in water. The instability index (II) provides an estimate of the stability of protein in a test tube. If (II) value of the protein is less than 40 that indicate protein is stable. According to the ExPASy's ProtParam tool output it is showing that RuBisCO is stable protein. Aliphatic index was found to be 74.26 while Grand average of hydropathicity (GRAVY) index was calculated as -0.417 demonstrating amino acid to be in soluble protein.

Table 3. Phosphorylation profile of RuBisCO protein using neural network approach. Specific residue positions in the query protein are shown to be phosphorylated based on a significant Score. *S* refers to phosphorylation on serine residue,*T* refers to phosphorylation on threonine residue and *Y* refers to phosphorylation on tyrosine.(CKI- Casein kinase 1, CKII- Casein kinase 2, Unsp- Unspecified protein, PKA- Protein kinase A, PKC- Protein kinase C, PKG- Protein kinase G, DNAPK- DNA-dependent protein kinase, ATM- Ataxia telangiectasia-mutated):

Protein name	Location	Residue	Context	Score	Kinase
RuBisCO	13	Y	YTPEYETKD	0.953	unsp
	15	T	PEYETKDTD	0.605	CKII
	15	T	PEYETKDTD	0.510	unsp

	27	T	AFRVTPQPG	0.713	unsp
	45	S	VAAESSTGT	0.617	CKI
	45	S	VAAESSTGT	0.504	unsp
	46	S	AAESSTGTW	0.573	CKI
	46	S	AAESSTGTW	0.532	PKA
	49	T	SSTGTWTTV	0.617	CKI
	52	T	GTWTTVWTD	0.517	PKC
	60	S	DGLTSLDRY	0.938	unsp
	60	S	DGLTSLDRY	0.682	PKC
	60	S	DGLTSLDRY	0.505	CKI
	81	Y	EENQYIAYV	0.912	unsp
	81	Y	EENQYIAYV	0.568	SRC
	96	S	FEEGSVTNM	0.571	CKI
	96	S	FEEGSVTNM	0.525	PKA
	149	Y	KLNKYGRPL	0.506	INSR
	165	S	KLGLSAKNY	0.950	unsp
	165	S	KLGLSAKNY	0.612	PKG
	165	S	KLGLSAKNY	0.531	PKC
	184	T	GLDFTKDDE	0.576	CKII
	184	T	GLDFTKDDE	0.529	cdc2
	192	S	ENVNSQPFM	0.642	DNAPK
	192	S	ENVNSQPFM	0.562	ATM
	210	Y	AEAIYKSQA	0.904	unsp
	212	S	AIYKSQAET	0.883	unsp
	212	S	AIYKSQAET	0.574	DNAPK
	212	S	AIYKSQAET	0.563	CKII
	216	T	SQAETGEIK	0.571	PKC
	223	Y	IKGHYLNAT	0.702	unsp

Protein phosphorylation is a type of post-translational modification which can turn a protein on and off, thus modifying its function and activity. Phosphorylation generally occurs on serine(S), threonine (T) and tyrosine (Y).The kinase replace natural hydroxyl groups on serine(S), threonine (T) and tyrosine (Y) with negatively charge phosphate. Here, NetPhos is an artificial neural network method used for predicts phosphorylation sites. It predict in independent sequences with sensitivity in the range from 69% to 96%.

Regions of RuBisCO sequence showed extensive phosphorylation on serine and residues. While least no phosphorylation capability of threonine and tyrosine residues was predicted. Computational prediction of potential ubiquitination sites has become a useful strategy for complete proteome annotation. "UbSite" ubiquitination prediction tool was used in this case. UbSite can be used to predict ubiquitination sites for multiple species. This server follow SVM algorithm. Ratio of positive and negative sample is 1:1 and window size is 41. The tool provides score, sensitivity and specificity. Specificity define possible Ubiquitination site. Table4. Show the result summery of Ubiquitination site for selected protein.

Table 4. Result summary of Predicted Ubiquitination site. Specific residue positions in the query protein are shown to be ubiquitination based on a significant score. According to the legend box, confidence level is also showing corresponding Ubiquitination site:

Protein name	Location	Ubiquitination Sites	Substrate Motifs	Confidence
RuBisCO	16	TPEYET **K** DT**D**ILA	0.422489	**Medium**
	112	GNVFGF **K** ALRALR	0.622333	**High**
	145	IQVERD **K** LNK**Y**GR	0.546072	**High**
	148	ERDKLN **K** **Y**GRPLL	0.47491	**Medium**
	159	LLGCTI **K** PKLGLS	0.338439	**Medium**
	167	KLGLSA **K** N**Y**GRAV	0.50677	**High**
	185	GGL**D**FT **K** DDENVN	0.429162	**Medium**
	220	AETGEI **K** GHYLNA	0.391826	**Medium**

Legend			
Confidence level	**Score range**	**Sensitivity**	**Specificity**
High	$0.67 \leq s$	37.54%	86.28%
Medium	$0.50 \leq s \leq 0.65$	67.06%	59.93%
Low	$s \leq 0.50$	86.48%	33.96%

Here, RuBisCO have highly possible Ubiquitination site on location 112,145 and 167(amino acid residue position). Prediction of potential methylation and acetylation of protein sequences was used PLMLA (Prediction of potential lysine methylation and lysine acetylation) tool. PLMLA is an in silico online tool using for prediction of potential lysine methylation and lysine acetylation site from protein sequences. Here, window size -6 to +6 is employed to construct the prediction model. The system efficiently returns the predictions, including protein name, the position of site, flanking amino acids, predicted result and SVM probability.

Table 5. Represent predicted result of methylation by PMLA:

Protein name	Position of site	Flanking residues	Predicted result	SVM Probability
RuBisCO	130	IPPAYI-**K**-TFQGPP	methylated lysine	0.544797
	145	IQVERD-**K**-LNKYGR	methylated lysine	0.500000
	148	ERDKLN-**K**-YGRPLL	methylated lysine	0.676530
	161	GCTIKP-**K**-LGLSAK	methylated lysine	0.566163
	167	KLGLSA-**K**-NYGRAV	methylated lysine	0.585978
	185	GGLDFT-**K**-DDENVN	methylated lysine	0.528486
	220	AETGEI-**K**-GHYLNA	methylated lysine	0.533680

Table 6. Represent predicted result of acetylation by PMLA:

Protein name	Position of site	Flanking residues	Predicted result	SVM Probability

RuBisCO	16	TPEYET-**K**-DTDILA	acetyllysine	0.548913
	65	TSLDRY-**K**-GRCYHI	acetyllysine	0.522093
	112	GNVFGF-**K**-ALRALR	acetyllysine	0.500000
	145	IQVERD-**K**-LNKYGR	acetyllysine	0.521565
	148	ERDKLN-**K**-YGRPLL	acetyllysine	0.552415
	159	LLGCTI-**K**-PKLGLS	acetyllysine	0.509113
	167	KLGLSA-**K**-NYGRAV	acetyllysine	0.514422
	211	CAEAIY-**K**-SQAETG	acetyllysine	0.544585

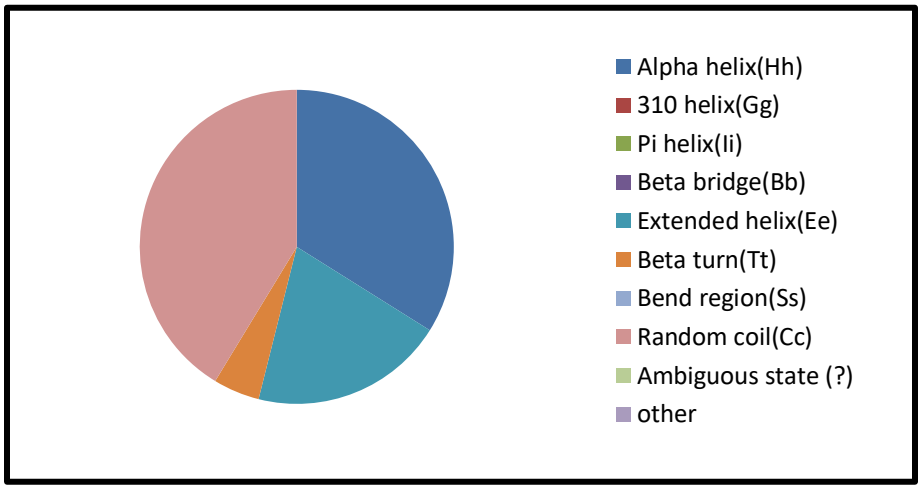
Structure analysis:

The secondary structure of targeted protein was predicted by SOPMA (Self Optimized Prediction Method with Alignment) which correctly predicts 69.5% of amino acids for a state description of the secondary structure prediction. The secondary structure indicates whether a given amino acid lies in a helix, strand or coil. Secondary structure features as predicted using SOPMA were represented for protein RBCL. The result revealed that random coils dominated among secondary structure elements followed by alpha helix, extended strand and beta turns for sequence. The secondary structure was predicted by using default parameters (Window width: 17, similarity threshold: 8 and number of states: 4)

Table 7. Calculated secondary structure elements by SOPMA.

Secondary structure	**Protein name**
	RuBisCO
Alpha helix(Hh)	33.91%
3_{10} helix(Gg)	0.00%
Pi helix(Ii)	0.00%
Beta bridge(Bb)	0.00%
Extended helix(Ee)	20.00%
Beta turn(Tt)	4.78%
Bend region(Ss)	0.00%

Random coil(Cc)	41.30%
Ambiguous state (?)	0.00%
Other	0.00%



The modelling of the three dimensional structure of the selected proteins was performed by homology modelling programme RaptorX. RaptorX predict tertiary structure and also assigns some confidence score to indicate the quality of predicted structure and model error at each residue. Here, "P" value indicate the relative global quality and" uGDT" indicate the absolute global quality. The smaller "P" value implies the higher quality the model. Score value is the sequence alignment score. Score may exceed the sequence length of protein due to estimation error. A model with both good P value and uGDT is very likely to be high quality. A model with good P value but poor uGDT may be low quality. A model with a good uGDT but poor P value may be of good quality although it may not be better than randomly generated model.

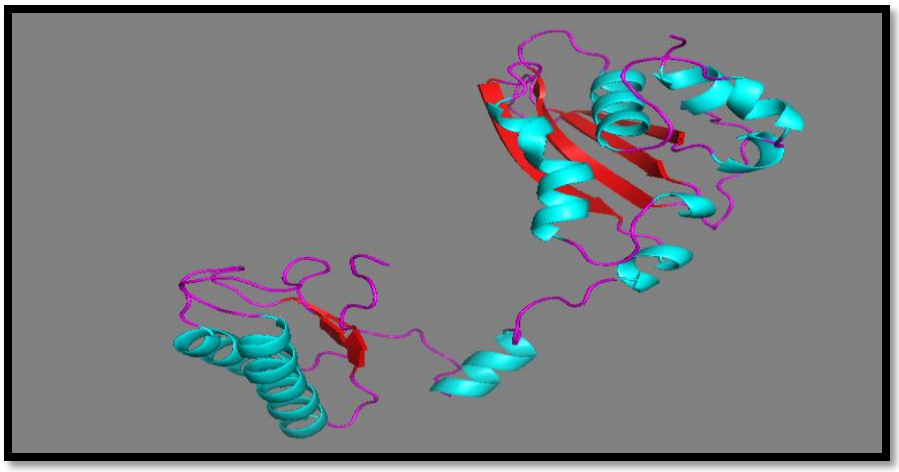


Figure 1. Stereo ribbon configuration view represent predicted model of the protein

In RaptorX for the model prediction of RBCL, 100% residues are covered and 3% positions predicted as disorder. The computated score, "P" value, "uGDT" value for selected proteins was shown. Here, P value of the protein is not more than 4 that indicate predicted model are good quality.

Table 8. Result of predicted model quality justifying parameter.

Protein name	score	P value	uGDT value	uSeqId value
RuBisCO	260	2.25e-12	230(100)	223(97)

Steric clash is one of the problems in homology models. Steric clashes arise for the unwanted overlap of any two nonbonding atoms in a protein structure. Here, resolved this problem by Chiron. So, it was clear that steric classes decrease after energy optimization.

Table 9. Before energy minimize and after energy minimize clash report

	Description	Initial clash report	Final clash report
	Total number of residues	230	230

RuBisCO	Total number of contacts	3599	3287
	Total number of clashes	162	100
	Total VDW repulsion energy	157.213 kcal/mol	58.9245 kcal/mol
	Clash ratio	0.0436825	0.0170365

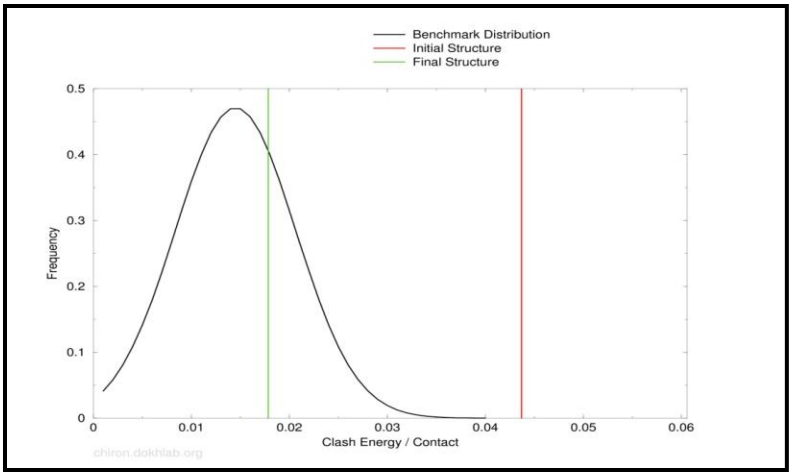


Figure 2. Chiron clash energy minimization summary; where red represents the class energy for the modeled structure, the green is for the final structure after minimization and the black is for a set of high resolution structures

The stereo chemical quality of the predicted models and accuracy of the protein model was evaluated again after the refinement process using Ramachandran Map calculations computed with the PROCHECK. The assessment of the predicted models generated by modeler was shown. The main chain parameters plotted are Ramachandran plot quality, peptide bond planarity, Bad bonded interactions, main chain hydrogen bond energy, C-alpha chirality and over-all G factor. In the Ramachandran plot analysis, the residues were classified according to its regions in the quadrilateral. The red regions in the graph indicate the most allowed regions whereas the yellow regions represent allowed regions. Glycine is represented by triangles and

other residues are represented by squares. The distribution of the main chain bond lengths and bond angles were found to be within the limits for these proteins. Ω angle standard deviation occurs due to peptide bond planarity and zeta angle deviation occurs due to α carbon tetrahydral distortion. Such figure assigned by Ramachandan plot represents a good quality of the predicted models. Ramachandan plot analysis based on structures of resolution of at least 2.0 Angstroms and good quality model would be expected to have over 90% in the most favorable region regions. Here, RuBisCO have 93.8% residues in most favorable region, 6.2% in additional allowed region regions.

Table 10. Model evaluated result of three protein's main chain parameters by PROCHECK:

Protein name	Stereo chemical parameter	No of data	Parameter value	Comparison value		No of band width from mean	remark
				Typical value	Band width		
RuBisCO	% residues	194	93.8	88.2	10.0	0.6	inside
	Ω angle standard deviation	228	10.2	6.0	3.0	1.4	worse
	Bad contact	2	0.9	1.0	10.0	-0.0	inside
	Zeta angle standard deviation	210	2.0	3.1	1.6	-0.7	inside
	H- bond energy deviation	130	0.8	0.7	0.2	0.4	inside
	Overall G - factor	230	-0.4	-0.2	0.3	-0.8	inside

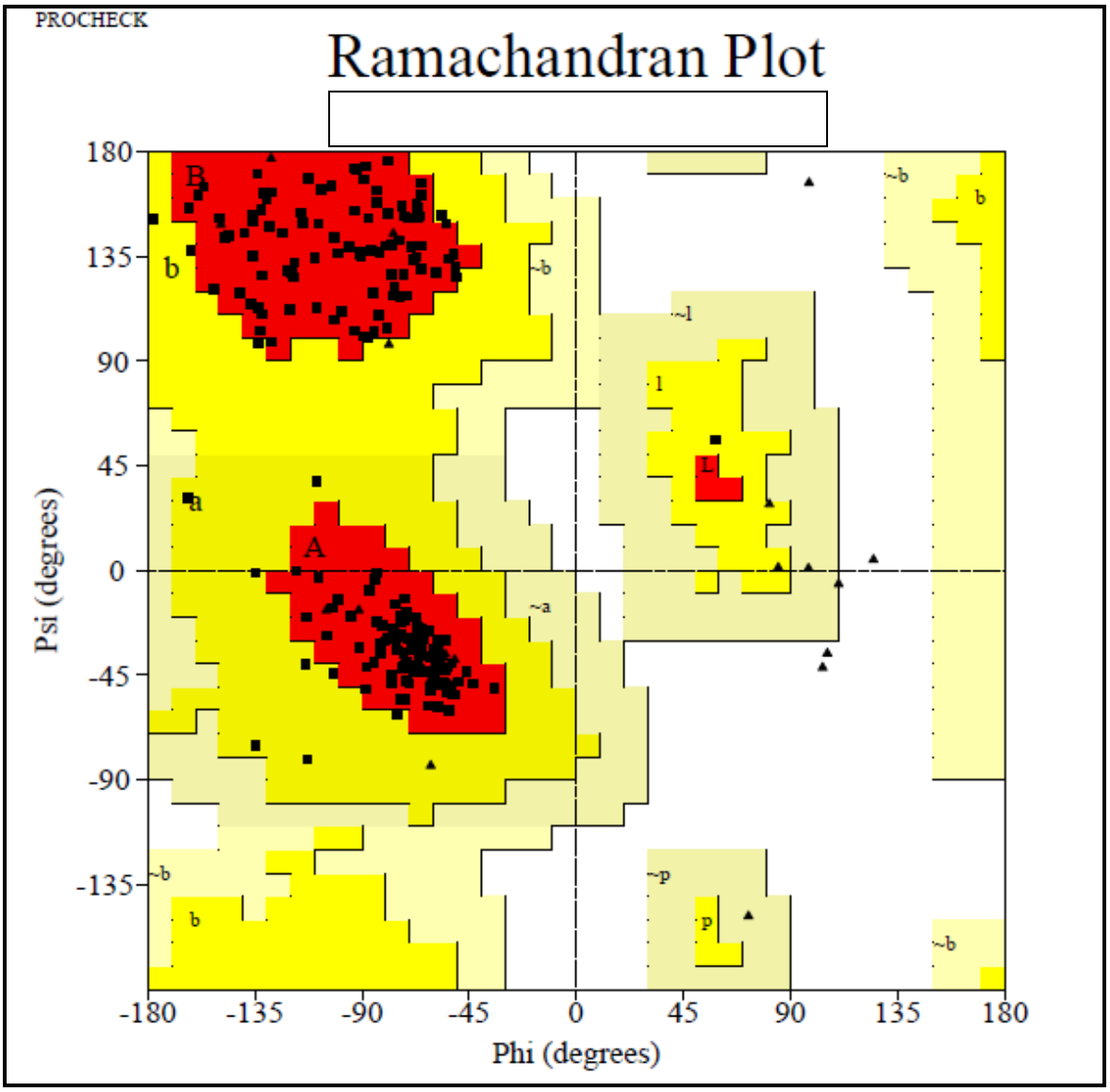
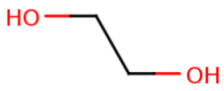
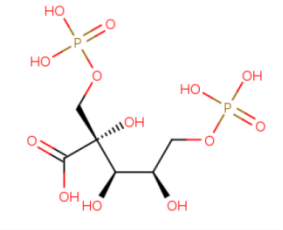


Figure 3. Ramachandan map of selected protein RuBisCO - 93.8% residues in most favorable region, 6.2% in additional allowed region. Glycine is represented by triangles and other residues are represented by squares

Table 11. Summary of binding site Prediction Results:

Protein name	Binding residues	Multiplicity	Possible ligand	Ligand structure
RuBisCO	T157 K159 K161 D187	197	MG	Mg^{2+}
	V1 K2 Y8 G48 T49 W50 T51 T52 V53 D56 L61	175	EDO (ethane-1,2-diol)	
	Y4 E36 A113	80		
	Y210 K211 A214	61		
	T49 W50 N107	94	CAP (2-carboxyarabinitol-1,5-diphosphate)	

Protein interactions are important from the aspect of the cellular function and determining how these proteins interact with their ligands and other small molecules. Predicting active site of the in silico modeled proteins are further of great aspect as it provides more precised characterization of protein from functional point of view. Active site of RuBisCO was predicted using RaptorX binding online software. RaptorX binding software predicted results were summarized.

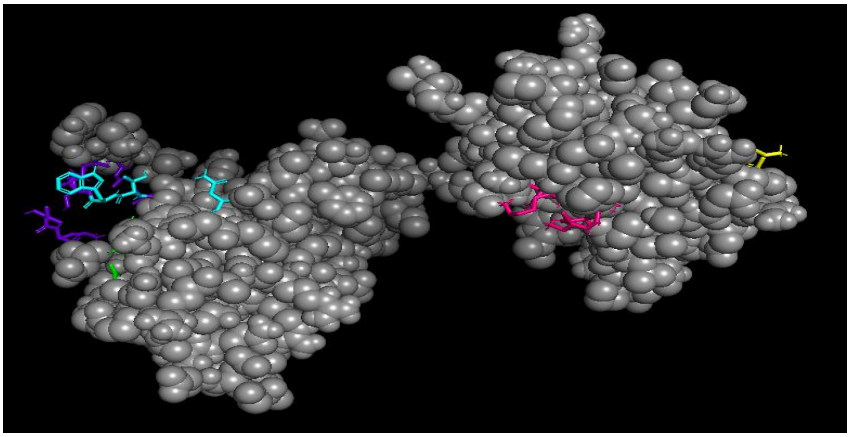


Figure 4. Represent three proteins label with its main active pocket. pink color code indicate first pocket which multiplicity is 197, blue color code indicate second pocket which multiplicity is 175, green color code indicate third pocket which multiplicity is 80, yellow color code indicate forth pocket which multiplicity is 61 and cyan color code indicate fifth pocket which multiplicity is 94.

For binding site prediction, Pocket Multiplicity is used to judge the quality of a predicted pocket as well as predicted binding site. It represents the frequency with which the selected pocket was found in a set of ligand-binding protein structures. When Multiplicity is above 40, there is a good chance that the predicted pocket is true.

CONCLUSION:

Ribulose bisphosphate carboxylase is a key enzyme produced by Pouzolzia zeylanica. This chloroplast protein plays an important role in plant metabolism and as well as enhance medicinal activity of the leaf of Pouzolzia zeylanica. In the present work, sequence analysis has been conducted to shed light on post translational modification of domains associated with RuBisCO protein and 3D structure study of the protein. This Computational work might be valuable contribution in the field of bioinformatics research and may get other idea about the protein structure. Hence this work will also help for detecting of such type of protein in vivo. [14-17]

REFERENCES:

1. Paul, S., Saha, D. (2012). In Vitro Screening of Cytotoxic Activities of Ethanolic Extract of Pouzolzia Zeylanica (L.) Benn, International Journal of Research Article Pharmaceutical Innovations,2(1):52-55

2. Li, P., Huo, L., Su, W., Lu, R., Deng, C., Liu, L., Deng, Y., Guo, N., Lu, C., He, C. (2011). Free radical-scavenging capacity, antioxidant activity and phenolic content of Pouzolzia zeylanica, J. Serb. Chem. Soc.,76 (5): 709–717

3. Yusuf,M., Wahab,M.A., Chowdhury,J.U., Begum, J. (2006). Ethno-Medico-Botanical Knowledge From Kaukhali Proper and Betbunia of Rangamati District, Bangladesh J. Plant Taxon.,13(1): 55-61

4. Taylor, T.C., Andersson, I. (1997). The structure of the complex between rubisco and its natural substrate ribulose 1,5-bisphosphate, Journal of Molecular Biology , 265(4):432-444

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

6. Prajapati,C., Bhagat,C. (2012). In-Silico Analysis and Homology modeling of Target Proteins for *Clostridium botulinum*, IJPSR,; Vol. 3(7): 2050-2056

7. Pace, C.N., Vajdos, F., Fee, L., Grimsley, G. and Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. Protein Science, 4: 2411-2423.

8. Sahay, Archna and Shakya, Madhvi. (2010). In silico Analysis and Homology Modelling of Antioxidant Proteins of Spinach, J Proteomics Bioinform, 3(5): 148-154.

9. Deleage G. (2017). ALIGNSEC: viewing protein secondary structure predictions within large multiple sequence, alignments, Bioinformatics,33(24): 3991–3992.

10. Feng Z., JianP., Joe D., Freed, K.F., Sosnick, T.R., and JinboXu. (2010). A probabilistic and continuous model of protein conformational space for template-free modeling, Journal of Computational Biology, 17(6):1-8.

11. Laskowski R. A., MacArthur M. W., et al. (1993). PROCHECK: a program to check the stereochemical quality of protein structures." J. Appl. Cryst, 26: 283-29.

12. Basharat, Z (2015). Computational sequence analysis and in silico modeling of a stripe rust resistance protein encoded by wheat TaHSC70 gene, Journal of Computational Biology and Bioinformatics Research, 7(1):1-7.

13. Efferth, T., Greten, J., Henry (2012). In Silico Analysis of Microarray-Based Gene Expression Profiles Predicts Tumor Cell Response to Withanolides, Microarrays (Basel), 44–63.

14. Lagunin, A.A., Goel, R.K, Gawande, D.Y, Pahwa, P., Gloriozova, T.A, Dmitriev, A.V., Ivanov, S.M, Rudik, A.V., Konova, V.I., Pogodin, P.V, Druzhilovsky ,D.S., Poroikov, V.V. (2014). Chemo and bioinformatics resources for In silico drug discovery from medicinal plants beyond their traditional use : a critical review, Natural Product Reports, 31(11):1585-1611.

15. Remya, C., Dileep, K. V., Variayr, E. J., Sadasivan, C. (2016). An In silico guided identification of nAChR agonists from *Withania somnifera*, Taylor & Francis, Frontiers in life science, 9: 3, 201–213.

16. Sharmila, R., Subburathinam, K.M., Aishwarya, S., Margret, A. A. (2013). In-Silico Analysis of Andrographolide against Cancer, International Journal of Pharmaceutical Sciences and Drug Research, 5(2): 56-61.

17. Fu, M., Niu, Y.Y., Yu, J., Kong Q.T. (2012). Study on the chemical constituents in Pouzolzia zeylanica. Zhong Yao Cai., 35 (11): 1778-17781.