Pak. J. Agri. Sci., Vol. 54(2), 395-406; 2017
ISSN (Print) 0552-9034, ISSN (Online) 2076-0906
DOI: 10.21162/PAKJAS/17.6187
http://www.pakjas.com.pk

A NOVEL STRUCTURAL AND FUNCTIONAL INSIGHT INTO CHLOROPLAST-ENCODED CENTRAL SUBUNIT OF DARK-OPERATED PROTOCHLOROPHYLLIDE OXIDOREDUCTASE (DPOR) OF PLANTS

Muhammad Tahir Ul Qamar and Muhammad Sarwar Khan*

Center of Agricultural Biochemistry and Biotechnology (CABB), University of Agriculture, Faisalabad, Pakistan
***Corresponding author's e-mail: sarwarkhan_40@hotmail.com**

Chlorophyll converts harvested light into chemical energy during the strategic process of photosynthesis in chloroplasts. Protochlorophyllide reduction in the chlorophyll formation is catalyzed by two complex enzymes; light-dependent protochlorophyllide oxidoreductase (LPOR) and dark-operated protochlorophyllide oxidoreductase (DPOR). Of these two, DPOR is a three-subunit complex in which ChlB plays a vital role in developing photosynthetically competent chloroplasts. What has not been reported before is the complete structural and functional annotation of ChlB subunit from plants. Sequence, structure and functional analyses of the ChlB subunit are performed using a blend of molecular biology and bioinformatics approaches to identify conserved residues and distribution of amino acids. Complete ChlB sequence analysis coupled with phylogenetic analysis, molecular docking and protein-protein interaction reveal that the ChlB is thermo-stable, acidic and hydrophilic in nature. The 3D structure (RMSD of 0.20Å) of ChlB is predicted and used as a target in docking and protein-protein interaction studies. Structural characterization of ChlB further elucidates that the amino acids Arg18, Asn175, Glu221 and Asp311 being important catalytic residues are involved in the basic function of ChlB and its interaction with other DPOR subunits. Further, the mutation analysis substantiates the central role of predicted catalytic residues in the structure of the ChlB. We conclude that the generated information will facilitate researchers in engineering chlorophyll pathway to improve photosynthesis in plants.

Keywords: Photosynthesis, chlorophyll engineering, chloroplast, *In silico* analyses

INTRODUCTION

Photosynthetic organisms contain one of the most important molecules, chlorophyll, that converts captured solar energy into its biochemical form. Chlorophyll is an important tetrapyrrole pigment, which is responsible for harvesting and conversion of solar energy into its chemical form during the process of oxygenic photosynthesis on our planet (Gust *et al.,* 2001; Nelson and Junge, 2015; Tsukatani and Masuda, 2015). Chlorophyll and different other pigments are produced as metabolic products (Reinbothe *et al.,* 2010) of a complex biosynthetic pathway. The first intermediate of chlorophyll synthesis pathway is d-aminolevulinic acid (ALA), which is a C5-compound and the pathway is termed as C5-pathway (Foyer and Noctor, 2000; Camp *et al.,* 2003; Tanaka and Tanaka, 2007; Blankenship, 2013). Protochlorophyllide is precursor of chlorophyll a that lacks the phytol side chain of chlorophyll. Protochlorophyllide forms chlorophyllide by reduction of its C17=C18 double bond and converts porphyrin ring to a chlorine ring during the biosynthesis of chlorophyll. Chlorine ring contain magnesium (Mg^{2+}) usually called chlorophylls and are the main photosynthetic pigments in chloroplast. Chlorophyllide then immediately converted into chlorophyll a as shown in Figure 1 (Tanaka and Tanaka, 2007; Musuda and Fujita, 2008; Reinbothe *et al.,* 2010; Nazir and Khan, 2012). Asymmetric conjugated double bond system is essential for efficient light absorption of chlorophyll a, because it increases the absorption ability by shifting towards stronger wavelength that is visible or closer to the ultraviolet. Therefore, manipulation in chlorophyll pigment has an important impact on light absorption and leads to a change in the absorption properties of tetrapyrrole ring systems for wavelengths from 627 to 665 nm (Lebedev and Timko, 1999; Muraki *et al.,* 2010; Reinbothe *et al.,* 2010).

The reduction of protochlorophyllide to chlorophyllide in photosynthetic organisms is catalyzed by two different enzymes: LPOR (light-dependent protochlorophyllide oxidoreductase), a nuclear-encoded light-dependent plastid-localized single subunit enzyme that requires light for its activation in angiosperms, and DPOR (dark-operated protochlorophyllide oxidoreductase), a light-independent, plastid-encoded enzyme that is composed of three subunits; ChlL, ChlN, and ChlB. From an evolutionary perspective, light-dependent chlorophyll biosynthesis is assumed to be a universal feature of oxygenic photosynthetic organisms (Schoefs and Franck, 2003; Eckhardt *et al.,* 2004); however, genes (*ChlL, ChlN,* and *ChlB*) encoding three protein subunits of DPOR are highly conserved in their nucleotide sequence in the plastid genome of most gymnosperms

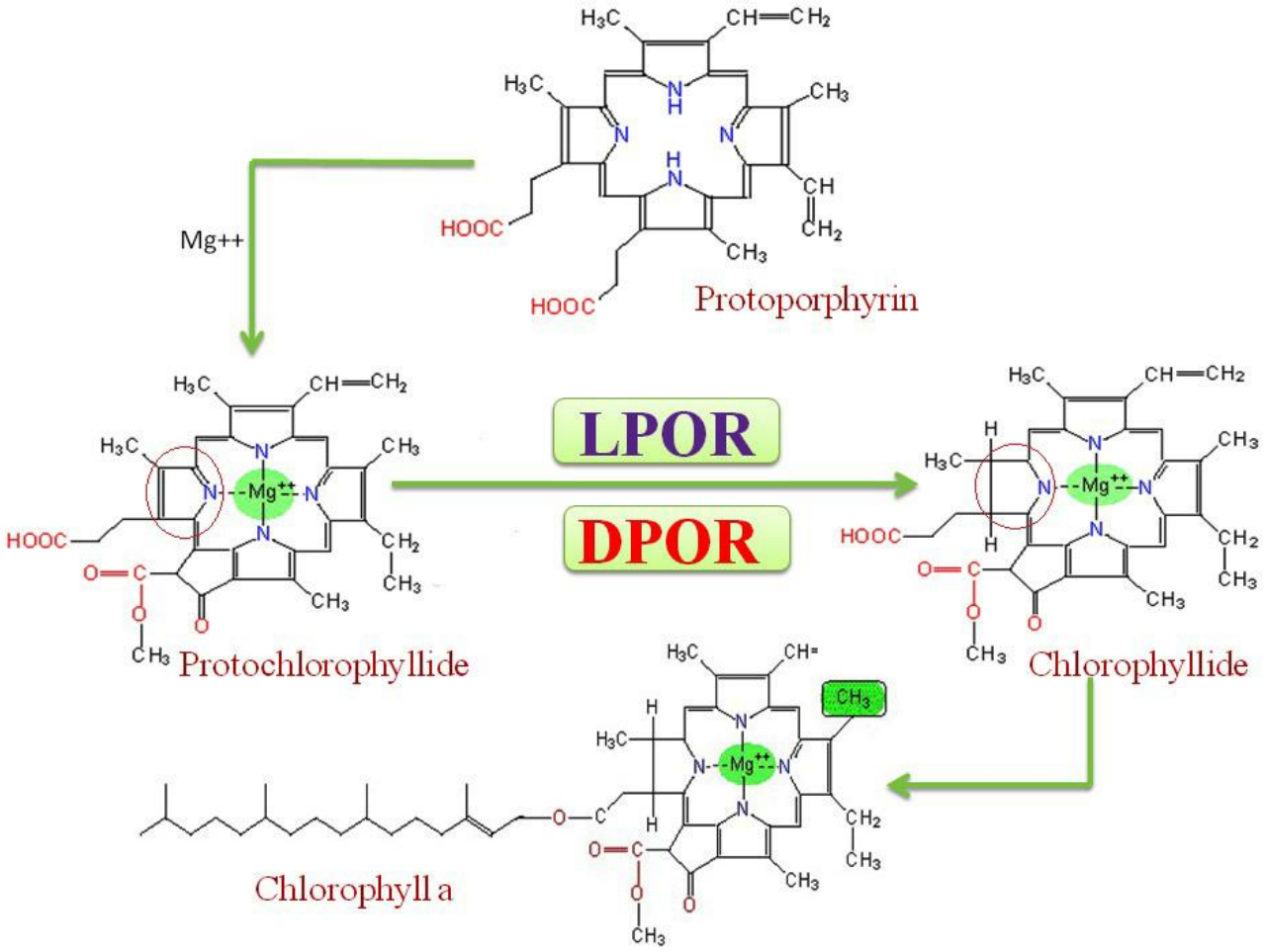

Figure 1. Mechanism of protochlorophyllide reduction by LPOR & DPOR.

(Wakasugi *et al.,* 1994), bryophyta (Ohyama *et al.,* 1986), pteridophyta (Yamada *et al.,* 1992) and algae, chlorophyta (Suzuki and Bauer, 1992; Li *et al.,* 1993) and rhodophyta (Reith and Munholland, 1993), and in the chromosomal DNA of cyanobacteria (Ogura *et al.,* 1992; Fujita, 1996) and an oxygenic photosynthetic bacteria (Burke *et al.,* 1993). These photosynthetic organisms can green and assemble functional chloroplasts in the dark and are ready for photosynthesis upon exposure to light (Yamazaki *et al.,* 2006; Kusumi *et al.,* 2006). However, darkness promotes etiolation in the seedlings of angiosperms because they lack DPOR genes. Consequently, the distribution of DPOR and LPOR entails that the two systems coexisted throughout evolution from the cyanobacteria to the gymnosperms but that the genes encoding DPOR subunits were lost during the evolution from gymnosperms to angiosperms and, as a result, lost their ability to green in the dark (Shi and Shi, 2006; Demko *et al.,* 2009). Hence, the reduction of protochlorophyllide to chlorophyllide in the absence of light in most dark-grown organisms, other than etiolated angiosperms, correlates with the presence of the *ChlL, ChlN*, and *ChlB* genes (Shi and Shi, 2006). To date, the sequenced plastid genomes have revealed that plastomes either encode all three subunits or completely lack these genes. Nevertheless, it has been observed experimentally that ChlB is most essential as compared to other subunits encoded by *ChlL* and *ChlN* genes for light independent protochlorophyllide reduction in chlorophyll biosynthesis (Karpinska *et al.,* 1997; Cahoon and Timko, 2000; Fujita *et al.,* 2015). Disruption of the *ChlB* gene from plastome of *Chlamydomonas reinhardtii* developed yellow mutants (Liu *et al.,* 1993), whereas null mutants of cyanobacterium failed to synthesize chlorophyll under dark conditions (Fujita *et al.,* 1996). However, disruption of *ChlB* from liverwort *Marchantia polymorpha* L. resulted in low chlorophyll synthesis in dark and revealed the importance of *ChlB*, which is particularly required for protochlorophyllide reduction under short day conditions (Ueda *et al.,* 2014).

These research findings clearly suggest that the ChlB has a pivotal role in the activity of DPOR. Apart from its important role in chlorophyll biosynthesis, *ChlB* also has different other functions, which enhance its significance. Ribulose 1, 5-bisphosphate carboxylase/oxygenase (*rbcL*) gene has been broadly used in phylogenetic analysis of plants as a molecular marker. Evolutionary rate of *ChlB* is much higher as compared to *rbcL*. Therefore, *ChlB* sequence can also be used as a molecular marker in phylogenetic analyses. Guo *et al*. (2006) used *ChlB* as a DNA marker for identification and discrimination of *Ephedra* species and the crude drugs derived from them, which has been used in Japanese/Chinese traditional medicine. However, ectopic expression of *ChlB*-coding gene in the chloroplasts of tobacco has resulted in root development when transgenic tobacco plants were grown under dark conditions. Further, the chlorophyll synthesis in transgenic plants compared to wild type was more efficient upon transfer from dark to light. These research results clearly indicated that the *ChlB* gene also contribute in regulation of certain developmental processes in addition to be a DPOR core subunit (Nazir and Khan, 2012).

Different genetic and molecular studies revealed that the LPOR and subunits of DPOR have no sequence similarity and suggested that the photosynthetic organisms evolved these mechanisms autonomously. Complete details of DPOR mechanism of action is not revealed to date. Few questions are still unanswered that how DPOR catalysis the stereospecific protochlorophyllide reduction, how its subunits interact with each-other, how it is evolved and how it structurally related to nitrogenases (Fujita and Bauer, 2000; Muraki *et al.,* 2010). Further, complete structural and functional annotation of ChlB subunit is also not elaborated up till now. Thus, there is a strong need to explore ChlB structure and function which would be helpful in understanding the mechanism of DPOR. Keeping in view the importance and applications, we have performed structural as well as functional annotation of *Pinus thunbergii* ChlB subunit using different bioinformatics tools. Gymnosperm *Pinus thunbergii,* a Japanese Black Pine has the ability to reduce protochlorophyllide to chlorophyllide in light independent fashion whereas all of the angiosperms have lost this ability during the process of evolution (Armstrong, 1998). The present study offered novel insights into the structural features of ChlB and mutation analyses substantiated the central role of predicted catalytic residues in the structure of the ChlB. We conclude that the generated information could facilitate researchers in engineering chlorophyll pathway to improve photosynthesis in plants.

MATERIALS AND METHODS

In this study, various desktop-based applications were employed to characterize ChlB sequence, homology modeling, molecular docking, protein-protein interaction and mutation analysis including ClustalX v2.0 (Larkin *et al.,* 2007), MEGA v6.0 (Tamura *et al.,* 2013), Modeller v9.11 (Webb *et al.,* 2014), UCSF Chimera (Pettersen *et al.,* 2004), Molecular Operating Environment (MOE, 2005) and PyMol (Seeliger and de-Groot, 2010).

***Sequence characterization*:** Total genomic DNA of *Pinus thunbergii* was extracted using CTAB–isopropanol precipitation method (Paterson *et al.,* 1993) and further used as template in PCR (*Eppendorf MasterCycler, F. Hoffmann-La Roche Ltd.,* Foster city, California, USA). The primers (S4: 5'-GAT ATC AAA TTA GCC CA-3'; S5: 5'-GAT ATC TTA ATG CTT CT-3') were designed to amplify *ChlB* gene with suitable restriction sites on either ends. The PCR program was set for 30 cycles at 95°C for 2 min., 56°C for 2 min. and 72°Cfor 3 min., with a step of 72°C for 10 min. The resulting 1.5 kb DNA fragments were purified using Gel-Elution Kit (Thermo Scientific, Waltham, Massachusetts, USA). The PCR products were then ligated into PTZ57R/T. Restriction analysis were performed to confirm the cloning of desired gene that was also sequenced for sequence confirmation. The nucleotide sequence of cloned *ChlB* gene was translated into protein sequence using Translation tool of JustBio online server (http://www.justbio.com/index.php?page=translator). The physicochemical properties of ChlB sequence were calculated by ProtParam tool of EXPASy server (Gasteiger *et al.,* 2003). In order to recognize the similar sequences, key/conserved residues accountable for catalytic activity of the ChlB and to infer phylogeny, multiple sequence alignment of ChlB with other closely related species having known ChlB protein was performed using ClustalX2 software (Larkin *et al.,* 2007). ESPript-3 (http://espript.ibcp.fr/ESPript/ESPript/; Robert and Gouet, 2014), SOPMA (Geourjon *et al.,* 1995) and PSIPRED Server (McGuffin *et al.,* 2000) were used to predict the secondary structure of *ChlB*.

***Structural characterization*:** Homology modeling was carried out using Modeller (v9.11), because a good quality template was available for structure prediction. The retrieved sequence of ChlB was used as query sequence in PSI-BLAST against Protein Data Bank (PDB) to search out a potential template for homology modeling (Bernstein *et al.,* 1978). After selecting template, the alignment between template and ChlB sequence was generated by align2d function of Modeller. Once a target-template alignment was constructed, Modeller calculated 3D models of the target completely automatically by using its auto model class. Over 30 models were built and analyzed for their quality and steriochemical properties through PROCHECK Ramachandran plot (http://services.mbi.ucla.edu/PROCHECK/; Laskowski *et al.,* 1996), ProSA-web Z-score (https://prosa.services.came.sbg.ac.at/prosa.php; Wiederstein *et al.,* 2007), Qmean plot

(https://swissmodel.expasy.org/qmean/cgi/index.cgi; Benkert *et al.,* 2009), ERRAT (http://services.mbi.ucla.edu/ERRAT/; Colovos *et al.,* 1993) and VERIFY 3D (http://services.mbi. ucla.edu/Verify_3D/; Liithy *et al.,* 1992). Further, Root Mean Squared Deviation (RMSD), superimposition of query and template structure analyses were performed using UCSF Chimera and FATCAT server (http://fatcat.burnham.org/; Ye and Godzik, 2004).

***Molecular docking studies*:** Molecular docking analyses were performed using MOE (Molecular Operating Environment) software. Best predicted ChlB model on the basis of model evaluation analysis was selected and optimized by minimizing its energy with parameters (Force Field: AMBER99, Gradient: 0.05). The minimized structure was used as the receptor protein for docking. MOE site finder tool was used to find out the active site where ligand can bind. To select a suitable ligand, ChlB was docked against whole ZINC database contacting 35 million potential compounds in 3D ready to dock format (Irwin and Shoichet, 2005). MOE docking program with parameters (Rescoring function: London dG, Placement: Triangle matcher, Retain: 10, Refinement: Force field, Rescoring 2: London dG) was used to bind the selected ligand with ChlB protein.

***Protein-Protein interaction analysis*:** The interaction studies were carried out using ClusPro (Comeau *et al.,* 2004). DPOR have three subunits ChlB, ChlL and ChlN. These subunits interact with each-other and play a critical role in chlorophyll biosynthesis (Nomata *et al.,* 2005; Heyes and Hunter, 2009; Fujita and Bauer, 2012). In this study, ChlB used as a receptor and ChlL, ChlN were used as ligands to study protein-protein interaction. No three-dimensional structures were found of *Pinus thunbergii* ChlL and ChlN proteins. Therefore, we predicted 3D structure of ChlL and ChlN by using homology modeling method through Modeller (v9.11) (Webb *et al.,* 2014). Furthermore, Pymol software (Seeliger and de-Groot, 2010) was used to analyze and visualize the interactions between receptor and ligand complexes.

***Mutation analysis*:** NetSurfP an online program was used to identify the surface accessibility of all predicted residues (Petersen *et al.,* 2009). After examining the surface accessibility, possible effects of mutations on stability and functionality of ChlB protein were determined using bioinformatics tools SNAP, predicting the effect of mutation on protein functionality (Bromberg and Rost, 2007) and I-Mutant2.0, predicting the effect of mutation on protein stability (Capriotti *et al.,* 2008).

RESULTS

***Sequence characterization*:** In this study, *ChlB* gene was isolated from *Pinus thunbergii* using gene-specific primers in Polymerase Chain Reaction. Amplified gene fragment was cloned into TA cloning vector (*PTZ57R/T, Fermentas,* Waltham, Massachusetts, USA). Ligation of amplified gene fragment into the TA cloning plasmid vector was confirmed using *Eco*RV restriction enzyme, engineered in both (forward & reverse) primers (Fig. 2).

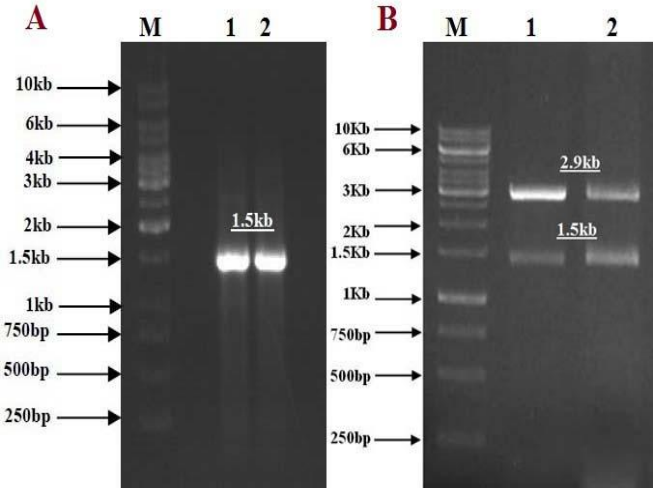

Figure 2. A) PCR amplification of *ChlB* gene, B) Restriction confirmation of cloned *ChlB* gene intopTZ57R/T vector with Ecor321 restriction enzyme.

Further in this study, the translated sequence of the *ChlB* gene was characterized using ProtParam tool of EXPASY (Gasteiger *et al.,* 2003). The calculated isoelectric point (6.16) showed that the ChlB is an acidic protein of 57741.1 Dalton molecular mass. The calculated isoelectric point is very important and it will significantly facilitate researchers in separation of ChlB from crude extract using polyacrylamide gel by isoelectric focusing. The measured extinction coefficient (43485) is useful to determine the concentration of ChlB in solution. Instability complex (38.09) and aliphatic index (89.49) values are indicating the stability of ChlB. The calculated values show that the ChlB is thermo-stable. GRAVY index (-0.243) helps in indicates the solubility of proteins and a negative value of GRAVY index revealed that the ChlB is hydrophilic in nature (Atsushi, 1980). Recognition of catalytic residues is a key to explore functions of the proteins and their phylogeny. Further, mutagenesis analyses of catalytic residues would be possible. Multiple sequence alignment helps to identify the conserved residues among closely similar or homolog sequences. The ChlB sequences from *Pinus_massoniana* (Accession No: YP_008082244.1), *Pinus_brutia* (Accession No: AET44790.1), *Pinus_taeda* (Accession No: YP_008082317.1), *Pinus_radiata* (Accession No: AET46004.1), *Pinus_pinea* (Accession No: AET46575.1), *Pinus_serotina* (Accession No: AET45645.1) and *Pinus_glabra* (Accession No: AET48443.1) were selected because they show significant identity and considerable

sequence length coverage. The result of multiple sequence analyses revealed noteworthy conserved key residues which have been highlighted with red color shown in Figure 3.

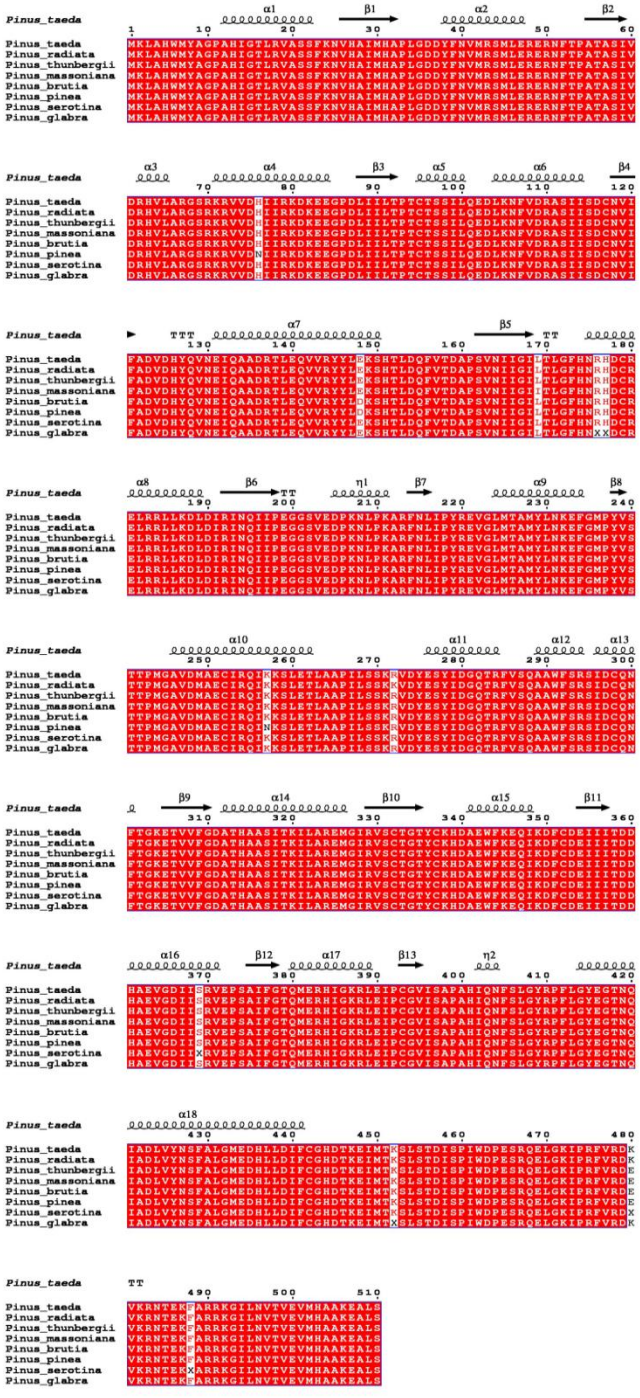

Figure 3. Multiple sequence alignment of ChlB from *P. thunbergii* with highly similar sequences of *P._massoniana* (YP_008082244.1), *P._brutia* (AET44790.1), *P._taeda* (YP_008082317.1), *P._radiata* (AET46004.1), *P._pinea* (AET46575.1), *P._serotina* (AET45645.1) and *P._glabra* (AET48443.1).

Accurate functional annotation for genes/proteins is an important step in the era of high throughput genome and gene sequencing. Most of functional annotation methods used today are comparative in nature and do not take benefit from phylogenetic perspective. Evolutionary analysis coupled with homology bases multiple sequence alignment greatly facilitate the researchers of biology in functional predictions of many genes/proteins (Eisen and Wu, 2002; Dereeper *et al.,* 2008). Phylogenetic dendogram was constructed through ClustalX2 software, using neighbor joining method based on multiple sequence alignment of ChlB protein, with other closely related families and analyzed with the help of MEGA v6.0 (Tamura *et al.,* 2013) software (Fig. 4).

Secondary structure information is very useful to determine protein structure, function and solubility. High fraction of helices in any protein structure makes it flexible for folding and increases its interaction with other proteins (Tokuriki and Tawfik, 2009). Different tools have been used to predict secondary structure of proteins which have success rates ranging from 56 to 70%. We choose SOPMA and PSIPRED Servers because these servers have gained additional 4% prediction power and attained 73.2% success rates as compared to other tools used now a day (Geourjon *et al.,* 1995; McGuffin *et al.,* 2002). Secondary structure predicted results of ChlB protein showed that it contains alpha helixes 46.27%, extended strands 15.10%, beta turns 7.84% and random coils 30.78%.

***Structural characterization*:** Known protein sequences in the result of genome sequencing projects are increasing day by day. But, only 1% sequences have been experimentally characterized for structure and functional determination. Computational proteins functional characterization is one of the fundamental research areas in biology which have the potential to bridge the proteins sequences and structures slit. Computational protein functional characterization is based on accurate 3D structure prediction of proteins by homology modeling method in absence of an experimentally determined structure of protein. Different tools have been developed for 3D structure prediction but Modeller v9.11 proved one of best tool for this task in many studies (Fiser and Sali, 2003; Eswar *et al.,* 2008). Modeller v9.11 was used for homology modeling. We selected template by using PSI-BLAST against PDB. The selected template (Chain B of *Thermosynechococcus elongates* dark operative protochlorophyllide oxidoreductase (chlN-ChlB) Complex; PDB ID: 2XDQ_B) showed total score 743, RMSD value with template sequence 2.40Å, identity 68%, E-value 0.0 and query coverage 100%. By using this template in Modeller total 30 models were developed. The best model was selected on the basis of structural evaluation with comparison to template and stereochemical analysis.

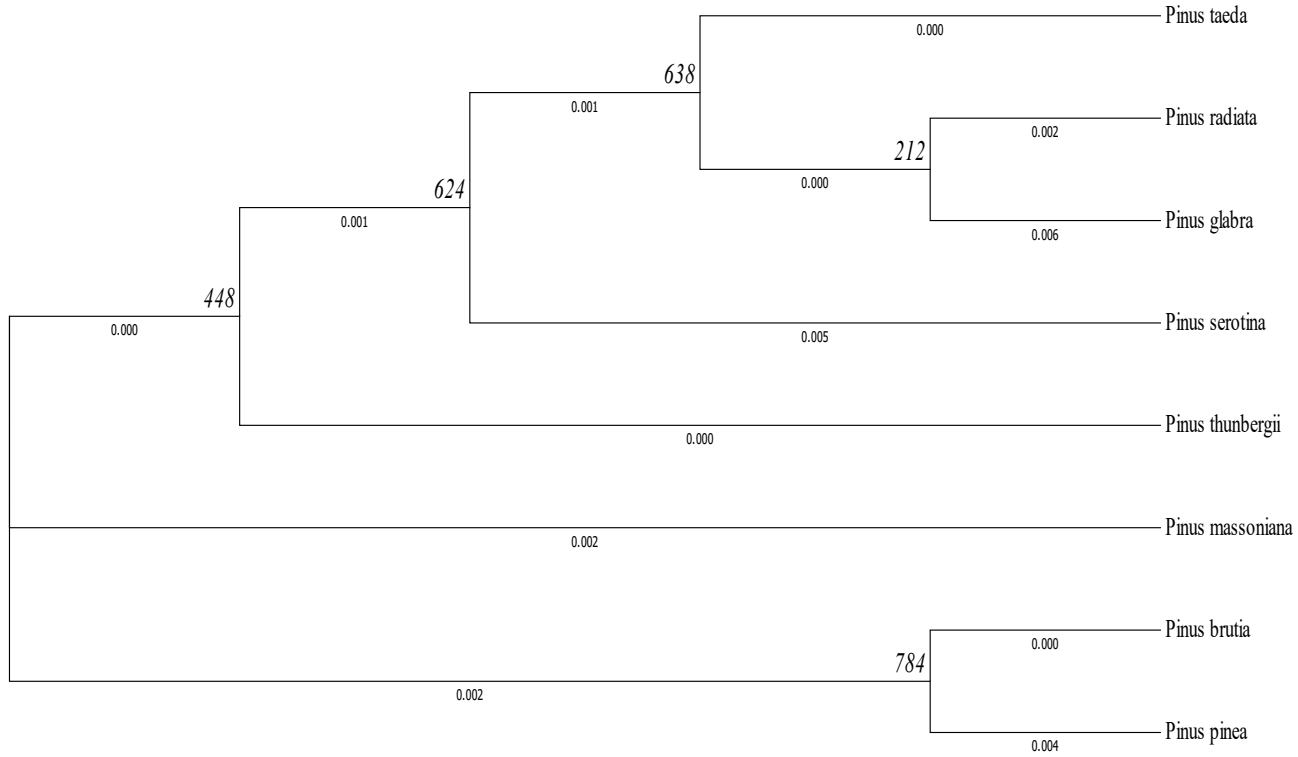

Figure 4. Phylogenetic tree based upon multiple sequence alignment of ChlB from *Pinus thunbergii* with highly similar sequences.

Structural evaluation and superimposition of selected model with the template was done by using UCSF Chimera software and FATCAT server and result showed a very low RMSD of 0.20Å. Z-score calculated with ProSA-web server. Z-score of 9.1 and RMSD value 0.20Å suggested that both template and ChlB protein have similar folds. To analyze the compatibility of predicted 3D model with its own amino acids sequence ERRAT and Verify-3D servers have been used. Results revealed that 99.2% of the residues had an average 3D-1D score of <0.2, predicting that the model is well-matched with its sequence. Ramachandran plot is based on the distribution of backbone dihedral angles and it is a simple method to judge the quality of a protein structure (Hooft *et al.,* 1997). ProCheck server was used to calculate Ramachandran plot, results showed that 96% of residues were in most favored region and 99.2% are in allowed region which strengthen our results and prove the quality of predicted ChlB 3D structure. However, Qmean plot and knowledge based energy graph also show favorable results. Detail of structural characterization is shown in Figure 5. To facilitate the other researcher, predicted model was submitted to Protein Model Database (PMD) and has been assigned the following PMID-0080059.

***Molecular docking studies*:** 4-[(2-oxobenzimidazol-5-yl)sulfonylamino]-N-(p-tolyl)benzamide (ZINC ID: 39795572, Molecular formula: $C_{21}H_{18}N_4O_4S$, Molecular weight: 422.46g/mol) was selected as a ligand for ChlB protein because it show strong hydrogen bonding with Arg18, Asn175, Glu221, Asp311 and make strong polar contacts with Glu131, Pro399, Ala398, His314, Leu171, Phe173, Met7, Ala312, Tyr336, Ala9, Gly172 amino acid residues of ChlB receptor protein. The calculated value of free binding energy of ligand with protein is -17.93 Kcal/mol and value of *S*-score is -21.52. *S*-score is the value calculated by built-in scoring functions of MOE on the basis of ligand binding affinity with receptor protein after docking. These results clearly indicate that the 4-[(2-oxobenzimidazol-5-yl)sulfonylamino]-N-(p-tolyl)benzamide is best ligand for ChlB. These results also give useful information about the residues in active site of ChlB which can be used in further In-vitro mutagenesis experiments. An interaction model of ligand complexed with ChlB protein has been shown in Figure 6. Furthermore, the predicted 3D structure was used to study ChlB interactions with DPOR other subunits through protein-protein interaction.

***Protein-Protein interaction analysis*:** Predicted model of ChlL and ChlN were submitted to Protein Model Database (PMD) and have been assigned the following PMD ID's 0080062 and 0080063 respectively. Protein-protein interaction analysis is very much useful to identify protein

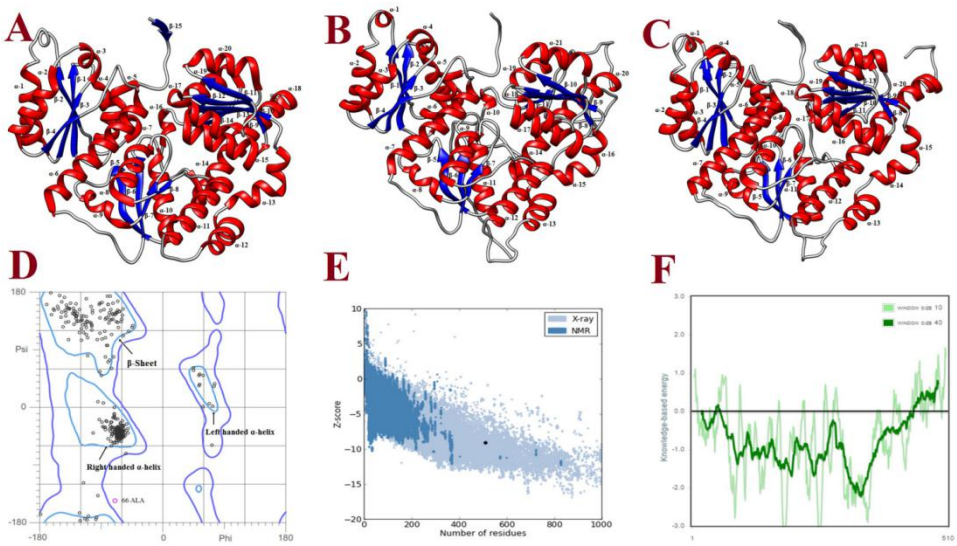

Figure 5. A) 3D structure of selected template (PDB ID: 2XDQ_B), B) Predicted 3D structure of ChlB encoded protein (PMID: 0080059), C) Superimposing of ChlB predicted model (green) and template (Blue) showing high similarity, D) Ramachandran plot of ChlB showing 99.2% amino acids in allowed region, E) Z-score plot ChlB showing quality of model, F) Knowledge based energy graph showing all residues of ChlB model at stable position (dark green line).

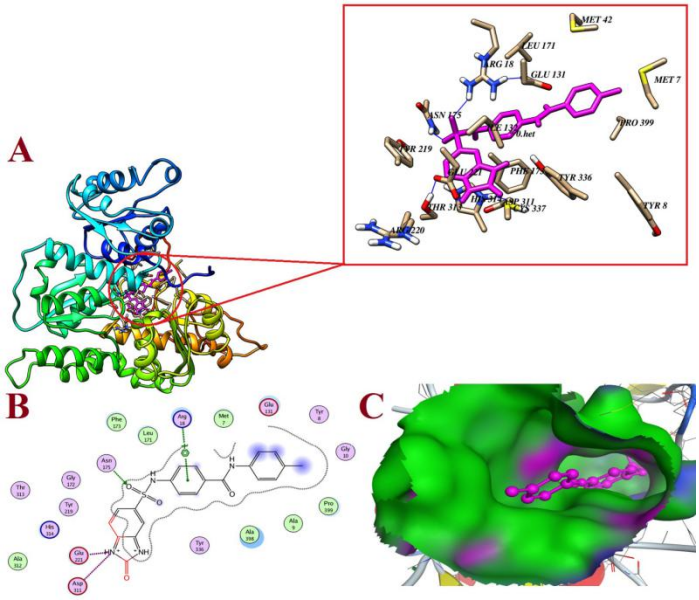

Figure 6. A) 3D binding mode of ligand (4-[(2-oxobenzimidazol-5-yl)sulfonylamino]-N-(p-tolyl)benzamide; magenta color) with ChlB and interaction view, B) 2-Dimensional ligand interaction view with ChlB residues, C) Binding pocket mode of ligand (magenta color) inside active site of *ChlB*.

function in different pathways. ClusPro server was used to carry out interaction analysis of DPOR subunits. It is the only tool available to date which offers fully automated online program for computational protein-protein interaction study (Comeau *et al.,* 2004). In the results predicted by ClusPro, ChlB were found to be interacted with ChlL and ChlN. Furthermore, PyMOL v1.7.4 was used to determine the name and number of residues of ChlB protein interacted with residue of ligand proteins (Table 1). Our results clearly highlighted that the ChlB is interacting with other sub-units of DPOR ChlN and ChlL which strengthen and proved the already purposed hypothesis that the ChlB is involve in light

Table 1. Interacting residues of ChlB protein with ChlL and ChlN.

Receptor	Ligand	Residue Interaction	
		Receptor	Ligand
ChlB	ChlL	Arg475, Gln299, Arg294, Arg264, Arg478, Arg483, Glu480, Arg383, Asn427, Ser295, Thr291, Gln288, Val286, Gln416, Arg18, Asn175, Glu221, Asp311	Glu259, Arg158, Asp126, Asp37, Glu193, Pro271, Asn270, Met1, Arg184, Asn281, Asn284, Asn151, Arg181, Arg210
ChlB	ChlN	Arg388, Leu3, Met7, Tyr38, Asn37, Asn51, Ile461, Arg79, Gln101, Glu49, Glu84, His51, Tyr409, Phe39, Arg18, Asn175, Glu221, Asp311	Glu62, Arg91, Gln52, Asn53, His19, Lys467, Val466, Asn17, Ser380, Glu72, Asn402, Thr462, Ile59, Tyr49

Table 2. Effects of mutations on protein structure predicted by I-Mutant 2.0

Predicted catalytic site residues	Decrease in stability if mutate with:	Increase in stability if mutate with:
Arg18	Val, Leu, Ile, Met , Phe, Tyr, Pro, Arg, Trp, Gly, Ala, Ser, Thr, Cys, His, Lys, Gln, Asn, Asp	None
Asn175	Phe, Tyr, Pro, Arg, Trp, Gly, Ala, Ser, The, Cys, His, Lys, Gln, Asn, Asp	Val, Leu, Ile, Met
Glu221	Trp, Gly, Ala, Ser, The, Cys, His, Lys, Gln, Asn, Asp	Val, Leu, Ile, Met, Phe, Tyr, Pro, Arg
Asp311	Trp, Gly, Ala, Pro, Arg, Lys, Asn	Val, Leu, Ile, Met, Phe, Tyr, Ser, The, Cys, His, Gln, Glu

independent photosynthesis and crucial for chlorophyll biosynthesis in dark.

***Mutation analysis*:** NetSurfP tool predicted the surface accessibility and provide information that Arg18 and Glu221 amino acids are exposed while Asn175 and Asp311 buried inside the core structure of *ChlB* protein. Therefore, it was concluded that any mutation in Arg18 and Glu221 may results in higher conformational as well as functional changes in ChlB protein as compared with Asn175 and Asp311 and effect protein stability. SNAP and I-Mutant 2.0 estimated the effect of each amino acid mutation on protein functionality and stability respectively. SNAP give results in numerical values form ranging from -100 to +100. A lower/negative value represents the neutral effect while a higher/positive value represents the functional aberration (Bromberg and Rost, 2007). The output of SNAP revealed that any mutation at Arg18, Glu221 and Asp311 positions potentially affect (<80) the function of ChlB protein. In case of Asn175, only Asn175 mutation affects function of protein while all other mutations were found neutral. I-Mutant 2.0 works on the basic principle of Gibbs free energy principle and calculates AAG values before and after mutation. It creates mutations in protein sequence by replacing the target amino acid with other nineteen amino acids and measures the effect of each mutation on protein stability (Capriotti *et al.,* 2008). The output of I-Mutant 2.0 revealed that any mutation at Arg18 position leads to decrease in stability and potentially damaging to overall protein structure. Meanwhile, mutations at Asn175, Glu221 and Asp311 position showed different response (Table 2).

DISCUSSION

Of the items, photosynthetic pigments (chlorophyll) are considered as very important and crucial targets for improvements (Khan *et al.,* 2007b; Nazir and Khan, 2012). During the biosynthesis of chlorophyll, protochlorophyllide reduces to chlorophyllide that converted to chlorophyll an instantly. The reduction of protochlorophyllide is very important step, which involves two different enzymes (LPOR and DPOR) depending upon availability of light (Armstrong, 1998; Reinbothe *et al.,* 2010). DPOR is a complex of three subunits encoded by chloroplast genes; *ChlB*, *ChlL* and *ChlN* and present in many photosynthetic organisms such as photosynthetic bacteria, green algae and gymnosperms (Reinbothe *et al.,* 2010; Muraki *et al.,* 2010). To date, few questions are remained to be answered. For example, how DPOR catalyses the stereospecific protochlorophyllide reduction, how its subunits interact with each-other, how it is evolved and structurally related to nitrogenases (Fujita and Bauer, 2000; Muraki *et al.,* 2010). Different studies have been carried out to reveal the roles of DPOR subunits. A number of independent studies have revealed that ChlB is essentially important compared to other two subunits (ChlL and ChlN) for light independent protochlorophyllide reduction, hence could be engineered to improve photosynthesis (Karpinska *et al.,* 1997; Cahoon and Timko, 2000; Fujita *et al.,* 2015). But the complete structural and functional annotation of ChlB subunit is yet to be elaborated.

Pace of submission of un-annotated sequences to different biological databases surpass as compared to ones annotated in the lab. Majority of recognized proteins are still to be characterized experimentally. Experimental methods are time consuming, have high costs and labor intensive. Different bioinformatics tools provide insight into the function of various proteins based on their physiochemical properties, sequences, evolutionary history, structure and protein-protein interaction. Therefore, we prefer computational methods to analyze ChlB subunit (Punta and Orfan, 2008; Pitre *et al.,* 2008; Dereeper *et al.,* 2008; Nawaz *et al.,* 2014). For the very first time in present study, we have carried out structural as well as functional annotation of plant's ChlB subunit using different computational biology and bioinformatics approaches. We divided our studies in five parts. In first part, we isolated, cloned, sequenced and characterized the *ChlB* and predicted its physiochemical properties, homology/similarity with closely related species, phylogeny and secondary structure. Multiple sequence alignment is an essential technique for structure and function prediction of protein and phylogenetic analysis (Edgar and Batzoglou, 2006). Information about secondary structure of protein is very helpful to get better target-template alignment during the process of structure prediction. In second part, we predicted 3D structure of ChlB followed by its validation to get more insight into ChlB structure/function. Homology modeling method was utilized to determine the 3D structure of ChlB of *Pinus thunbergii*. Homology modeling is a computational technique in which three-dimensional (3D) protein structure build by using experimentally (NMR, X-Ray crystallography) determined structures of related family members as templates (Bordoli *et al.,* 2008). Although overall structure was well predicted but further analysis were performed to check model quality and found reasonable. In third part, we docked predicted structure with a suitable ligand to identify active site residues which could be target in experimental analysis. Molecular docking involves multiple computational methods to predict the true binding orientations of small molecules within their protein targets (Lengauer and Rarey, 1996). In fourth, we check ChlB interaction with other subunits of DPOR and predict the interacting residues. 3D structure of other DPOR subunits ChlL and ChlN were first predicted by homology method and their interaction with ChlB subunit was determined computationally. Many cellular functions are based on interactions of different proteins. A large number of experimental methods have been applied so far to discover protein-protein interactions. But all experimental techniques are costly, take too much time and need high expertise. However, computational techniques propose solutions for these problems (Pitre *et al.,* 2008). We therefore consider bioinformatics tools for protein-protein interaction prediction. In fifth and last part, we determined the effects of mutations (in predicted active site residues) on protein structure and verify the potential of our prediction. Functional annotation through in vitro mutagenesis is a popular methodology to determine the function of an un-annotated gene/protein sequence (Alonso *et al.,* 2003). But these experimental methods are time consuming and have high costs. We used *in silico* mutagenesis approach in this study to verify further the activity and importance of predicted catalytic site residues in ChlB functioning. Resulted information will ultimately facilitate researchers working on in vitro mutagenesis. Catalytic site amino acid position in protein structure is very much important. The tendency of mutation in exposed amino acids is higher than the buried due to their involvement in core interaction formation necessary for stability of protein. Resulted information showed that mutations at Arg18, Asn175, Glu221 and Asp311 positions could significantly disrupt the overall structure of ChlB protein. These mutation analysis results greatly facilitate researcher working in vitro. The presence of Arg amino acid in the first 20 residues of nuclear encoded plastidic proteins is rare. In-vitro mutagenesis of chloroplastic proteins revealed the role of Arg residue. Mutation in Arg residue disrupts the import and processing of proteins in chloroplast (Pujol *et al.,* 2007). The presence of Arg at position 18 in *Pinus thunbergii ChlB* protein and also its involvement in catalytic site suggest that the predicted residues could be further studied for engineering chlorophyll biosynthesis. The results of our study proposed novel insights into the structural features of ChlB and gives worthwhile information. Hence, we conclude that the information obtained from present study will greatly facilitate the researchers working on DPOR mechanism analysis and possibly be regarded as a step forward to engineer chlorophyll biosynthesis for improved photosynthesis.

Acknowledgments**:** Authors would like to acknowledge Higher Education Commission (HEC), Pakistan and University of Agriculture Faisalabad (UAF), Pakistan for providing facilities for this study.

REFERENCES

Alonso, J.M., A.N. Stepanova, T.J. Leisse, C.J. Kim, H. Chen, P. Shinn and C. Gadrinab. 2003. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Sciences. 301:653-657.

Armstrong, G.A. 1998. Greening in the dark: light-independent chlorophyll biosynthesis from anoxygenic photosynthetic bacteria to gymnosperms. J. Photochem. Photobiol. B. Biol. 43:87-100.

Atsushi, I.K.A.I. 1998. Thermostability and aliphatic index of globular proteins. J. Biochem. 88:1895-1898.

Benkert, P., M. Kunzli and T. Schwede. 2009. QMEAN server for protein model quality estimation. Nucleic Acids Res. 37:510-514.

Bernstein, F.C., T.F. Koetzle, G.J. Williams, E.F. Meyer, M.D. Brice, J.R. Rodgers and M. Tasumi. 1978. The Protein Data Bank: a computer-based archival file for macromolecular structures. Arch. Biochem. Biophys. 185:584-591.

Blankenship, R.E. 2013. Molecular Mechanism of Photosynthesis, 2nd Ed. Wiley-Blackwell, USA.

Bock, R and M.S. Khan. 2004. Taming plastids for a green future. Trends Biotechnol. 22:311-318.

Bordoli, L., F. Kiefer, K. Arnold, P. Benkert, J. Battey and T. Schwede. 2008. Protein structure homology modeling using SWISS-MODEL workspace. Nat. Protoc. 4:1-13.

Bromberg, Y. and B. Rost. 2007. SNAP: predict effect of non-synonymous polymorphisms on function. Nucleic Acids Res. 35:3823-3835.

Burke, D.H., M. Albert and J.E. Hearst. 1993. bchF NBH bacteriochlorophyll synthesis genes of *Rhodobacter capsulatus* and identification of the third subunit of light-independent protochlorophyllide reductase in bacteria and plants. J. Bacteriol.175:2414–2422.

Cahoon, A.B. and M.P. Timko. 2000. Yellow-in-the-dark mutants of *Chlamydomonas* lack the CHLL subunit of light-independent protochlorophyllide reductase. Plant Cell 12:559-568.

Camp, R.G., D. Przybyla, C. Ochsenbein, C. Laloi, C. Kim, A. Danon and K. Apel. 2003. Rapid induction of distinct stress responses after the release of singlet oxygen in Arabidopsis. Plant Cell 15:2320-2332.

Capriotti, E., P. Fariselli, I. Rossi and R.A. Casadio. 2008. Three-state prediction of single point mutations on protein stability changes. BMC Bioinformatics 9:1-9.

Colovos, C. and T.O. Yeates. 1993. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 2:1511-1519.

Comeau, S.R., D.W. Gatchell, S. Vajdaand and C.J. Camacho. 2004. ClusPro: a fully automated algorithm for protein–protein docking. Nucleic Acids Res. 32:96-99.

Seeliger, D. and B.L. de Groot. 2010. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. J. Computer-aided Mol Des. 24:417-422.

Demko, V., A. Pavlovib, D. Valkova, L. Slovakova, B. Grimm and J. Hudak. 2009. A novel insight into the regulation of light-independent chlorophyll biosynthesis in *Larix decidua* and *Picea abies* seedlings. Planta 230:165-176.

Dereeper, A., V. Guignon, G. Blanc, S. Audic, S. Buffet, F. Chevenet and O. Gascuel. 2008. Phylogeny.fr:robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 36:465-469.

Eckhardt, U., B. Grimm and S. Hortensteiner. 2004. Recent advances in chlorophyll biosynthesis and breakdown in higher plants. Plant Mol. Biol. 56:1–14.

Edgar, R.C. and S. Batzoglou. 2006. Multiple sequence alignment. Curr. Opin. Struct. Boil. 16:368-373.

Eisen, J.A. and M. Wu. 2006. Phylogenetic analysis and gene functional predictions: phylogenomics in action. Theor. Popul. Biol. 61:481-487.

Eswar, N., D. Eramian, B. Web, M.Y. Shen and A. Sali. 2008. Protein structure modeling with MODELLER. Methods Mol. Biol. 426:145-159.

Eswar, N., B. Webb, M.A. Marti-Renom, M.S. Madhusudhan, D.Eramian, M.Y. Shen and A. Sali. 2006. Comparative protein structure modeling using Modeller. Curr. Protoc. Bioinformatics 5:1-47.

Fiser, A. and A. Sali. 2003. Modeller: generation and refinement of homology-based protein structure models. Meth. Enzymol. 374:461-491.

Foyer, C.H. and G. Noctor. 2000. Oxygen processing in photosynthesis: regulation and signaling. New Phytol. 146:358-388.

Fujita, Y. 1996. Protochlorophyllide reduction: a key step in the greening of plants. Plant Cell Physiol. 37:411-421.

Fujita, Y., R. Tsujimoto and R. Aoki. 2015. Evolutionary aspects and regulation of tetrapyrrole biosynthesis in *Cyanobacteria* under aerobic and anaerobic environments. Life 5:1172-1203.

Fujita, Y. and C.E. Bauer. 2000. Reconstitution of light-independent protochlorophyllide reductase from purified BchL and BchN-BchB subunits *in vitro* confirmation of nitrogenase-like features of a bacteriochlorophyll biosynthesis enzyme. J. Biol. Chem. 275:23583-23588.

Fujita, Y. and C.E. Bauer. 2012. The light-independent protochlorophyllide reductase: A nitrogenase-like enzyme catalyzing a key reaction for the porphyrin handbook: Chlorophylls and Bilins: Biosynthesis. Synthesis and Degradation. 13:109-133.

Fujita, Y., H. Takagi and T. Hase. 1996. Identification of the *ChlB* gene and the gene product essential for the light-independent chlorophyll biosynthesis in the cyanobacterium *Plectonema boryanum*. Plant Cell Physiol. 37:313-323.

Gasteiger, E., A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel and A. Bairoch. 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 31:3784-3788.

Geourjon, C. and D. Gilber. 1995. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput. Appl. Biosci.11:681-684.

Guo, Y., A. Tsuruga, S. Yamaguchi, K. Oba, K. Iwai, S, Sekita and H. Mizukami. 2006. Sequence analysis of chloroplast *ChlB* gene of medicinal *Ephedra* species

and its application to authentication of *Ephedra* Herb. Biol. Pharma. Bull.29:1207-1211.

Gust, D., T.A. Moore and A.L. Moore. 2001. Mimicking photosynthetic solar energy transduction. Acc. Chem. Res. 34:40-48.

Heyes, D.J. and C.N. Hunter. 2005. Making light work of enzyme catalysis: protochlorophyllide oxidoreductase. Trends. Biochem. Sci.30:642-649.

Heyes, D.J. and C.N. Hunter. 2009. Biosynthesis of Chlorophyll and Barteriochlorophyll. In Tetrapyrroles. Springer New York; pp.235-249.

Hooft, R.W., C. Sander and G. Vriend. 1997. Objectively judging the quality of a protein structure from a Ramachandran plot. Comput. Appl. Biosci. 13:425-430.

Irwin, J.J. and B.K. Shoichet. 2005. ZINC: a free database of commercially available compounds for virtual screening. J. Chem. Inf. Model 45:177-182.

Johnson, J.L., W. Indermaur and K.V. Rajagopalan. 1991. Molybdenum cofactor biosynthesis in *Escherichia coli.* Requirement of the *ChlB* gene product for the formation of molybdopterin guanine dinucleotide. J. Biol. Chem. 266:12140-12145.

Karpinska, B., S. Karpinski and J.E. Hallgren. 1997. The *ChlB* gene encoding a subunit of light- independent protochlorophyllide reductase is edited in chloroplasts of conifers. Curr. Genet. 31:343-347.

Khan, M.S. 2007a. Engineering photorespiration in chloroplasts: a novel strategy for increasing biomass production. Trends Biotechnol. 25:437-440.

Khan, M.S., W. Hameed, M. Nozoe and T. Shiina. 2007b. Disruption of the *psbA* gene by the copy correction mechanism reveals that the expression of plastid-encoded genes is regulated by photosynthesis activity. J. Plant Res.120:421-430.

Kusumi, J., A. Sato and H. Tachida. 2006. Relaxation of function constraints on light-independent protochlorophyllide oxidoreductase in *Thuja*. Mol. Biol. Evol. 23:941-948.

Larkin, M.A., G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam and D.G. Higgins. 2007. Clustal W and Clustal X version 2.0. Bioinformatics. 23:2947-2948.

Laskowski, R.A., J.A.C. Rullmann, M.W. MacArthur, R. Kaptein and J.M. Thornton. 1996. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR. 8:477-486.

Lebedev, N. and M. Timko. 1999. Protochlorophyllide oxidoreductase B-catalyzed protochlorophyllide photoreduction in vitro: insight into the mechanism of chlorophyll formation in light-adapted plants. Proc. Natl. Acad. Sci. USA 96:9954–9959.

Lengauer, T. and M. Rarey. 1996. Computational methods for biomolecular docking. Curr. Opin. Struct. Biol.6:402-406.

Li, J., M. Goldschmidt-Clermont and P. Timko. 1993. Chloroplast-encoded *ChlB* is required for light-independent protochlorophyllide reductase activity in *Chlamydomonas reinhardtii*. Plant Cell 5:1817-1829.

Liithy, R., J.U. Bowie and D. Eisenber. 1992. Assessment of protein models with three-dimensional profiles. Nature 356:83-85.

Liu, X.Q., H. Xu and C. Huang. 1993. Chloroplast *ChlB* is required for light independent chlorophyll accumulation in *Chlamydomonas reinhordhi*. Plant Mol. Biol. 23:297–308.

Masuda, T. and Y. Fujita. 2008. Regulation and evolution of chlorophyll metabolism. Photochem. Photobiol. Sci.7:1131-1149.

Masuda, T. and K. Takamiya. 2004. Novel insights into the enzymology, regulation and physiological functions of light-dependent protochlorophyllide oxidoreductase in angiosperms. Photosynth. Res. 81:1-29.

McGuffin, L.J., K. Bryson and D.T. Jones. 2000. The PSIPRED protein structure prediction server. Bioinformatics 16:404-405.

The Molecular Operating Environment (MOE). 2005.06. Available online at at http://www.chemcomp.com

Muraki, N., J. Nomata, K. Ebata, T. Mizoguchi, T. Shiba, H. Tamiaki and Y. Fujita. 2010. X-ray crystal structure of the light-independent protochlorophyllide reductase. Nature. 465:110-114.

Nawaz, M., N. Iqbal, S. Idrees and I. Ullah. 2014. DREB1A from Oryza sativa var. IR6: homology modelling and molecular docking. Turk. J. Bot. 38:1095-1102.

Nazir, S. and M.S. Khan. 2012. Chloroplast-encoded *ChlB* gene from *Pinus thunbergii* promotes root and early chlorophyll pigment development in *Nicotiana tabaccum*. Mol. Biol. Rep. 39:10637-10646.

Nelson, N. and W. Junge. 2015. Structure and energy transfer in photosystems of oxygenic photosynthesis. Ann. Rev. Biochem. 84:659-683.

Nomata, J., L.R. Swem, C.E. Bauer and Y. Fujita. 2005. Over expression and characterization of dark-operative protochlorophyllide reductase from *Rhodobacter capsulatus*. Biochim. Biophys. Acta Bioenerget. 1708:229-237.

Nomata, J., T. Ogawa, M. Kitashima, K. Inou and Y. Fujita. 2008. NB protein (BchN-BchB) of dark-operative protochlorophyllide reductase is the catalytic component containing oxygen-tolerant Fe-S clusters. FEBS Lett. 582:1346-1350.

Ogura, Y., M. Takemura, K. Oda, K. Yamato, E. Ohta, H. Fukuzawa and K. Ohyam. 1992. Cloning and nucleotide sequence of afrxCORF469 gene cluster of Synechocystis PCC6803: Conservation with liverwort

chloroplast *frxC*-ORF465 and nif operon. Biosci. Biotechnol. Biochem. 56:788-793.

Ohyama, K., H. Fukuzawa, T. Kohchi, H. Shirai, T. Sano, S. Sano, K. Umesono, Y. Shiki, M. Takeuchi, Z. Chang, S. Aota, H. Inokuchi and H. Ozeki. 1986. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature 322:572-574.

Paterson, A.H., C.L. Brubaker and J.F. Wendel. 1993. A rapid method for extraction of cotton (*Gossypium* spp.) genomic DNA suitable for RFLP or PCR analysis. Plant Mol. Biol. Rep. 11:122-127.

Petersen, B., T.N. Petersen, P. Andersen, M. Nielsen and C. Lundegaard. 2009. A generic method for assignment of reliability scores applied to solvent accessibility predictions. BMC Str. Biol. 9:1.

Pettersen, E.F., T.D Goddard, C.C Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng and T.E. Ferrin. 2004. UCSF Chimera-a visualization system for exploratory research and analysis. J. Comput. Chem. 25:1605-1612.

Pitre, S., M. Alamgir, J.R. Green, M. Dumontier, F. Dehne and A. Golshani. 2008. Computational methods for predicting protein–protein interactions. Adv. Biochem. Engin/Biotechnol. 110:247-267.

Pujol, C., L. Marechal-Drouard and A.M. Duchene. 2007. How can organellar protein N-terminal sequences be dual targeting signals? In silico analysis and mutagenesis approach. J. Mol. Biol. 369:356-367.

Punta, M. and Y. Ofran. 2008. The rough guide to in silico function prediction, or how to use sequence and structure information to predict protein function. PLoS Comput. Biol. 4:1-7.

Reinbothe, C., M. El-Bakkouri, F. Buhr, N. Muraki, J. Nomata, G. Kurisu and S. Reinbothe. 2010. Chlorophyll biosynthesis: spotlight on protochlorophyllide reduction. Trends Plant Sci. 15:614-624.

Reith, M. and J. Munholland. 1993. A high-resolution gene map of the chloroplast genome of the red alga *Porphyra purpurea*. Plant Cell 5:465-475.

Robert, X.P. and P. Gouet. 2014. Deciphering key features in protein structures with the new END script server. Nucleic Acids Res. 42:W320-W324.

Schoefs, B. and F. Franck. 2003. Protochlorophyllide Reduction: Mechanisms and evolution. Photochem. Photobiol. 78:543-557.

Shi, C. and X. Shi. 2006a. Expression switching of three genes encoding light-independent protochlorophyllide oxidoreductase in *Chlorella protothecoides*. Biotechnol. Lett. 28:261–265.

Shi, C. and X. Shi. 2006b. Characterization of three genes encoding the subunits of light-independent protochlorophyllide reductase in *Chlorella protothecoides* CS-41. Biotechnol. Progr. 22:1050-1055.

Suzuki, J.Y. and C.E. Bauer. 1992. Light-independent chlorophyll biosynthesis: Involvement of the chloroplast gene chlL (frxC). Plant Cell 4:929–940.

Tamura, K., G. Stecher, D. Peterson, A. Filipski and S.Kumar. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30:2725-2729.

Tanaka, R. and A. Tanaka. 2007. Tetrapyrrole biosynthesis in higher plants. Ann. Rev. Plant Biol. 58:321-346.

Tokuriki, N. and D.S. Tawfik. 2009. Protein dynamism and evolvability. Science 324:203-207.

Tsukatani, Y. and S. Masuda. 2015. Elucidation of genetic backgrounds necessary for chlorophyll a biosynthesis toward artificial creation of oxygenic photosynthesis. Orig. Life Evol. Biosph. 45:367-369.

Ueda, M., A. Tanaka, K. Sugimoto, T. Shikanai and Y. Nishimura. 2014. *ChlB* requirement for chlorophyll biosynthesis under short photoperiod in *Marchantia polymorpha* L. Genome Biol. Evol. 6:620-628.

Wakasugi, T., J. Tsudsuki, S. Ito, K. Nakashima, T. Tsudsuki and M. Sugiura. 1994. Loss of all ndh genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*. Proc. Natl. Acad. Sci. USA. 91:9794-9798.

Wiederstein, M. and M.J. Sippl. 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res.35:407-410.

Willows, R.D. 2003. Biosynthesis of chlorophylls from protporphyrin IX. Nat. Prod. Rep. 20:1-16.

Yamada, K., M. Matsuda, Y. Fujita, H. Matsubara and M. Sugai. 1992. A frxC homolog exists in the chloroplast DNAs from various pteridophytes and in gymnosperms. Plant Cell Physiol. 33:325-327.

Yamazaki, S., J. Nomata and Y. Fujita. 2006. Differential operation of dual protochlorophyllide reductases for chlorophyll biosynthesis in response to environmental oxygen levels in the cyanobacterium *Leptolyngbya boryana*. Plant Physiol. 142:911-922.

Ye, Y. and A. Godzik. 2004. FATCAT: a web server for flexible structure comparison and structure similarity searching. Nucleic Acids Res. 32:582-585.