South African Journal of Botany 141 (2021) 219–226

Contents lists available at ScienceDirect

South African Journal of Botany

journal homepage: www.elsevier.com/locate/sajb

# Functional annotation and sequence-structure characterization of a hypothetical protein putatively involved in carotenoid biosynthesis in microalgae

Parminder Kaur Narang[a,b], Jyotirmayee Dey[a], Soumya Ranjan Mahapatra[a], Mrinmoy Ghosh[c], Namrata Misra[a,c], Mrutyunjay Suar[a,c], Vijay Kumar[d,*], Vishakha Raina[a,*]

[a] School of Biotechnology, Kalinga Institute of Industrial Technology (KIIT), Deemed to be University, Bhubaneswar 751024, India
[b] SGTB Khalsa College, Delhi University, Delhi 110007, India
[c] KIIT-Technology Business Incubator (KIIT-TBI), Kalinga Institute of Industrial Technology (KIIT), Deemed to be University, Bhubaneswar 751024, India
[d] Department of Biotechnology, Lovely Faculty of Technology and Sciences, Lovely Professional University, Phagwara, Punjab 144402, India

ARTICLE INFO

*Article History:*
Received 31 March 2021
Accepted 7 April 2021
Available online xxx

Edited by Vishakha Raina

*Keywords:*
Carotenoids, Microalgae, Functional annotation, Chlamydomonas reinhardtii

ABSTRACT

Microalgae are widely used for commercial production of carotenoids and are an excellent model for the study of carotenogenesis metabolism in plants. Genetic engineering of microalgae to increase or knockdown enzymatic activities is one viable approach to accumulate desired carotenoids with higher productivity. Recently, various genome sequencing projects have led to accumulation of plethora of algal genomic data; however, majority of the protein coding genes are yet to be functionally characterized and these gene products are termed as hypothetical proteins. Identifying and functional annotation of these proteins can pave the way for better understanding of algal carotenoid biosynthetic pathway and can indicate targets for genetic modification. In the present study, bioinformatics analysis was employed to predict the function of a hypothetical protein from the model microalga *Chlamydomonas reinhardtii*. The retrieved sequence was functionally and structurally characterised through determination of their physico-chemical properties, sub-cellular localisation, conserved domain and motif search. Conclusively, the results possibly define the putative role of the protein as phytoene desaturase enzyme catalysing the rate limiting step in biosynthesis of important carotenoid pigment known as lycopene. Furthermore, homology modelling and docking analysis helped to determine the three-dimensional structure and underpinning residues in the active site region that are involved in interaction with the cofactor flavin adenine dinucleotide. The analysis reported here can be potentially utilised for further experimental validation and to expedite genetic improvement efforts for increased carotenoids productivity.



## 1. Introduction

Recently carotenoids from natural sources are in huge demand in food, nutraceutical and cosmetic industries (Noviendri et al., 2011; Saini and Keum 2017; Sathasivam and Ki. 2018). Various carotenoids pigments are widely used as active ingredients with biological activity because of their color and nutritional properties (Noviendri et al., 2011; Fernandez-Garcia et al., 2012). Additionally, due to antioxidant properties, carotenoids are also used as food preservation agents to slow down the oxidation process (Vilchez et al., 2011; Rashid and Azlina 2020). The animal feed sector, particularly aquaculture is one of the highest consumers of carotenoids as dietary supplements to enhance desired colour of various farmed fish (Novoveska et al., 2019; Nakano and Wiegertjes. 2020). Currently, microalgae are the most sustainable natural source of the production of wide variety of carotenoids including astaxanthin, $\beta$-carotene, lutein, lycopene, canthaxanthin and zeaxanthin (Das et al.,2007; Rammuni et al., 2019). However, for commercial large-scale production of carotenoids, genetic engineering is the most promising approach (Fraser et al., 2009; Varela et al., 2015).

Despite the high value of carotenoids and the advantages of algal platforms, there are few reported genetic engineering efforts towards carotenoids augmentation in these organisms (Gimpel et al., 2015; Leu and Boussiba 2014). Against this backdrop, a deeper understanding of genes and enzymes underpinning the carotenoid biosynthetic pathway is essential to design strategies for over expression of target enzymes (Rosenberg et al., 2008). However, unlike several comparative genomics and *in silico* studies on lipid biosynthetic pathway have been recently reported in microalgae (Misra et al., 2012; Misra et al., 2013; Misra et al., 2014) not many efforts have been undertaken to

* Corresponding authors.
*E-mail addresses:* vijay.srm23@gmail.com (V. Kumar), vraina@kiitbiotech.ac.in (V. Raina).

https://doi.org/10.1016/j.sajb.2021.04.014
0254-6299/

unravel the structure-function of crucial enzymes that play a vital role in carotenogenesis (Misra and Panda, 2013). Furthermore, approximately half of the protein encoding genes in most algal genomes are classified as hypothetical proteins (HPs) whose function have not been experimentally determined and this plethora of uncharacterized proteins probably have their own significance in the total proteomic platform of an organism (Sahoo et al., 2020). In-depth physicochemical characterization and elucidation of three-dimensional structure of HPs can pave the way for better understanding of carotenoids metabolic pathway and also aid in genetic engineering strategies.

Phytoene desaturase (PDS) is an essential carotenoid biosynthetic enzyme in plants and microalgae that introduces two double bonds into the symmetric, colorless phytoene substrate which ultimately leads to the red-colored lycopene (Zhu et al, 2005; Koschmieder et al., 2017). Several experimental studies in the past have corroborated that PDS is a promising target for increased synthesis of carotenoids (Mann et al., 1994). *Chlamydomonas reinhardtii* is the model algae for performing genetic manipulation in carotenoid accumulation (Cordero et al., 2011; Liu et al., 2013). The present study aims to carry out the characterization of the hypothetical protein A8J3K3 from *C. reinhardtii* using a combination of bioinformatics tools based on homology search against functionally identified proteins. Conserved domain and motif analysis confirmed that the putative role of HP is PDS enzyme in the carotenoid biosynthetic pathway. Physicochemical analysis, the three-dimensional structure elucidations provided insights into its binding interaction with ligand, flavin adenine dinucleotide (FAD).

## 2. Materials and methods

### 2.1. Sequence retrieval

The amino acid and nucleotide sequence of a hypothetical protein (UniProt Accession Id: A8J3K3) from *C. reinhardtii was* retrieved from its whole genome database (https://mycocosm.jgi.doe.gov/Chlre4/Chlre4.home.html) utilising the known homologous sequence by BLAST search. The complete methodology and various bioinformatics tools used in this study is schematically represented in Fig. 1. The multiple sequence alignment of homologous sequences was obtained using BioEdit biological sequence alignment editor (Hall et al., 2011).

### 2.2. Functional prediction

Pfam, SUPERFAMILY and CATH were used to predict the conserved domains present in order to characterize protein function at the molecular level. Pfam database (https://pfam.xfam.org/) provides a complete and accurate classification of protein families and domains. SUPERFAMILY (http://www.supfam.org/SUPERFAMILY/), a database based on a collection of hidden markov model unravels the protein domain annotation at both superfamily and family levels. CATH (https://www.cathdb.info/) is a database for hierarchical domain classification of protein structures at four major levels, Class (C), Architecture (A), Topology (T) and Homologous superfamily (H).

### 2.3. Sub-cellular localization prediction

Determining sub-cellular localization is crucial for understanding the organelle location of the protein which provides clue for protein function analysis. TargetP (http://www.cbs.dtu.dk/services/TargetP/), ChloroP (http://www.cbs.dtu.dk/services/ChloroP/) were employed for prediction of subcellular localization of the eukaryotic proteins using default parameters.

### 2.4. Analysis of physico-chemical properties

The ProtParam (http://web.expasy.org/protparam/) tool of ExPASy was used for the analysis of the physical and chemical properties of the targeted protein sequence. The properties including aliphatic index (AI), grand average of hydropathy (GRAVY), iso-electric point (pI), Instability Index (II) and molecular weight were analyzed.

### 2.5. Secondary structure analysis

The SignalP 5.0 server (http://www.cbs.dtu.dk/services/SignalP/) was used to detect the signal peptides and also to localize the cleavage site in the peptide chain from amino acid sequences**.** Solubility of the protein is predicted by the Scratch Protein Predictor tool (http://scratch.proteomics.ics.uci.edu/).The SCRATCH software suite includes predictors for secondary structure, relative solvent accessibility, disordered regions, and disulfide bridges. The number of transmembrane helices present in integral membrane protein was predicted by HMMTOP Server, v.2.0 (http://www.enzim.hu/hmmtop/), a widely used bioinformatics tool, for prediction of the propensity of a protein to be a membrane protein.The secondary structure of a protein was analyzed using SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) tool. The disordered regions of protein were identified by PSIPRED tool (http://bioinf.cs.ucl.ac.uk/psipred/) with default parameters. CYSPRED (http://gpcr.biocomp.unibo.it/cgi/predictors/cyspred/pred_cyspredcgi.cgi) tool was used to predict disulfide bonds between cysteine residues which play an important role in the folding of the protein to form a functional and stable conformation.

### 2.6. Homology modeling and docking

For template selection, BlastP search was performed against the PDB database (http://www.rcsb.org/pdb/) to find X-ray crystallographic structures with maximum identity and lower E-value. Also, SWISS-MODEL (https://swissmodel.expasy.org/) was used for template selection. Finally, the crystal structure of gamma-carotenoid desaturase derived at 1.97Å (PDB ID: 4REP) and having 23% sequence identity (85% sequence similarity) was used as a template for modeling. Target -template alignment was determined using BioEdit tool. The tertiary structure was predicted using ROBETTA server (Kim et al., 2004) and visualised using Pymol (Yuan et al., 2016). The stereochemical quality of the predicted structure was assessed using various structure validation tools like Verify3D (Eisenberg et al., 1997), Errat (Lengths and Angles, 2018). Ramachandran plot was generated using the Procheck server (Laskowski, Roman A., et al. 1993). Active site of the protein was determined by the computed atlas of surface topography of proteins (CASTp) server (http://sts.bioe.uic.edu/castp/index.html?201l), which provides an online resource for locating, delineating, and measuring concave surface regions on the three-dimensional structures of proteins. Further, the structure of FAD in SD format was retrieved from Pubchem and converted into 3D structures in pdbqt format through open babel module of molecular modeling programme (O'Boyle et al., 2011). The Autodock vina was used for docking environment at its default parameters (Trott et al.,2010). The input protein and ligands were generated in pdbqt format with the help of Autodock tools (Morris et al., 2009) and the grid box was defined around the entire active site as predicted by CastP tool to accommodate ligands. Gasteiger charges were applied and added polar hydrogen to the receptor and ligands. The receptor was kept rigid during the docking procedure. The grid box was defined around the active site in the range of $104 \times 84 \times 110$ for protein to accommodate ligand. The vina parameter “exhaustiveness” was set to the value of 8.

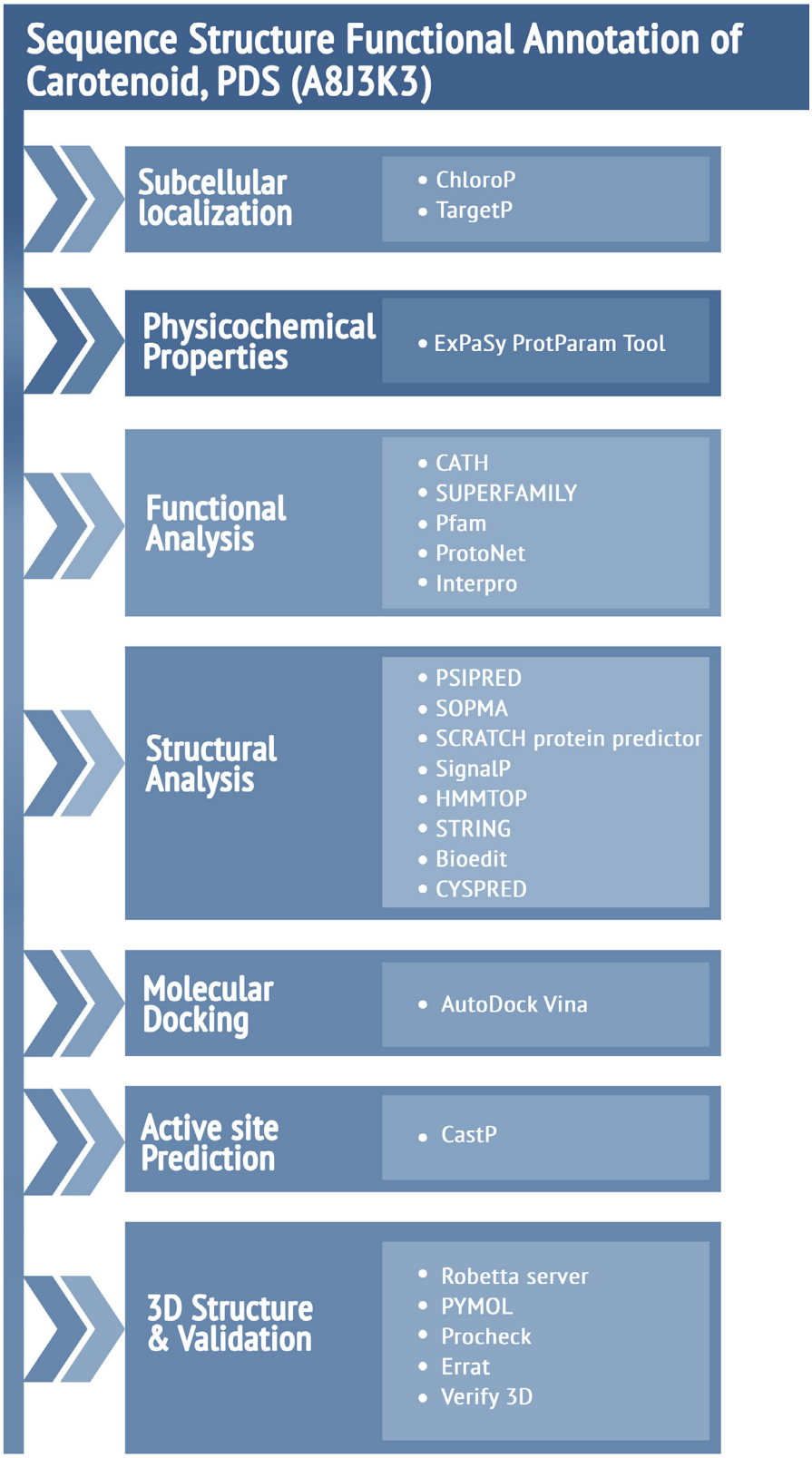


**Fig. 1.** Schematic representation of bioinformatics approach employed for sequence-structure annotation of the protein.

## 3. Results and discussion

### 3.1. Sequence alignment and functional analysis

Functional domains and motifs are conserved signatures of protein families and can often be used as reference for the prediction of protein function at the cellular level. Various publicly available databases such as Pfam, SUPERFAMILY, CATH, ProtoNet, InterPro was used for characterization of domain architecture and MEME tool was used for motif discovery and the results are summarized in Supplementary Table 1. The Pfam databases suggests that the hypothetical protein (A8J3K3) contains the domain for flavin containing amine


**Fig. 2.** Sequence alignment of A8J3K3 with known homologous sequences from *P. ananatis (*4DGK-pdb id); *A. thaliana*(Q07356); *D. salina*(O23915*); B. prasinos* (K8FAQ2);*V. carteri* (D8TP57); *C. sorokiniana*(A082P6TR05); *N. gaditana* (W7TCB4); *O. lucimarinus*(A4S416); *C. merolae* (M1US84). The secondary structure elements of A8J3K3 have been indicated above the alignment and the coloured bar underneath the alignment indicates the domain organisation with the FAD-binding domain (green), the substrate-binding domain (pink), and the non-conserved 'helical' or 'membrane-binding' domain (cyan). Disordered regions in the structure are represented by a purple circle and residues involved in FAD binding as predicted using are indicated by red triangles.

oxido reductase belonging to amino oxidase protein family (PF01593) ranging from the 55-544 amino acids. SUPERFAMILY also predicted that the protein possesses the FAD/NAD (P)-binding domain. The enzyme PDS catalyzes the symmetric introduction of the double bonds required to form the fully desaturated lycopene and contains flavin adenine dinucleotide (FAD) as the sole protein-bound redox-cofactor (Gemmecker et al., 2015). The ProtoNet tool also confirmed that the protein belongs to Cluster 4150575 which includes proteins exhibiting oxido-reductase activity.

Sequence alignment of crystal structure of PDS (PDB Id:4DGK; P21685) enzyme of *P. ananatis* and similar sequences identified through BLAST searches, particularly from various algal species and Arabidopsis allowed identification of regions involved in membrane binding, FAD binding and substrate binding. The alignments results (Fig. 2) indicated that the sequences share high conservation in the aforementioned regions responsible for enzyme catalysis and function as compared to the rest of the regions. The region spanning the FAD binding domain in the A8J3K3 protein of *C. reinhardtii* is occupied by amino acids at positions 66-95, 114-143, 292-382, 410-581. Likewise, the substrate binding (96-113, 144-174, 288-291, 383-409) and membrane binding residues (175-287) in the protein were found to possess high sequence identity. MEME tool identified the signature Glycine rich GXGXXG motif conserved in all the homologous sequences which is commonly present in the NAD(P)H-binding domain and FAD-bindingdomain protein family (Dym and Eisenberg, 2001). The role of the Gly residues in the conserved central GxGxxG is well known understood (Wierenga et al., 1986). The strictly conserved Ghy residues permits a tight turn of the main chain, which is crucial for positioning the second Gly that binds to the FAD while the third Gly allows close packing of the helix with the $\beta$-sheet (Dym and Eisenberg, 2001). Taken together, these results indicate that the hypothetical protein A8J3K3 is a putatively PDS enzyme in the carotenoid biosynthetic pathway (Busch et al.,2002).

### 3.2. Physico-chemical analysis

The ProtParam tool computed various physiochemical properties of the predicted PDS protein of *C. reinhardtii* (A8J3K3), such as its molecular weight, theoretical pI, amino acid composition, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY), as shown in Supplementary Table 2. The ProtParam tool calculated a pI of 8.89 and a molecular weight of 60169.30 for the protein. The instability index refers to the stability of a protein in a test tube where a value above 40 signifies instability whereas value less than 40 signifies a stable protein. The studied putative PDS protein was observed to be stable with Instability Index value as 38.08.The aliphatic index of a protein demonstrates the relative volume occupied by aliphatic side chains and the GRAVY value is defined by the sum of hydropathy values of all amino acids divided by the protein length. The higher Aliphatic Index of value 90.27 shows that the protein is thermostable in nature. Further, the GRAVY value was found to be -0.084 suggesting that the putative PDS to be a hydrophilic protein.

### 3.3. Sub cellular localisation and secondary structure analysis

Sub cellular localisation tools like ChloroP, TargetP and transmembrane prediction tool HMMTOP confirmed that the protein is located in plastid and is membrane bound which in agreement with experimental reports (Yruela et al., 2010; Schaub et al., 2012). The two transmembrane helices spanning 48-65 and 238-257 residues are highlighted in the Fig. 2. SCRATCH protein predictor also confirmed that the protein is insoluble in nature with probability 0. 666463. Signal P tool predicted that the putative PDS has no signal peptide, which shows that the protein is not involved in the secretary pathway. Secondary structure analysis using the tool SOPMA showed that the protein contains 37.10% alpha helix, 16.31 % beta sheet and 40.50% random coil. Further analysis revealed that the FAD-binding domain is composed of nineteen stranded sheets and a twenty-one-helix bundle as reported earlier in a similar study on structural alignment of homologous PDS sequences containing the FAD-binding Rossmann fold proteins (Schaub et al., 2012). The secondary structure elements present in the FAD domain have been indicated above the alignment (Fig. 2). The Sub cellular localisation and Secondary structure results are summarised in Supplementary Table 1. The PSIPRED tool identified the disorder regions which are indicated as blue dots in the sequence alignment Fig.2. CYSPRED predicted the cysteine residues which are capable of forming disulphide bonds and play a pivotal role in the stability of the protein (Supplementary Table 1).

### 3.4. Three-dimensional structure and interaction with FAD ligand

BlastP and SWISS Model server was used for template selection. Based on high sequence similarity, the crystal structure of gamma-carotenoid desaturase derived at 1.97Å (PDB ID: 4REP) and having 23% sequence identity (85% sequence similarity) was used as a template for modeling. Fig. 3 showing sequence alignment between the

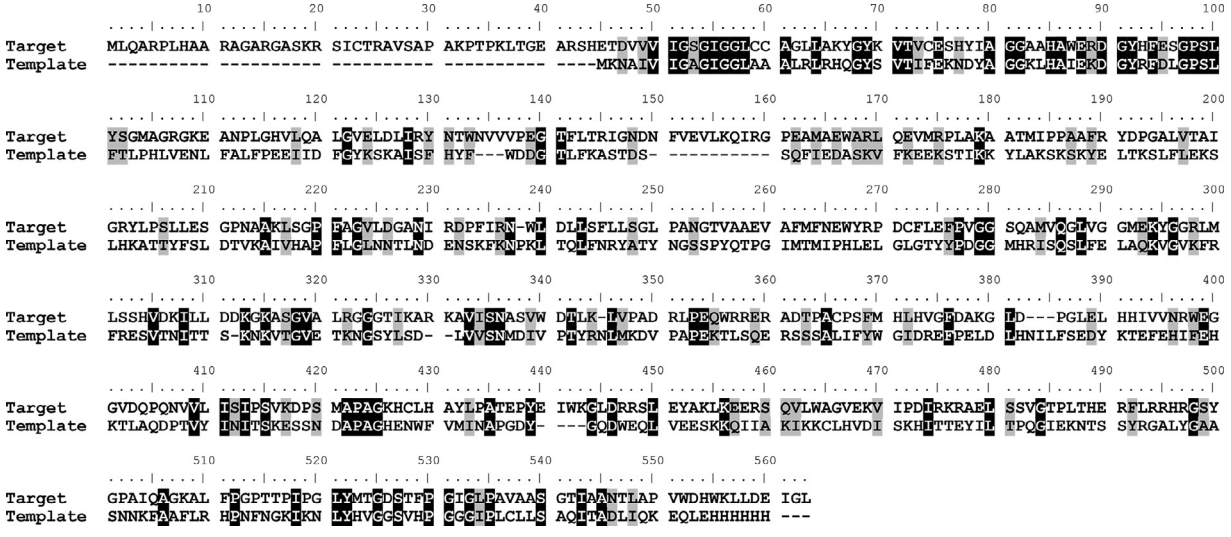


**Fig. 3.** Multiple sequence alignment of the protein sequence of *C. reinhardtii* with the template structure (Pdb Id: 4REP).the identical residues are shaded in black colour while the similar amino acids in grey shade.

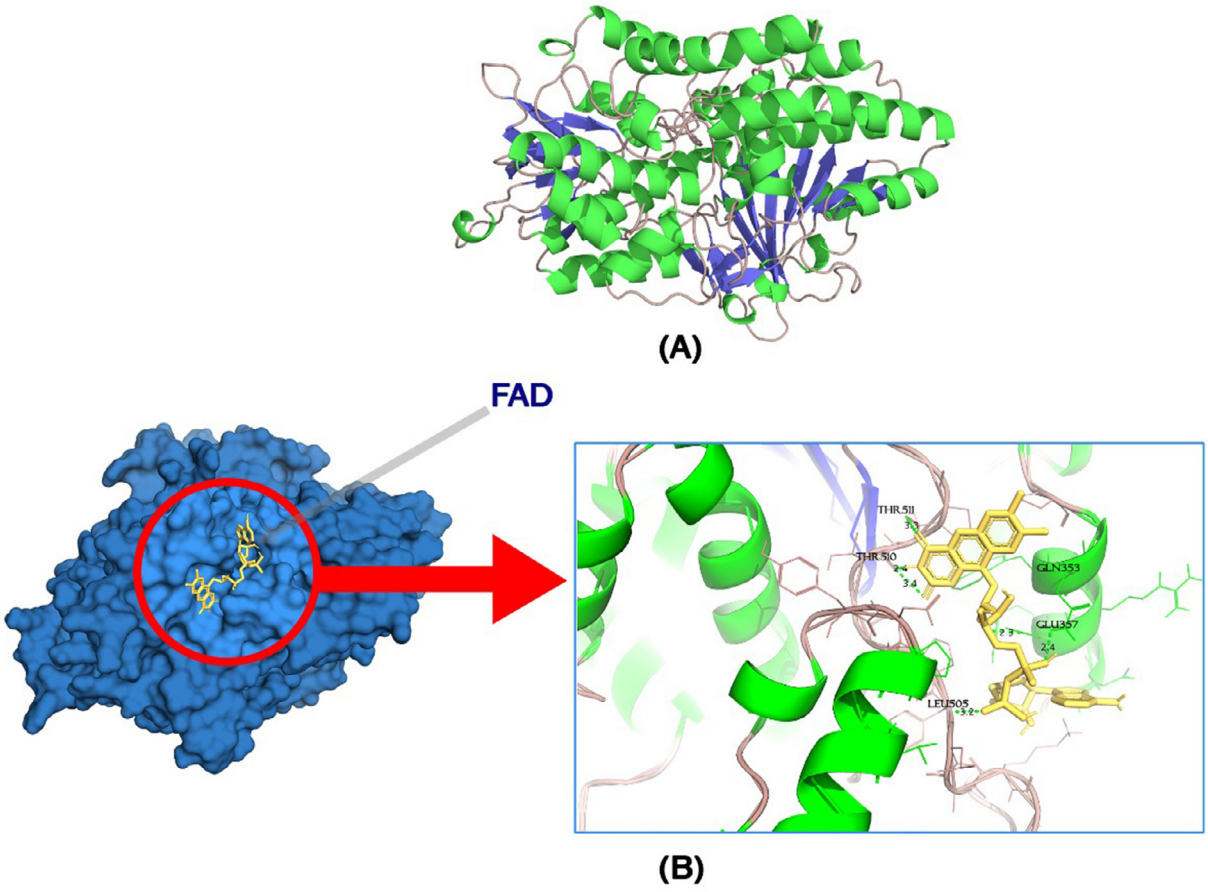


**Fig. 4.** (A) Predicted three-dimensional structural model visualised using PyMol **(B)** Docked complex showing the amino acids involved in FAD cofactor binding. The hydrogen bonds are shown in black lines with bond distances indicated in angstrom.

target and template confirmed high conservancy in both the sequences and thus suitable to be used as template for further analysis. Prior to modeling, the attached ligand was removed from the crystal structure and the apoenzyme was subsequently used as template for further analysis. Subsequently, the three-dimensional structure was generated using ROBETTA server (Fig. 4A) and validated using various structure validation tools like Procheck, Errat, and Verify3D (Supplementary Table3). The Ramachandran plot showed that 87.5%, 11.4%, 0.7% and 0.4% of residues fall in most favoured, additional allowed, generously allowed and disallowed region, respectively. The high

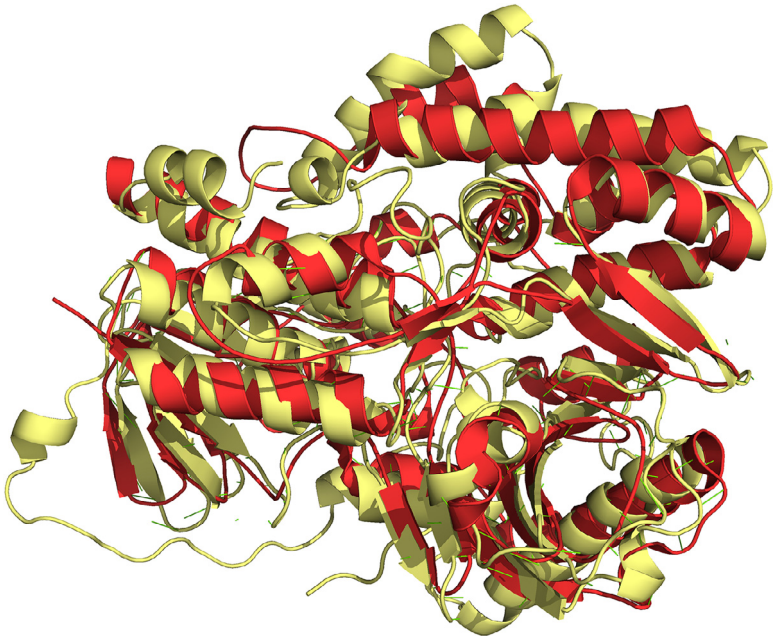

**Fig. 5.** Superimposed three-dimensional structure model of template 4REP (Red) and target protein (Yellow).

scores of ERRAT and VERIFY3D demonstrated the overall accuracy of the modelled structures (Supplementary Fig. 1). Further, the template structure was superimposed on to the modeled structure of *C. reinhardtii* phytoene desaturase enzyme using the PyMOL tool. The root mean square deviation (RMSD) between the main chain atoms of the model and template was observed to be 2.360 Angstrom (Fig. 5) which confirmed the reliability and accuracy of the modeled structure for further analysis.

The protein was found to contain 25 alpha helices and 19 beta sheets. The catalytic pocket as predicted by CastP showed a hydrophobic pocket of area 435.425SA and volume 200.834SA.The residues present in the active site includes $HIS^{8}$, $ARG^{11}$, $ALA^{12}$, $ALA^{26}$, $VAL^{27}$, $ILE^{51}$, $GLY^{52}$, $SER^{53}$, $GLY^{54}$, $ILE^{55}$, $CYS^{74}$, $GLU^{75}$, $SER^{76}$, $HIS^{77}$, $GLY^{81}$, $ALA^{83}$, $ALA^{84}$, $PRO^{98}$, $SER^{99}$, $LEU^{100}$, $TRP^{266}$, $SER^{302}$, $HIS^{303}$, $VAL^{304}$, $ASP^{305}$, $ASN^{335}$, $ALA^{336}$, $ASP^{340}$, $LYS^{343}$, $LEU^{344}$, $GLN^{353}$, $GLU^{357}$, $LEU^{488}$, $ARG^{489}$, $ARG^{490}$, $SER^{494}$, $TYR^{495}$, $LEU^{505}$, $THR^{510}$, $THR^{511}$, $GLY^{520}$, $ASP^{521}$, $ILE^{527}$, $GLY^{528}$, $LEU^{529}$, $VAL^{532}$. Further, in order to check the binding interaction of FAD ligand in the predicted protein model, molecular docking was performed employing Autodock Vina. After docking procedures, 9 docking poses were generated with predicted binding affinity values in the range of -8.1 and -6.9 Kcal/mol with different orientations and configurations of ligands. The best dock pose with highest binding energy was analyzed further by Pymol. The docking study also showed that the FAD ligand binds tightly in the FAD binding domain of the protein (Fig.4 B). The residues involved in the five hydrogen bond interactions are THR 511, THR 510, GLU 357, LEU 505, and GLN 353 with the distance 3.3 Å, 2.4 Å, 2.3 Å, 3.2 Å, and 2.4 Å, respectively which lies in $\alpha$14 and between $\alpha$20 and $\beta$19. Furthermore, the FAD ligand molecule docked into the predicted active site pocket of the protein as predicted by CastP tool and also aligned well with the conformation of the crystal structure of the template which corroborates the optimum orientation of the developed ligand-receptor docked structure.

## 4. Conclusion

The in-depth sequence-structure analysis revealed that the hypothetical protein A8J3K3 from *C. reinhardtii* has a major role in the carotenoid biosynthetic pathway as PDS enzyme. Information about the biophysical properties, sub cellular localization, transmembrane topology, conserved motif architecture, FAD binding domain, and substrate binding region will certainly facilitate further experimental analysis using this enzyme. Further the three dimensional model generated is the first attempt to delineate the structural conformation and active site region of PDS from microalgae. The employed *in silico* approach for functional annotation of hypothetical proteins in this work can be utilised as framework for characterizing other vital enzymes of algal carotenoid pathway.

## Funding statement

The present study is an inhouse exploratory research work, authors received no funding support from an external source.

## Declaration of Competing Interest

All authors declared that there are no conflicts of interest.

## Acknowledgments

We acknowledge the help rendered by Mr. Krishn Kumar Verma, Scientific Visualiser of KIIT-Technology Business Incubator, Bhubaneswar for schematic representation of all figures provided in the manuscript.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.sajb.2021.04.014.

## References

Abd Rashid, S.N.A., Malik, S.A., Embi, K., Ropi, N.A.M., Yaakob, H., Cheng, K.K., Leong, H.Y., 2020. Carotenoids and antioxidant activity in virgin palm oil (VPO) produced from palm mesocarp with low heat aqueous-enzyme extraction techniques. Mater. Today: Proc.. https://doi.org/10.1016/j.matpr.2020.10.616.

Busch, M., Seuter, A., Hain, R., 2002. Functional analysis of the early steps of carotenoid biosynthesis in tobacco. Plant Physiol. 128, 439–453. https://doi.org/10.1104/pp.010573.

Cordero, B.F., Couso, I., León, R., Rodríguez, H., Vargas, M.Á., 2011. Enhancement of carotenoids biosynthesis in Chlamydomonas reinhardtii by nuclear transformation using a phytoene synthase gene isolated from Chlorella zofingiensis. Appl. Microbiol. Biotechnol. 91, 341–351. https://doi.org/10.1007/s00253-011-3262-y.

Das, A., Yoon, S.H., Lee, S.H., Kim, J.Y., Oh, D.K., Kim, S.W., 2007. An update on microbial carotenoid production: application of recent metabolic engineering tools. Appl. Microbiol. Biotechnol. 77, 505. https://doi.org/10.1007/s00253-007-1206-3.

Dym, O., Eisenberg, D., 2001. Sequence-structure analysis of FAD-containing proteins. Protein Sci. 10, 1712–1728. https://doi.org/10.1110/ps.12801.

Fernández-García, E., Carvajal-Lérida, I., Jarén-Galán, M., Garrido-Fernández, J., Pérez-Gálvez, A., Hornero-Méndez, D., 2012. Carotenoids bioavailability from foods: From plant pigments to efficient biological activities. Food Res. Int. 46, 438–450. https://doi.org/10.1016/j.foodres.2011.06.007.

Fraser, P.D., Enfissi, E.M., Bramley, P.M., 2009. Genetic engineering of carotenoid formation in tomato fruit and the potential application of systems and synthetic biology approaches. Arch. Biochem. Biophys. 483, 196–204. https://doi.org/10.1016/j.abb.2008.10.009.

Gemmecker, S., Schaub, P., Koschmieder, J., Brausemann, A., Drepper, F., Rodriguez-Franco, M., Beyer, P., 2015. Phytoene desaturase from Oryza sativa: oligomeric assembly, membrane association and preliminary 3D-analysis. PLoS One 10, e0131717. https://doi.org/10.1371/journal.pone.0131717.

Gimpel, J.A., Henríquez, V., Mayfield, S.P., 2015. In metabolic engineering of eukaryotic microalgae: potential and challenges come with great diversity. Front. Microbiol. 6, 1376. https://doi.org/10.3389/fmicb.2015.01376.

Hall, T., Biosciences, I., Carlsbad, C., 2011. BioEdit: an important software for molecular biology. GERF Bull. Biosci. 2, 60–61.

Koschmieder, J., Fehling-Kaschek, M., Schaub, P., Ghisla, S., Brausemann, A., Timmer, J., Beyer, P., 2017. Plant-type phytoene desaturase: functional evaluation of structural implications. PLoS One 12, e0187628. https://doi.org/10.1371/journal.pone.0187628.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283–291. https://doi.org/10.1107/S0021889892009944.

Lengths, M.C., Angles, M.C., 2018. Limitations of structure evaluation tools errat. Quick Guideline Comput. Drug Des. 16, 75.

Leu, S., Boussiba, S., 2014. Advances in the production of high-value products by microalgae. Ind. Biotechnol. 10, 169–183. https://doi.org/10.1089/ind.2013.0039.

Liu, J., Gerken, H., Huang, J., Chen, F., 2013. Engineering of an endogenous phytoene desaturase gene as a dominant selectable marker for Chlamydomonas reinhardtii transformation and enhanced biosynthesis of carotenoids. Process Biochem. 48, 788–795. https://doi.org/10.1016/j.procbio.2013.04.020.

Mann, V., Pecker, I., Hirschberg, J., 1994. Cloning and characterization of the gene for phytoene desaturase (Pds) from tomato (Lycopersicon esculentum). Plant Mol. Biol. 24, 429–434. https://doi.org/10.1007/BF00024111.

Misra, N., Panda, P.K., Parida, B.K., 2014. Genome-wide identification and evolutionary analysis of algal LPAT genes involved in TAG biosynthesis using bioinformatic approaches. Mol. Biol. Rep. 41, 8319–8332. https://doi.org/10.1007/s11033-014-3733-1.

Misra, N., Panda, P.K., 2013. In search of actionable targets for agrigenomics and microalgal biofuel production: sequence-structural diversity studies on algal and higher plants with a focus on GPAT protein. OMICS 17, 173–186. https://doi.org/10.1089/omi.2012.0094.

Misra, N., Panda, P.K., Parida, B.K., 2013. Agrigenomics for microalgal biofuel production: an overview of various bioinformatics resources and recent studies to link OMICS to bioenergy and bioeconomy. Omics 17, 537–549. https://doi.org/10.1089/omi.2013.0025.

Misra, N., Panda, P.K., Parida, B.K., Mishra, B.K., 2012. Phylogenomic study of lipid genes involved in microalgal biofuel production—candidate gene mining and metabolic pathway analyses. Evol. Bioinformatics 8,. https://doi.org/10.4137/EBO.S10159 EBO-S10159.

Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., Olson, A.J., 2009. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J. Comput. Chem. 30, 2785–2791. https://doi.org/10.1002/jcc.21256.

Nakano, T., Wiegertjes, G., 2020. Properties of carotenoids in fish fitness: a review. Marine Drugs 18, 568. https://doi.org/10.3390/md18110568.

Noviendri, D., Hasrini, R.F., Octavianti, F., 2011. Carotenoids: sources, medicinal properties and their application in food and nutraceutical industry. J. Med. Plants Res. 5, 7119–7131. https://doi.org/10.5897/JMPRx11.011.

Novoveská, L., Ross, M.E., Stanley, M.S., Pradelles, R., Wasiolek, V., Sassi, J.F., 2019. Microalgal carotenoids: a review of production, current markets, regulations, and future direction. Marine Drugs 17, 640. https://doi.org/10.3390/md17110640.

O'Boyle, N.M., Banck, M., James, C.A., Morley, C., Vandermeersch, T., Hutchison, G.R., 2011. Open Babel: an open chemical toolbox. J. Cheminformatics 3, 33. https://doi.org/10.1186/1758-2946-3-33.

Rammuni, M.N., Ariyadasa, T.U., Nimarshana, P.H.V., Attalage, R.A., 2019. Comparative assessment on the extraction of carotenoids from microalgal sources: Astaxanthin from H. pluvialis and $\beta$-carotene from D. salina. Food Chem. 277, 128–134. https://doi.org/10.1016/j.foodchem.2018.10.066.

Rosenberg, J.N., Oyler, G.A., Wilkinson, L., Betenbaugh, M.J., 2008. A green light for engineered algae: redirecting metabolism to fuel a biotechnology revolution. Curr. Opin. Biotechnol. 19, 430–436. https://doi.org/10.1016/j.copbio.2008.07.008.

Sahoo, S., Mahapatra, S.R., Das, N., Parida, B.K., Rath, S., Misra, N., Suar, M., 2020. Functional elucidation of hypothetical proteins associated with lipid accumulation: prioritizing genetic engineering targets for improved algal biofuel production. Algal Res. 47, 101887. https://doi.org/10.1016/j.algal.2020.101887.

Saini, R.K., Keum, Y.S., 2017. Progress in microbial carotenoids production. Indian J. Microbiol. 57, 129–130. https://doi.org/10.1007/s12088-016-0637-x.

Sathasivam, R., Ki, J.S., 2018. A review of the biological activities of microalgal carotenoids and their potential use in healthcare and cosmetic industries. Marine Drugs 16, 26. https://doi.org/10.3390/md16010026.

Schaub, P., Yu, Q., Gemmecker, S., Poussin-Courmontagne, P., Mailliot, J., McEwen, A.G., Beyer, P., 2012. On the structure and function of the phytoene desaturase CRTI from Pantoea ananatis, a membrane-peripheral and FAD-dependent oxidase/isomerase. PLoS One 7, e39550. https://doi.org/10.1371/journal.pone.0039550.

Trott, O., Olson, A.J., 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 31, 455–461. https://doi.org/10.1002/jcc.21334.

Varela, J.C., Pereira, H., Vila, M., León, R., 2015. Production of carotenoids by microalgae: achievements and challenges. Photosynth. Res. 125, 423–436. https://doi.org/10.1007/s11120-015-0149-2.

Vílchez, C., Forjan, E., Cuaresma, M., Bédmar, F., Garbayo, I., Vega, J.M., 2011. Marine carotenoids: biological functions and commercial applications. Marine Drugs 9, 319–333. https://doi.org/10.3390/md9030319.

Wierenga, R.K., Terpstra, P., Hol, W.G., 1986. Prediction of the occurrence of the ADP-binding $\beta\alpha\beta$-fold in proteins, using an amino acid sequence fingerprint. J. Mol. Biol. 187, 101–107. https://doi.org/10.1016/0022-2836(86)90409-2.

Yuan, S., Chan, H.S., Filipek, S., Vogel, H., 2016. PyMOL and Inkscape bridge the data and the data visualization. Structure 24, 2041–2042.

Yruela, I., Arilla-Luna, S., Medina, M., Contreras-Moreira, B., 2010. Evolutionary divergence of chloroplast FAD synthetase proteins. BMC Evol. Biol. 10, 311. https://doi.org/10.1186/1471-2148-10-311.

Zhu, Y.H., Jiang, J.G., Yan, Y., Chen, X.W., 2005. Isolation and characterization of phytoene desaturase cDNA involved in the $\beta$-carotene biosynthetic pathway in Dunaliella salina. J. Agric. Food Chem. 53, 5593–5597. https://doi.org/10.1021/jf0506838.