In silico Structural Determination of Fatty acid Hydrolysis and Synthesizing Enzymes for Biofuel Production in Microalgae

S. Sathya[1], S. Meenakshi [2] and S. Srisudha[3]

Research Scholar, Research centre of Botany and Microbiology, Lady Doak College, Madurai – 625002, Tamilnadu, India[1]

Research Scholar, Research centre of Botany and Microbiology, Lady Doak College, Madurai – 625002, Tamilnadu, India[2]

Associate Professor, Research centre of Botany and Microbiology, Lady Doak College, Madurai – 625002, Tamilnadu, India[3]

Abstract: Twenty five sequences of algae were retrieved from NCBI for lipase, fatty acid synthase and glycerol 3-phosphate dehydrogenase. Physicochemical parameters were determined using pepstats, ProtParam, ProtScale and Network protein sequence analysis. Conserved regions were identified by multiple sequence alignment using Clustal W. The primary structure analysis of three enzymes indicates that the percentage of Methionine residues is approximately 4% among the selected microalgae. Secondary structure analysis revealed that the proteins are largely alpha helical; this was supported by the homology modeling of three enzymes of *Chlamydomonas reinhardtii* (L2), *Osterococcus lucimarnus* (F7) and *Osterococcus taurii* (G7). Structural validation tools viz., Ramachandran Plot, VERIFY 3D and ERRAT indicates the overall accuracy of the developed models. More active sites are present in *Osterococcus lucimarnus* (F7) than *Chlamydomonas reinhardtii* (L2) and *Osterococcus taurii* (G7). The results revealed that the structural models developed from this study will give useful information for biofuel production.

Key words: Biofuel, Lipase, Fatty acid synthase, Glycerol 3 phosphate dehydrogenase.

I. INTRODUCTION

Biofuels can be solids, liquids or gasses so long as they are derived directly from biological sources. Here microalgae as a biological sources for the synthesis of Biofuels production. Liquid and gaseous biofuels generally require more refining, and include bioethanol, biodiesel, and engine-combustible hydrocarbons as well as methane from anaerobic digestion. To survive in ecosystems, algae produce a myriad of lipids. These include structural lipids for cellular membranes, as well as lipids for nutrient storage [22, 23]. Optimizing microalgal biofuel production using metabolic engineering tools requires an in depth understanding of the structure-function relationship of genes involved in lipid biosynthetic pathway. The enhancement of lipid production in microalgal cells under controlled stress conditions and engineering metabolic pathways are promising strategies to obtain large amounts of standard biofuel for industry [1]. Storage lipids differ from both structural and signaling lipids in that they are mainly composed of glycerol esters of fatty acids, also known as triacylglycerol (TAG). These lipids are generally stored in a specialized compartment found in most oleaginous plant cells [2]. Lipases widely exist in microbe and plant species, but their structure is poorly characterized. Triglycerides are hydrolyzed by lipases to glycerol and fatty acids and provide energy. Generally, algae use starch as their primary storage vehicle, however, in some strains large quantities of TAG accumulate under specific growing conditions [26]. Fatty acid synthase (FAS) complex is a major player in de novo fatty acid synthesis. Triglycerides are stored in intracellular structures called oil bodies [3]. Three-dimensional structure of proteins gives valuable insight into the molecular organization, function and also for effective drug designing experiments. In the absence of an experimentally determined crystal structure, homology modeling could provide a rational opportunity to obtain a reasonable 3D model [4]. Glycerol-3-phosphate dehydrogenase (GPDH) is an enzyme that catalyzes the reversible redox conversion of dihydroxyacetone phosphate to glycerol-3-phosphate. GPDH serves as a major link between carbohydrate metabolism and lipid metabolism. The structural insight obtained from *in silico* study may

provide useful clues for further advanced biotechnological studies of strategic site-specific genetic and metabolic engineering of microalgae for enhanced biofuel production [5]. The present study will improve our understanding of the global lipid metabolic pathway and contribute to the engineering of regulatory networks of micro algal strains for higher production of biofuel.

II. METHODOLOGY

Protein Sequences

Twenty five protein sequences were retrieved from the National Center for Biotechnology Information (NCBI) for each enzyme such as such as Lipase, fatty acid synthase and glycerol-3-phosphate dehydrogenase [http://www.ncbi.nlm.nih.gov/]. The sequences were converted to FASTA format using the ReadSeQ sequence conversion server [6]. The selected protein sequences and their accession numbers are represented in Table I.

Sequence Analysis
Sequence Alignments

Multiple sequence alignments were performed with ClustalW [7]. Phylogenetic analysis of protein sequences of all the three enzymes, were generated using the alignment obtained with ClustalW. Each enzyme contains 25 protein sequences.

Structural Analysis
Physico-Chemical Characterization and Secondary Structure Prediction

The amino acid composition of lipase, fatty acid synthase and glycerol-3-phosphate dehydrogenase sequences was computed using the pepstats analysis tool [8]. The physico-chemical parameters such as the molecular weight, isoelectric point (pI), extinction coefficient, half-life, aliphatic index, amino acid property, instability index and Grand Average Hydropathicity (GRAVY) were calculated using the ProtParam tool of the Expasy proteomics server. Secondary structure consensus prediction was performed using the Network Protein Sequence Analysis (NPSA) server and the most predicted structure is presented as double prediction method (DPM) [9]. Protscale was used for computing and representation was done in the form of a two-dimensional plot of the profile produced by any amino acid scale on a selected protein. The Hydrophobicity and Hydrophilicity analysis of an amino acid sequence of protein sequences was fulfilled with Protscale program [10].

Structural Assessment of the Developed Model

The protein structure of all the three enzyme homology model building program was visualized using SWISS-MODEL with energy minimization parameters [11]. The refine models and the standalone version of Swiss PDB viewer [12] was used to generate publishable images of the enzyme models. The reliability of the 3-D models of microalgal enzymes were constructed using various structure validation tools viz., PROCHECK [13], WHATCHECK [14], Verify 3D [15], ERRAT [16] and Ramachandran plot [17,18]. Structural and packing architecture of the modeled structure were computed by calculating the total volume, accessible surface area, backbone and side chain dihedral angle, hydrogen bonding partners and hydrogen bond energy using VADAR 1.4 programme [19].

Active Site Prediction

Identification of substrate accessible pockets and functional residues present in the active sites of the selected protein sequences were determined with Computed Atlas of Surface Topography of Proteins (CASTp) programme [20].

III. RESULTS AND DISCUSSION

Twenty five protein sequences were retrieved from the National Center for Biotechnology Information (NCBI) for each enzyme such as Lipase, fatty acid synthase and glycerol-3-phosphate dehydrogenase (Table I). On the basis of primary structure analysis, the protein sequences were selected for each enzyme. The primary structure analysis of 75 protein sequences for three enzymes indicated, low percentage of Methionine residues approximately 4% among all microalgae. Lower methionine content increased the synthesis of fatty acids in microalgae. The most abundant amino acid was leucine (10-11 %.) in all the three enzymes. Production of unusual fatty acids in transgenic hosts can induce antagonistic pathways reducing the effects of genetic manipulation, which must be addressed to maximize production efficiency [26]. Leucine residues were found as dimers in cyanobacterial PEPC's and it was one of the most significant residues in all the cyanobacterial strains [21]. The maximum molecular weight was observed in lipase (1, 57,906gmol^{-1}), fatty acid synthase (8, 48,291.1gmol^{-1}) and glycerol-3-phosphate dehydrogenase (77,125gmol^{-1}) among the all selected protein sequences. The lower cost of TAG fuels on the transesterfication and purification of FAMEs greatly enhances the market potential of such biodiesels. Biodiesel production, and may be applicable through genetic manipulation of lipid biosynthetic pathways [26]. Physico-Chemical Characterization showed an outline of genetic manipulations that

have been performed in microalgae and the resulting changes in fatty acid composition and content. These tests must prove that strains with high lipid content and economically viable compared to elite, high-yield commercial varieties [26]. The computed isoelectric point (pI) was below 7 indicating that the proteins can precipitate in acidic buffers. pI value was high in lipase (10.19), fatty acid synthase (8.57) and glycerol-3 phosphate dehydrogenase (9.31) and this pI value indicated that these three enzymes were actively involved in hydrolysis and lipid biosynthesis. Instability index (Ii) of fatty acid synthase (26.32) was less than lipase (34.14) and glycerol-3 phosphate dehydrogenase (39.54). Instability index (Ii) greater than 40, indicated that, the protein may be unstable. Aliphatic index (Ai) gave the measure of relative volume occupied by alanine, valine, isoleucine and leucine and it ranged from 92.85 to 104.85 for all the three enzymes. The relative high Ai values indicated that, microalgae may be stable over a large range of temperatures. The GRAVY values for all the three enzymes ranged from -0.0169 to -0.471 and it indicated that the proteins can interact favourably with water. The GRAVY index characterizes the solubility of the protein. The lipid biosynthetic

Species	Accession no	Species	Accession no	Species	Accession no
Micromonas sp. RCC299(L1)	ACO70423	*Ostreococcus lucimarinus CCE9901* (F1)	ACO66304	*Dunaliella salina* (G1)	AAX56341
Chlamydomonas reinhardtii(L2)	EDP04732	*Ostreococcus taurii* (F2)	EEC50984	*Thalassiosira pseudonana CCMP1335*(G2)	EED95644
Ostreococcus taurii(L3)	CAL56469	*Micromonas sp. RCC299* (F3)	XP_002180902	*Chlamydomonas reinhardtii*(G3)	EDP05322
Micromonas pusilla CCMP1545(L4)	EEH54218	*Micromonas sp. RCC299* (F4)	ABO99784	*Coccomyxa subellipsoidea C-169*(G4)	EIE25294
Ostreococcus taurii(L5)	XP_003079965	*Coccomyxa subellipsoidea C-169* (F5)	EIE24233	*Dunaliella viridis*(G5)	ACD84643
Chlorella variabilis(L6)	EFN51998	*Chlamydomonas reinhardtii* (F6)	EDP01579	*Dunaliella viridis*(G6)	ACD84644
Thalassiosira oceanic(L7)	EJK68811	*Micromonas sp. RCC299* (F7)	ACO63966	*Ostreococcus tauri*(G7)	CAL51808
Isochrysis galbana(L8)	CAO78537	*Micromonas pusilla CCMP1545* (F8)	XP_002185095	*Volvox carteri f. nagariensis*(G8)	EFJ49807
Thalassiosira pseudonana CCMP1335(L9)	EED89966	*Thalassiosira pseudonana CCMP1335* (F9)	EED88129	*Chlamydomonas reinhardtii*(G9)	XP_001693544
Nannochloropsis gaditana CCMP526(L10)	AFJ69551	*Phaeodactylum tricornutum* (F10)	XP_002183028	*Chlorella variabilis*(G10)	EFN58098
Ostreococcus taurii(L11)	XP_003079395	*Ostreococcus taurii* (F11)	CAL57604	*Chlorella variabilis*(G11)	EFN55198
Ostreococcus taurii(L12)	CBN77732	*Ostreococcus taurii* (F12)	CAL57571	*Phaeodactylum tricornutum CCAP 1055/1*(G12)	EEC42726
Ostreococcus taurii(L13)	CAL50026	*Ostreococcus taurii* (F13)	CAL56184	*Phaeodactylum tricornutum CCAP 1055/1*(G13)	EEC51773
Microcystis aeruginosa NIES-843(L14)	BAG02443	*Micromonas sp. RCC99* (F14)	ACO65933	*Phaeodactylum tricornutum CCAP 1055/1*(G14)	EEC42727
Micromonas pusilla CCMP1545(L15)	XP_003061588	*Micromonas pusilla CCMP1545* (F15)	EEH59313	*Dunaliella salina*(G15)	ABI94007
Micromonas sp. RCC299(L16)	XP_002509165	*Micromonas pusilla CCMP1545* (F16)	EEH55336	*Thalassiosira oceanica*(G16)	EJK77600
Micromonas sp. RCC299(L17)	ACO64108	*Phaeodactylum tricornutum CCAP 1055/1* (F17)	EEC45368	*Thalassiosira oceanica*(G17)	EJK64545
Chlamydomonas	XP_0017023	*Micromonas pusilla*	EEH51062	*Thalassiosira*	EED96360

reinhardtii(L18)	79	*CCMP1545* (F18)		*pseudonana CCMP1335*(G18)	
Isochrysis galbana(L19)	CCB63200	*Phaeodactylum tricornutum CCAP 1055/1* (F19)	NP_053806	*Micromonas sp. RCC299*(G19)	ACO66397
Thalassiosira pseudonana CCMP1335(L20)	XP_002292770	*Thalassiosira oceanica* (F20)	EJK64649	*Micromonas pusilla CCMP1545*(G20)	EEH55245
Ostreococcus taurii(L21)	XP_003082612	*Thalassiosira oceanica* (F21)	EJK58743	*Chlamydomonas reinhardtii*(G21)	EDP06377
Ostreococcus taurii(L22)	XP_003074695	*Chlorella variabilis* (F22)	EFN53668	*Chlamydomonas reinhardtii*(G22)	EDP09536
Chlorella variabilis(L23)	EFN54122	*Chlorella variabilis* (F23)	EFN55645	*Chlamydomonas reinhardtii*(G23)	EDP09534
Chlamydomonas reinhardtii(L24)	EDP08159	*Chlamydomonas reinhardtii* (F24)	EDO96631	*Ostreococcus lucimarinus CCE9901*(G24)	ABO93860
Chlorella variabilis(L25)	EFN59908	*Thalassiosira oceanica* (F25)	EJK72067	*Ostreococcus taurii*(G25)	CAL50284

Table I: SELECTED PROTEIN SEQUENCES FOR LIPASE, FATTY ACID SYNTHASE AND GLYCEROL-3-PHOSPHATE DEHYDROGENASE RETRIEVED FROM NCBI

proteins showed large negative values. This showed that, these proteins are relatively more hydrophobic when compared to proteins with less negative values [1].

The results of the secondary structure consensus highlighted that, 75 protein sequences were largely alpha helical with less than16% beta structure. Maximum alpha helix observed for the three enzymes were 89.20%, 79.67% and 82.13% respectively. These results revealed that, random coil was predominant, followed by alpha helices and extended strands in the majority of sequences. The fuel properties of biodiesel are closely related to its fatty acid composition. Altering the fatty acid profile with the change of carbon chain length and the number of double bonds, could lead to a better-quality and production of inexpensive biodiesel [26]. The Hydrophobicity exhibited the maximum score value and hydrophilicity showed minimum score in Protscale server. Maximum score of lipase (3.367) at the position of 575, Fatty acid synthase (3.689) at the position of 4735 and 4736 and glycerol-3-phosphate dehydrogenase (3.300) at 72th position were observed. The acetyl TAG has a lower viscosity than common TAGs, and the potential to be used directly as biodiesel [27]. This specific acetyl DAGAT has been isolated from *Euonymus alatus*, and data on the oil properties of transgenic plants are much anticipated [28].

Sequence Alignment

The sequence alignment indicated that fatty acid synthase and glycerol-3-phosphate dehydrogenase shared the most identity with each other. Maximum similarity percentage (67%) was observed in fatty acid synthase and glycerol-3-phosphate dehydrogenase. A more promising method of genetic engineering has been successfully established in the chloroplast of *Chlamydomonas reinhardtii* [30]. Lipase showed only 35% similarity. Phylogenetic tree for lipase showed, 97-100% of similarity in relation to cyanobacteria than diatoms. In fatty acid synthase, *Osterococcus taurii* – G7 (67%) and *Chlamydomonas reinhardtii* – L2 (69%) exhibited a low percentage of similarity in comparison to the other algal strains. In glycerol-3-phosphate dehydrogenase, highly similar sequences were found in all the algal strains except *Phaeodactylum* sp (67%) and *Chlorella* sp (69%). With this primary sequence data and the functional characterization of algal genes determined that the key regulators of algal lipid biosynthesis *in silico*. First gene expression profile of *C. reinhardtii* under hydrogen producing conditions, which was recently reported [29].

Structural Assessment of the Developed Model

Chlamydomonas reinhardtii (L2) in lipase (92.5%), *Osterococcus lucimarinus* (F7) in fatty acid synthase (94.5%) and *Osterococcus taurii* (G7) in glycerol-3-phosphate dehydrogenase (95.7%) showed above 90% of favoured region (Fig. 1) in the Ramachandran plot. This suggested that the developed model was good. Several transformation techniques have been developed to genetically engineer *C. reinhardtii* to express recombinant proteins from both the chloroplast and nuclear genomes [26].
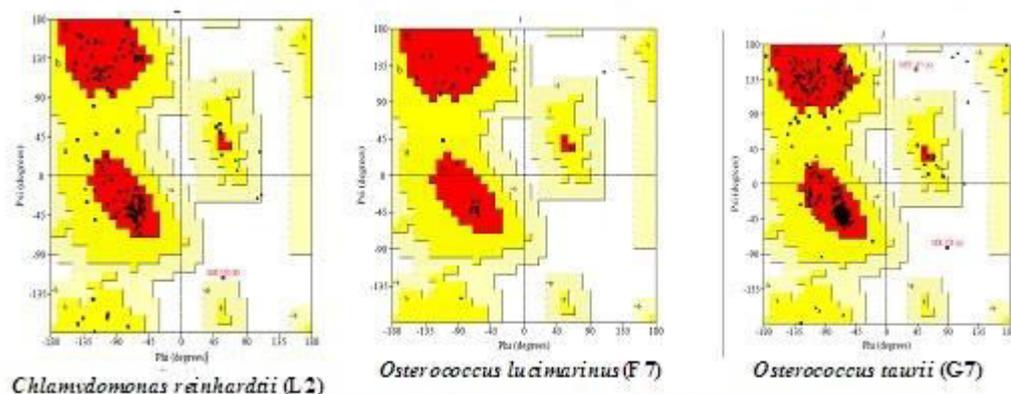

Fig.1: RAMACHANDRAN PLOT CALCULATION FOR 3-D MODEL USING PROCHECK.

In verify 3D model, the score above 0.2 indicated that the buit model was good. In ERRAT, overall quality factor above 50% indicated that the constructed model was good. The maximum ERRAT values were observed in *Chlamydomonas reinhardtii* (L2) for lipase (68.5%), *Osterococcus lucimarinus* (F7) for fatty acid synthase (100%) and *Osterococcus taurii* (G7) for glycerol-3-phosphate dehydrogenase (81.9%). Based on the physico-chemical parameters and validation tools, best sequences were selected from all the three enzymes. *Chlamydomonas* was also the first alga to be genetically transformed and a draft sequence of the whole genome has recently been determined [24]. Although it does not typically accumulate lipids under ideal conditions, metabolic engineering can be used to transform this alga into an oleaginous factory [25]. The selected microalgal strains were L2 (*Chlamydomonas reinhardtii*), F7 (*Osterococcus lucimarinus*) and G7 (*Osterococcus taurii*). The homology models of the selected algal strains were predicted using SWISS-MODEL (Fig. 2) and the predicted model was visualized by Swiss PDB viewer. The Verify 3D method assesses the protein structures using three-dimensional profiles. This program analyzed the compatibility of an atomic model (3D) with its own amino acid sequence (1D). The scores ranged from -1 (bad score) to +1 (good score) [12] (Table II). In addition, the good structural and packing architecture as predicted through VADAR programme, conclusively confirmed the reliability of the developed structure for further *in silico* studies (Table III).
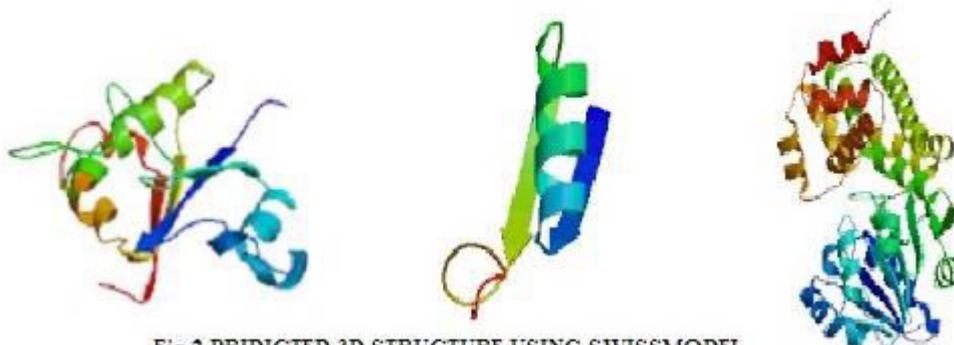
Fig.2 PRIDICTED 3D STRUCTURE USING SWISSMODEL

Microalgal strains	MHBD (Å)	MHBE	MHPh	MHPs	MCG+	MCG-	TASA (Å)	MRV (Å)	TV (Å)
Chlamydomonas reinhardtii (L2)	2.2	-1.9	-67.6	-37.7	-63.7	65.7	8200.4	121.4	16472.3
Osterococcus lucimarinus (F7)	2.2	-1.8	-66.4	-47.4	-67.0	60.8	3229.0	111.8	4471.5
Osterococcus taurii (G7)	2.2	-1.7	-67.4	-38.6	-64.7	64.5	14680.0	130.9	43053.7

Table. II: PROVE PLOT OF ATOMIC Z-SCORE

Microalgal Strains	Z-score mean	Z-score stddev	Z-score RMS	scored atoms	% outliers	Verify 3D Plot
Chlamydomonas reinhardtii(L2)	0.269	1.454	1.477	403	4.500	0.47
Osterococcus lucimarinus(F7)	-0.048	1.194	1.184	54	0.000	0.34
Osterococcus taurii(G7)	0.231	1.588	1.603	1019	7.200	0.34

Table. III: STRUCTURAL CHARACTERISTICS AND FEATURES OF THE MODEL AS PREDICTED USING VADAR 1.4 PROGRAMME

MHBD- Mean hydrogen bond distance; **MHBE-** Mean hydrogen bond energy; **MHPh-** Mean Helix Phi; **MHPs-** Mean Helix Psi; **MCG+ -** Mean Chi Gauche+; **MCG- -** Mean Chi Gauche-; **TASA-** Total accessible surface area; **MRV-** Mean residue volume; **TV-** Total volume (packing)

Active Site Prediction

The overall secondary structural comparison between the modeled active site regions of the selected three protein sequences was predicted by CASTp server. In our computational model the largest cavity with 71 amino acid residues, 842.5 Å2 area and 1361.9 volume Å3 was presumed to be the active site region of the *Ostreococcus taurii* (G7) (Fig.3). Expression levels vary greatly depending on a number of factors including auto-attenuation of exogenous sequences, codon usage bias, GC content, and proteasome mediated degradation [31].
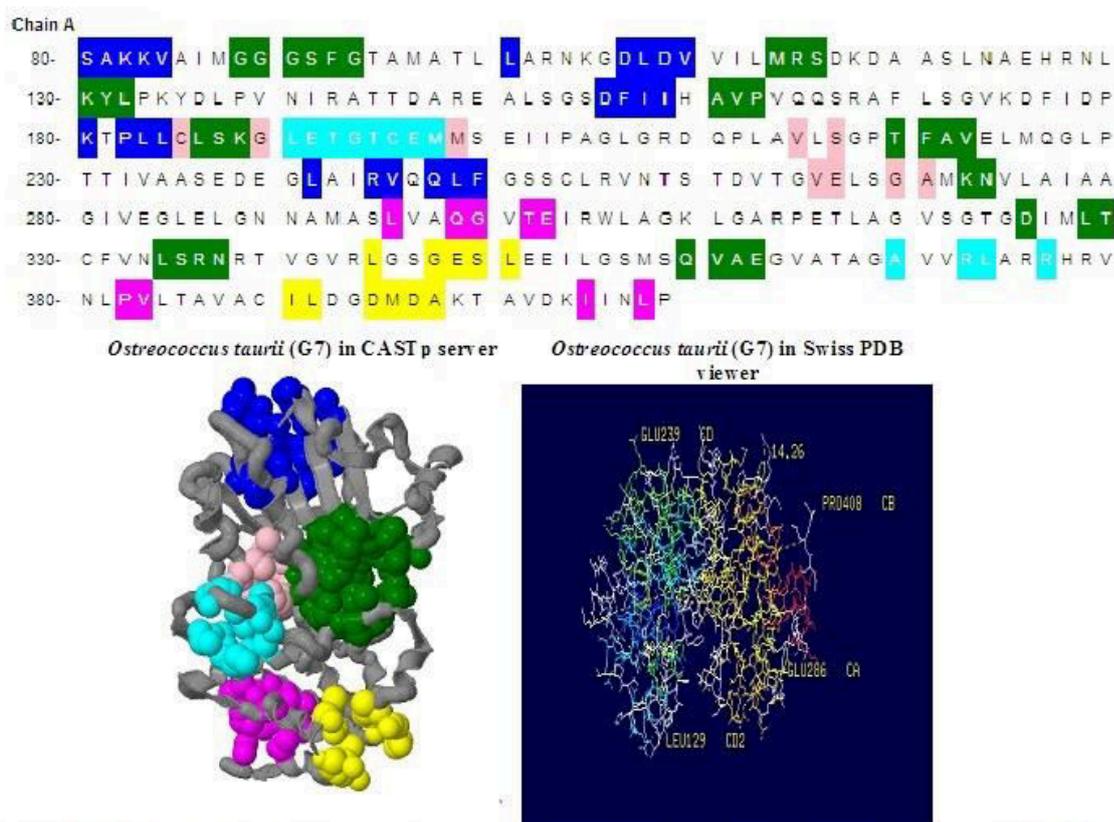

Fig. 3: ACTIVE SITE RESIDUES FOR *Ostreococcus taurii* (G7) USING CASTp SERVER. THE ACTIVE SITE RESIDUES ARE SHOWN IN GREEN COLOUR

IV. CONCLUSION

Identification of genes responsible for oil accumulation is a pre requisite to target microalgae for yields of biofuel precursors using metabolic engineering. Secondary structure analysis revealed that the proteins are largely alpha helical; this was supported by the homology modeling of three enzymes of *Chlamydomonas reinhardtii* (L2), *Osterococcus lucimarinus* (F7) and *Osterococcus taurii* (G7). High scoring values obtained by protein validation tools viz., Ramachandran Plot, VERIFY 3D and ERRAT indicated the overall accuracy of the developed models. More active sites are present in *Osterococcus lucimarnus* (F7) than in *Chlamydomonas reinhardtii* (L2) and *Osterococcus taurii* (G7). The results revealed that although each of the algal species maintain the basic genomic sequences required for lipid biosynthesis, they possess additional lineage-specific gene groups. More genomic, proteomic and metabolomic studies on algae lipid biosynthesis should also be nearing completion. This *in silico* study might be useful in manipulating enzymes involved in synthesis of fatty acids for algal biofuel engineering.

REFERENCES

[1] Misra, N., Panda, P. K., Parida, B. K., and Mishra, B. K., "Phylogenomic Study of Lipid Genes Involved in Microalgal Biofuel Production—Candidate Gene Mining and Metabolic Pathway Analyses". *Evolutionary Bioinformatics,* vol.8, pp.545–564, 2012.

[2] Murphy, D. J., "The biogenesis and functions of lipid bodies in animals, plants and microorganisms". Progress in Lipid Research, vol.40(5), pp.325-438, 2001.

[3] Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., "Basic local alignment search tool". *Journal of Molecular Biology*, vol.215, pp.403–10, 1990.

[4] Messaoudi, A., Belguith, H., and Hamida, J.B., "Three-Dimensional Structure of *Arabidopsis thaliana* Lipase Predicted by Homology Modeling Method". *Evolutionary Bioinformatics*, vol.7, pp. 99–105, 2011.

[5] Baral, M., Misra, N., Panda, P.K., and Thirunavoukkarasu, M., "*In silico* structural determination of GPAT enzyme from *Ostreococcus lucimarinus* for biotechnological application of microalgal biofuel production". *Biotechnol. and Biotechnol. Eq,* vol.26(1), pp.2794-2800, 2012.

[6] Gilbert, D., "Sequence file format conversion with command-line readseq, in Current Protocols in Bioinformatics", A. D., Baxevanis, G. A., Petsko, L. D., Stein, and G. D. Stormo, Eds., John Wiley & Sons, Hoboken, NJ, USA, 2003.

[7] Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., and McGettigan, P.A., "Clustal W and clustal X version 2.0". *Bioinformatics*, vol.23, pp.2947-2948, 2007.

[8] Rice, P., Longden, I., and Bleasby, A., "Emboss: The European molecular biology open software suite". *Trends Genet*, vol.16, pp.276-277, 2000.

[9] Deleage, G., and Roux, B., "An algorithm for protein secondary structure prediction based on class prediction". *Protein Eng*., vol.1, pp.289-294, 1987.

[10] Kyce, J., Doolittle, R.F., "A simple method for displaying the hydropathic character of a protein", *J. Mol. Biol*., vol.157(6), pp.105-132, 1982.

[11] Arnold, K., Bordoli, L., Kopp, J., and Schwede, T., "The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics", vol.22, pp.195-201, 2006.

[12] Guex, N, Peitsch, M.C., "Swiss-Model and the Swiss-Pdb Viewer: an environment for comparative protein modeling". *Electrophoresis*. Vol.18(15), pp.2714–23, 1997.

[13] Laskowiski, R.A, MacArthur, M.V., Moss, D.S., and Thornton, J.M., "PROCHECK: A program to check the stereochemical quality of protein structures". *J. Appl. Cryst*., vol.26, pp.283-291, 1993.

[14] Hooft, R.W.W., Vriend, G., Sander, C., Abola, E.E., "Errors in protein structures". *Nature*, vol.381, pp.272-292, 1996.

[15] Eisenhaber, F., Imperiale, F., Argos, P., and Frommel, C., "Prediction of secondary structural content of proteins from their amino acid composition alone. I. New analytic vector decomposition methods," *Proteins*, vol.25(2), pp.157–168, 1996.

[16] Colovos, C., and Yeates, T.O., "Verification of protein structures: patterns of non bonded atomic interactions" *Protein. Sci*, vol.2, pp.1511-1519, 1993.

[17] Lovell, S.C., Davis, I.W., Arendall, W.B., De Bakker, P.I., Word, J.M., *et al.* "Structure validation by Cα geometry:CΨ and Cβ deviation". *Proteins*, vol.50, pp.437-450, 2003.

[18] Vriend, G., "What If: A molecular modeling and drug design program". *J. Mol. Graphics*, vol.8, pp.52-56, 1990.

[19] Willard, L., Ranjan, A., Zhang, H., Monzavi, H., Boyko, R.F., Sykes, B.D., and Wishart, D.S., "VADAR: a web server for quantitative evaluation of protein structure quality". *Nucleic Acids Res*., vol.31, pp. 3316-3319, 2003.

[20] Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., and Liang, J., "CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues". *Nucleic Acids Res*., vol.34, pp.W116-W118, 2006.

[21] Smith, AA., Coomes, M.W., and Smith, T.E., "Isolation and sequence of the phosphoenolpyruvate carboxylase gene of the marine cyanobacterium *Synechococcus* PCC 7002", *J. Boil. Sci.,* vol.8, pp.1261-1270, 2008.

[22] Guschina, I.A., and Harwood, J.L., "Lipids and lipid metabolism in eukaryotic algae" Progress in Lipid Research, vol.45(2), pp.160-186, 2006.

[23] Thompson, G.A.J., "Lipids and membrane function in green algae". *Biochimica et Biophysica Acta,* vol.1302(1), pp.17-45, 1996.

[24] Merchant, S.S., Prochnik, S.E., Vallon, O., Harris, E.H., and Karpowicz, S.J., Witman, G.B., Terry, A., Salamov, A., Fritz-Laylin, L.K., Maréchal-Drouard, L., *et al.* "The *Chlamydomonas* Genome Reveals the Evolution of Key Animal and Plant Functions". *Science*, vol.318(5848), pp.245-250, 2007.

[25] Wang, Z.T., Ullrich, N., Joo, S., Waffenschmidt, S., and Goodenough, U., "Algal lipid bodies: stress induction, purification, and biochemical characterization in wild-type and starchless *Chlamydomonas reinhardtii*". *Eukaryot Cell*, vol.8(12), pp.1856-1868, 2009.

[26] Wei-Luen, Yu., William, A., Nathan, G.S., Michael, J.H., Stephen, P.M., and Michael, D.B., "Modifications of the metabolic pathways of lipid and triacylglycerol production in microalgae", *Microbial Cell Factories*., vol.10, pp.91, 2011.

[27] Durrett, T.P., Benning, C., and Ohlrogge, J., "Plant triacylglycerols as feedstocks for the production of biofuels". *Plant Journal*, vol.54(4), pp.593-607, 2008.

[28] Milcamps, A., Tumaney, A.W., Paddock, T., Pan, D.A., Ohlrogge, J., and Pollard, M., "Isolation of a gene encoding a 1,2-diacylglycerol-sn-acetyl-CoA acetyltransferase from developing seeds of *Euonymus alatus*". *Journal of Biological Chemistry*, vol.280(7), pp.5370-5377, 2005.

[29] Mus, F., Dubini, A., Seibert, M., Posewitz, M.C., and Grossman, A.R., "Anaerobic acclimation in *Chlamydomonas reinhardtii* - Anoxic gene expression, hydrogenase induction, and metabolic pathways". *Journal of Biological Chemistry*, vol.282(35), pp.25475-25486, 2007.

[30] Mayfield, S.P., Manuell, A.L., Chen, S., Wu, J., Tran, M., Siefker, D., Muto, M., Marin-Navarro, J., "*Chlamydomonas reinhardtii* chloroplasts as protein factories". *Current Opinion in Biotechnology*, vol.18(2), pp.126-133, 2008.

[31] León-Bañares, R., González-Ballester, D., Galván, A., and Fernández, E., "Transgenic microalgae as green cell-factories". TRENDS in Biotechnology, vol.22(1), pp.45-52, 2004.