Research Journal of Recent Sciences _____ E-ISSN 2277-2502
Vol. **5(9),** 30-37, September **(2016)** *Res. J. Recent Sci.*

Docking Studies of Heme Ligand onto the predicted 3D Structure of *Fatty Acid Desaturase 2* from Rat

Mutangana Dieudonne*, Simurabiye Jean Baptiste, Uwiringiyimana Thadee and Rwibasira Peter
University of Rwanda, College of Science and Technology, Department of Biology, Avenue de l'Armée, Po. Box 3900, Kigali-Rwanda
mature02@gmail.com

Available online at: www.isca.in, www.isca.me
Received 16th August 2016, revised 28th August 2016, accepted 31th August 2016

Abstract

Fatty acid desaturase 2 is a membrane bound enzyme of fatty acid desaturase family. It is encoded by a gene Fads2 located on the chromosome 1 of the rat genome. Like other membrane bound protein it has important physiological and industrial importance. Most importantly, the synthesis of polyunsaturated fatty acids requires the presence of fatty acid desaturase 2. However, the lack of its three-dimensional structure hinders the understanding of its biological function at molecular level. Sequence analysis was done using sequence alignment and phylogenetic which shows the evolutionary inferences between the sequence under study and its homologous proteins retrieved from PDB databank. To investigate in its functions, docking studies of heme ligand onto the predicted 3D structure of Fatty acid desaturase 2 from rat was conducted using computational methods. The ligand got docked onto the binding sites of the predicted model with more or less similar interacting residues when compare to interacting residues between heme and the experimentally determined structure of the template protein used while building the 3D model of fatty acid desaturase 2 through homology modeling technique.

Keywords: Docking, ligand, Model, 3Dimensional structure, Interacting residues, Binding domain.

Introduction

Fatty acid desaturase 2 is a membrane bound enzyme, also called delta (6) fatty acid desaturase (D6D), encoded by the *FADS2* gene located on the chromosome 1 of the rat genome[1]. This enzyme is involved in lipid metabolic pathway of highly unsaturated fatty acids (HUFAs) through substrates binding[2]. HUFAs are known to be important in membrane fluidity as well as in the inflammation processes and brain development[3]. It is demonstrated that this enzyme catalyzes both Linoleic acid (LA) and α linoleic acid (ALA) into γ linoleic acid and stearidonic acid respectively[4,5]. Watanabe and his co-authors worked on the designation of residues involved in the specific interaction of membrane binding fatty acids[6]. The same study shown elements needed for the conformational stabilization of the enzyme[6]. In another research, the involvement of HIS-box, a ligand binding motif required for electron transfer, was investigated[7]. In mammals, FADS2 plays a major role in cell signaling by acting as secondary messengers which activate the transcription factor[3]. FADS2 is also linked with both the decrease and increase of lipoprotein[3]. Activation of HUFAs are considered to have critical importance in the development of vertebrates and in the prevention of cancers, cardiovascular disease, and diabetes[8]. The investigation using site-directed mutagenesis techniques revealed that the activity of D6D depends on histidine of cytochrome domain[9]. Though the interaction between this domain and its homologs, significantly increases the enzyme activity; the function of HPGG motif is not yet clear and the role of cytochrome b5 to increase the desaturation activity is not known[9]. Many other investigations have been conducted on fatty acid desaturase 2 using both *in vivo* and *in vitro* techniques to find out its function but none of them characterized a clear structure function relationship. This reveals the necessity of structure-function characterization of this so important enzyme. This study aims to predict the 3-dimentional structure of the enzyme and evaluate the structure-function relationship using *in silico* techniques.

Methodology

Sequence information of Fatty acid desaturase 2 (accession number: q9z122) was retrieved from UniProt database[10]. The FASTA format was submitted to SWISS MODEL server to search for potential template structures[11-14] and PSI-BLAST iterated mode with PDB as database, was activated[15]. Theoretical structure of fatty acid desaturase 2 was predicted either by threading (I TASSER server)[16] or homology modelling (SWISS MODEL server)[11] approaches, based on sequence identity. Based on output generated by sequence alignment, multiple sequence alignment of fatty acid desaturase 2 and its top homologous proteins, was performed by accessing Clustal omega[17]. To evaluate the evolutionary relationship between fatty acid desaturase 2 and its homologous proteins, phylogenetic analysis was conducted using PHILIP package which generate a CONSENSE tree[18] via *seqboot*, *promlk* and *consense* programs of the package.

The three dimensional (3D) structure of fatty acid desaturase 2 was predicted by submitting the primary sequence information of this protein along with the template retrieved by PSI-BLAST

tool, to SWISS MODEL server in the automated mode[12]. The 3D structure initially obtained was submitted to ERRAT server[19] for quality factor control. Loop regions showing high percentage error was then refined using MODLOOP server[20], based on the output generated by ERRAT server. The refined 3D model was subjected to energy minimization using Deep view package then submitted to PROCHECK server[21] for checking its stereochemistry status[22].

Docking studies were undertaken by opening the energy minimized 3D model of fatty acid desaturase 2 and the 3D structure of *heme* ligand, into HEX software. This exercise was started by activating the shape and electrostatic energy of the software and both conformation and interaction energies were recorded at the completion of docking process.

Results and Discussion

Molecular modeling: PSI BLAST tool was able to retrieve sequence homologs from PDB databank having more that 30 % sequence identity which is essential for homology modeling (Figure-1). Among the retrieved homologs, 1CYO chain A was selected to serve as model structure for structure prediction. Multiple sequence alignment has shown conserved regions between the query sequence and top 5 structural homologs retrieved by PSI BLAST tool.

From Figure-2 showing multiple sequence alignment, some conserved domains were retrieved. These are "EEIQKH", "KVY", "HPG", "DAT" and "IGEL". Maithri and her collaborators have used Multiple sequence alignment to analyze sequence similarity between aminopeptidase N from *H. armigera* with selected PDB homologs using ClustalW[23].

The evolutional relationship between the query sequence and its homologs was analyzed. Phylip package was able to construct a consensus tree which shows the revel of relationship among compared homologous proteins. It appears that two sets of proteins ("1CYO-1HKO and 3OZZ") are closely related to fatty acid desaturase 2 and for this reason; it was chosen to serve as template while predicting the model structure of fatty acid desaturase through comparative modeling. 1CYO-1HKO are sister clades as they are placed on the same tree branch and 3OZZ can be considered as precursor of fatty acid desaturase 2. While 2I96 could be seen as the out group, a set of proteins placed above fatty acid desaturase 2 (seven proteins in a red box) can evolutionally be considered as delivered from fatty acid desaturase. There are published articles in which phylogenetic analysis was used as one way to study evolutional similarities between the query and target sequences of proteins[24,25].

Run PSI-Blast iteration 22 with max 500 Go

Sequences producing significant alignments with E-value BETTER than threshold

Select: All None Selected:0 Yellow: sequences scoring below threshold on previous iteration

Alignments Download GenPept Graphics Distance tree of results Multiple alignment

Description	Max score	Total score	Query cover	E value	Ident	Accession	Select for PSI blast	Used to build PSSM
Chain A, Nmr Structure Of Bovine Cytochrome B5	145	145	22%	4e-42	36%	1HKO_A	✓	✓
Chain A, Solution Structure Of The Oxidized Microsomal Human Cytochrome B5	142	142	22%	3e-41	35%	2I96_A	✓	✓
Chain A, The Solution Nmr Structure Of The B Form Of Oxidized Rat Microsomal Cytochrome B5, Minimized Average Structure	142	142	22%	4e-41	33%	1BFX_A	✓	✓
Chain A, Solution Structure Of The Water-Soluble Fragment Of Rat Hepatic Apocytochrome B5	142	142	22%	4e-41	33%	1I87_A	✓	✓
Chain A, Bovine Cytochrome B(5)	140	140	21%	1e-40	36%	1CYO_A	✓	✓
Chain A, Crystal Structure Of The Oxidized Form Of The Solubilized Domain Of Porcine Cytochrome B5 In Form 1 Crystal	139	139	22%	3e-40	35%	3X32_A	✓	✓
Chain A, Solution Nmr Structure Of Full-length Oxidized Microsomal Rabbit Cytochrome B5	140	140	22%	8e-40	35%	2M33_A	✓	✓
Chain A, Solution Structure Of Reduced Microsomal Rat Cytochrome B5, Nmr, Minimized Average Structure	137	137	21%	2e-39	34%	1AQA_A	✓	✓
Chain A, Solution Structure Of A67v Mutant Of Rat Ferro Cytochrome B5	137	137	21%	2e-39	35%	1JEX_A	✓	✓
Chain A, Solution Structure Of F35v Mutant Of Rat Ferro Cytochrome B5, A Conformation, Ensemble Of 20 Structures	136	136	21%	5e-39	34%	1IB7_A	✓	✓
Chain A, Solution Structure Of Oxidized Microsomal Rabbit Cytochrome B5, Factors Determining The Heterogeneous Binding Of The Heme	136	136	21%	6e-39	36%	1DO9_A	✓	✓
Chain A, Solution Structure Of Oxidized Rat Microsomal Cytochrome B5 In The Presence Of 2 M Guanidinium Chloride: Monitoring The Early Steps In Pr	135	135	21%	7e-39	34%	1BLV_A	✓	✓
Chain A, Structure Of Septuple Mutant Of Rat Outer Mitochondrial Membrane Cytochrome B5	135	135	20%	8e-39	32%	2I89_A	✓	✓
Chain A, Rat Outer Mitochondrial Membrane Cytochrome B5	133	133	20%	4e-38	32%	1AWP_A	✓	✓
Chain B, Structure Of A Cytochrome B5 Core-Swap Mutant	133	133	18%	6e-38	37%	3OZZ_B	✓	✓
Chain A, 2a Resolution Structure Of Rat Type B Cytochrome B5	133	133	18%	8e-38	32%	3MUS_A	✓	✓

Figure-1
Top sixteen homologs retrieved by PSI BLAST tool at the completion of 21st iteration

```
CLUSTAL O(1.2.2) multiple sequence alignment


sp|Q9Z122|FADS2_RAT      MGKGGNQGEGSTELQAPMPTFRWEEIQKHNLRTDRWLVIDRKVYNVTKWSQRHPGGHRVI
gi|2I96_A                ------------------YYTLEEIQKHNHSKSTWLILHHKVYDLTKFLEEHPGGEEVL
gi|1HKO_A                --------------------TLEEIQKHNNSKSTWLILHYKVYDLTKFLEEHPGGEEVL
gi|1CYO_A                --------------SKAVKYYTLEEIQKHNNSKSTWLILHYKVYDLTKFLEEHPGGEEVL
gi|1BFX_A                ----------AEQSDKDVKYYTLEEIQKHKDSKSTWVILHHKVYDLTKFLEEHPGGEEVL
gi|1I87_A                ----------AEQSDKDVKYYTLEEIQKHKDSKSTWVILHHKVYDLTKFLEEHPGGEEVL
                                   ******:  .. *:::. ***::**: :.****..*:

sp|Q9Z122|FADS2_RAT      GHYSGEDATDAFRAFHLDLDFVGKFLKPLLIGELAPEEPSLDRGKSSQITE
gi|2I96_A                REQAGGDATENFEDVGHSTDA-REMSKTFIIGELHPDDRPKLNKPPE----
gi|1HKO_A                REQAGGDATENFEDVGHSTDA-RELSKTFIIGELHPDDRSKITKPSE----
gi|1CYO_A                REQAGGDATENFEDVGHSTDA-RELSKTFIIGELHPDDRSKITKPSE----
gi|1BFX_A                REQAGGDATENFEDVGHSTDA-RELSKTYIIGELHPDDRSKIAKPSE----
gi|1I87_A                REQAGGDATENFEDVGHSTDA-RELSKTYIIGELHPDDRSKIAKPSE----
                          . :* ***: *. .  . *   :: *  :**** *::        .
```

Figure-2
Multiple sequence alignment showing fully conserved motifs (underlined in red) among the alignment proteins

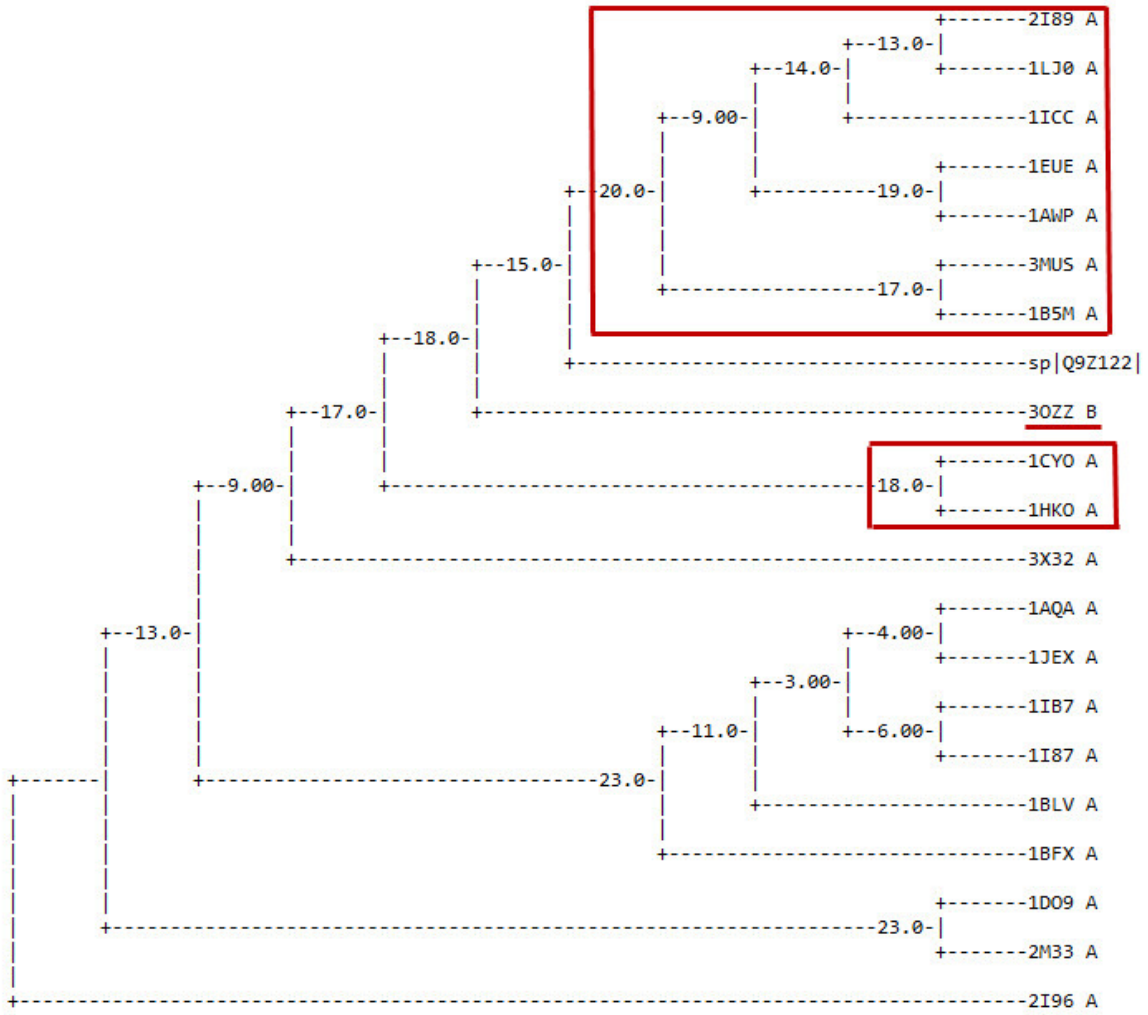


Figure-3
Consensus unrooted tree generated for fatty acid desaturase 2 from rat using *Promlk* program of PHYLIP.

The number on the branches indicates the number of times a given branch appeared among the tees out of 100 trees generated by the package.

Structure prediction: Both the results of sequence alignment and phylogenetic analysis gave insight on the template to be use for three-dimensional structure prediction of fatty acid desaturase 2. Based on both analyzes, 1CY0 was selected to as the template for homology modelling. This protein was selected based on the fact that it was determined using x-ray crystallography with a better resolution and it is among those with higher sequence identity, compared to other proteins candidates to serve as template for homology modeling exercise. Swiss model was able to generate a model of fatty acid desaturase 2. This model is generally considered to be the best one as it recorded the highest C-score. A model with a higher C score indicates a better structural match with the template structure used to build the model. This also is of great importance in anticipating the protein functions by utilizing the template protein[26]. The model obtained was submitted to ERRAT server for quality factor. This server shown that the quality was of accepted range and there was no loop to be refined (Figure-4).

The energy minimized structure was further validated using PROCHECK server. The later revealed that 89.2 % residues were in the favored region and the remaining 10.8% were in the allowed region with a reasonable overall G factor (-0.6). A G value of not less than -0.5 is indicates the satisfaction quality of the predicted model[27]. Prabhavathi *et al*., used ROCHECK to check the quality of the predicted thioredoxin (TRX) protein[28].

Both the predicted model of fatty acid desaturase and the template structure used to predict it were superimposed. They got more or less superimposed with overall RMSD of 0.333 Å. The structural sequence alignment shows that more regions of both proteins got fully aligned (Fig 7).In our previous study, we used structural superimposition as one of ways to model the interaction between MC2R and ACTH models from human[29].

Docking studies: Hex software was successful in docking the ligand (Heme) onto the predicted model of fatty acid desaturase 2 and has generated ten poses. Among all poses, the first one having a higher docking energy (-365.65 kcal mol^{-1}) was taken for further analysis. Active residues of the complex analyzed by Deep view software shown 20 active residues. These revealed interacting residues are: Leu[37,77,86,89], Ile[39,60], Tyr[44,63], Val[46,59,82], Trp[49], His[53], Pro[54], Gly[55], Ala[68,71], Phe[72,75,81]. Majority of these interacting residues being nonpolar "**Leu, Val, Pro**, **Gly**, **Ala**, **Phe, Trp** and **Ile**". Interacting residues between the native template structure 1CYO, used for homology modeling of fatty acid desaturase 2 and native heme ligand were analyzed for comparison. While the predicted model interacts with the ligand with 20 residues, the template structure interacts with 23 which are: Leu[23,**24**,32,46,70], Tyr[30], Phe[35,58,74], His[39,63] Pro[40], Gly[41,62], Val[45,61], Gln[49], Ala[54,67], Asn57, Ser[64,71], and Asn[66].Umme and her colleagues docked epigallocatechin (EGC) onto the middle domain of theoretical structure of cytoplasmic Hsp90 dimer of Arabidopsis thaliana, in their study "molecular dynamics simulation of homology modeled cytosolic hsp90 iso form from *Arabidopsisthaliana"*[30]

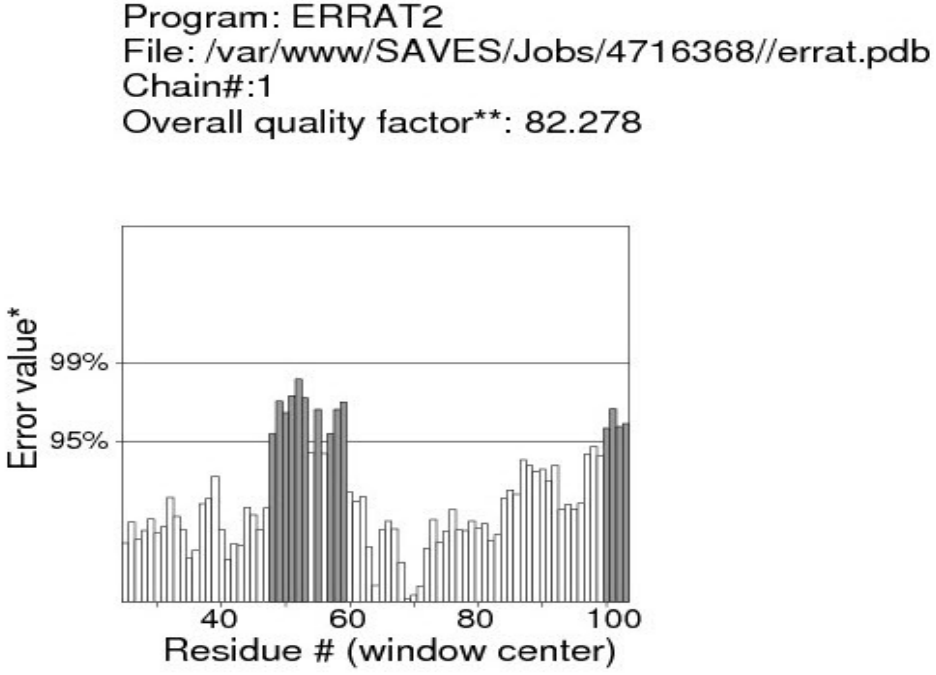
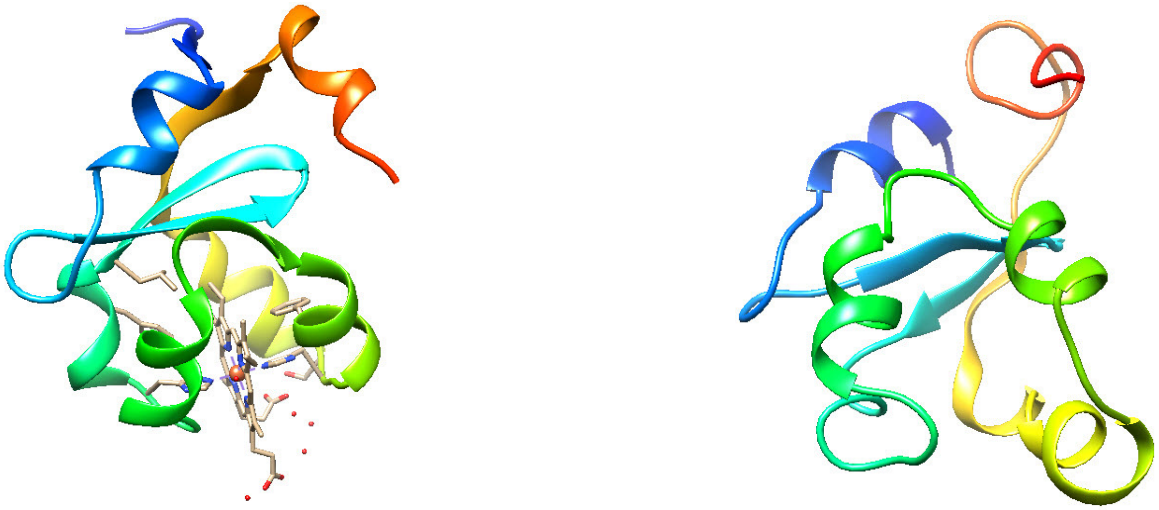


Figure-4
Errat plot showing the quality factor of the energy minimized predicted model of fatty acid desaturase 2



(a) (b)

Figure-5
(a) template structure used for structure prediction of the model (b) of fatty acid desaturase 2 from rat. The image was generated by CHIMERA package

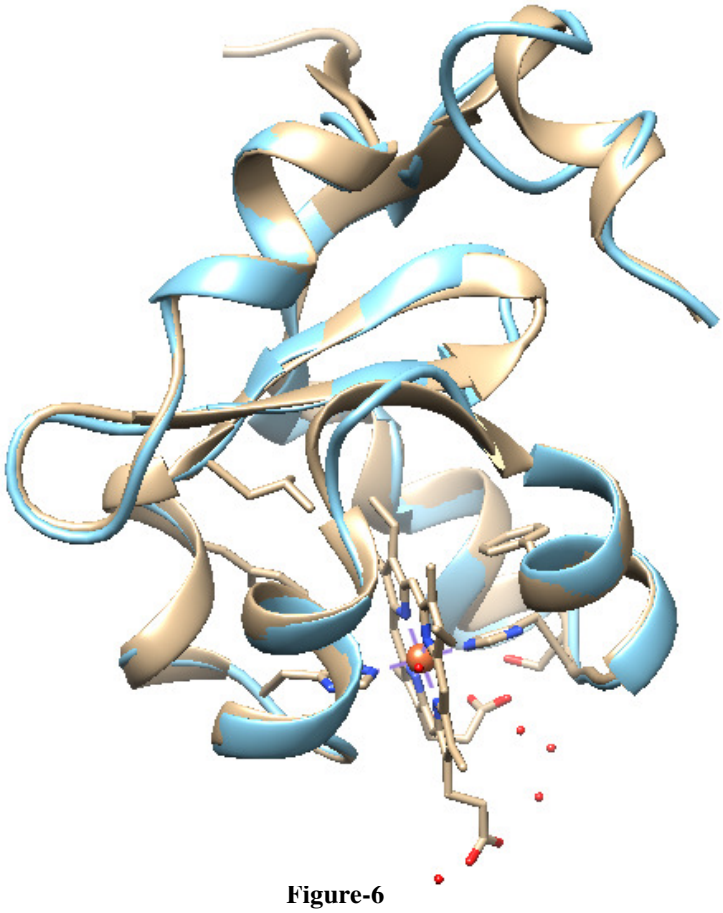


Figure-6
Superimposition of the model structure predicted onto the template structure (1CYO colored in cyan). The image was generated using CHIMERA package

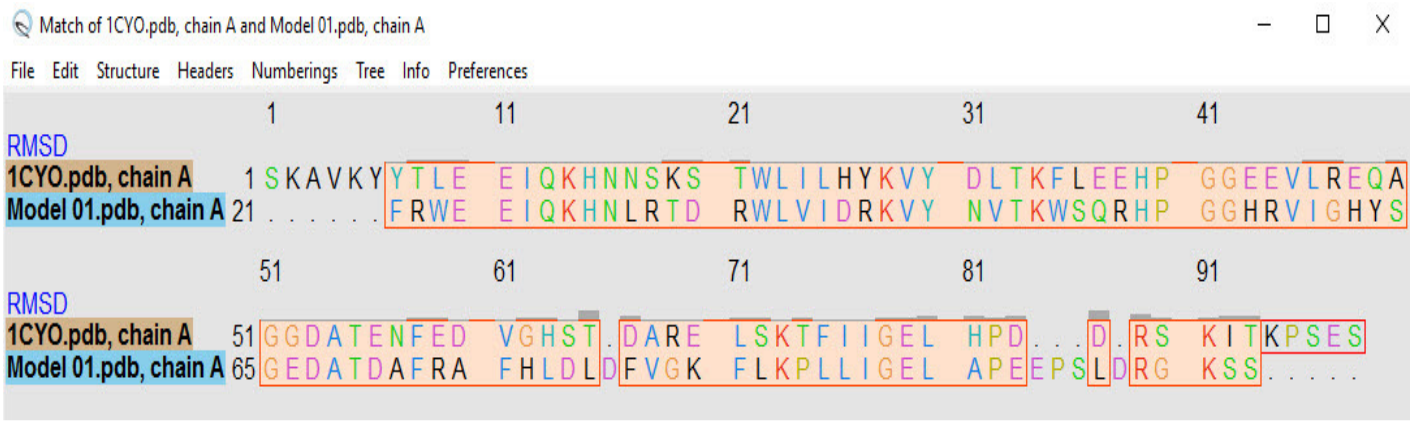


Figure-7
Structural alignment of the template structure (1CYO) and the predicted model of fatty acid desaturase 2 from rat. The image was generated using CHIMERA package

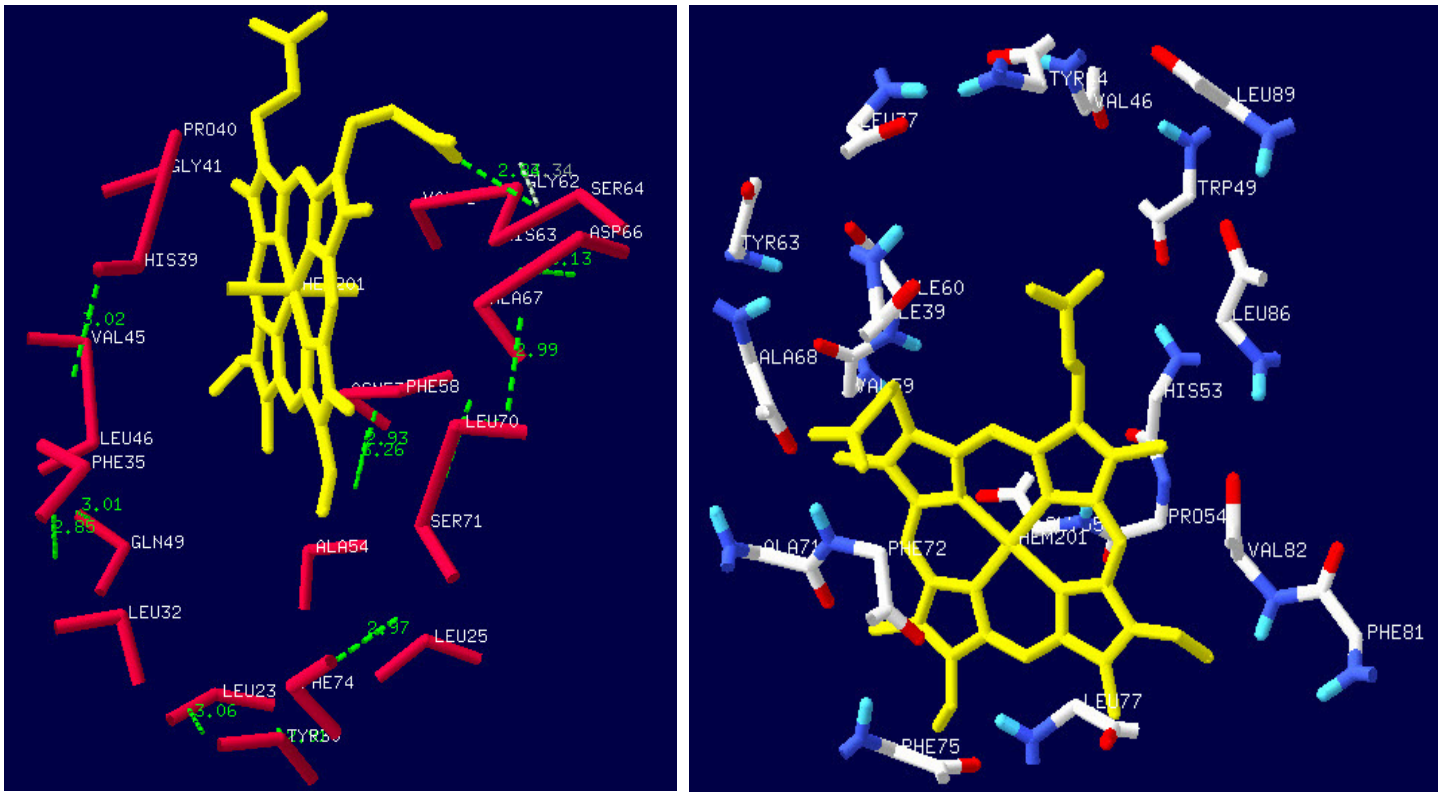


Figure-8
Interacting residues between (a) HEME ligand with the template structure and (b) HEME with the predicted model of fatty acid desaturase 2 from rat. Image was generated using PyMol software

Conclusion

The predicted structure of fatty acid desaturase 2 contains a cytochrome b5 binding domain. This domain acts as the catalytic region since it is similar to 1CY0, the protein used to predict the 3D model of fatty acid desaturase 2. The overall function should be electron transfer during the process of desaturation. Docking studies confirmed the affinity of heme ligand towards the predicted model. While twenty residues of the predicted 3D model were seen to be interacting with heme ligand, 23 residues of the protein used as template structure for homology modeling exercise; were interacting with the ligand. All interacting residues from both the model and template are believed to be of the same binding domain. This study revealed that fatty acid desaturase 2 from rat has the binding site of heme ligand.

References

1. Aki T, Shimada K, Inagaki K, Higashimoto H, Kawamoto, S, Shiget S, Ono K and Suzuki O. (1999). Molecular cloning and functional characterisation of rat Δ-6 fatty acid desaturase. *Biochem Biopyhs Res Commun*, 255, 575-579

2. Katsuya I, Tsunehiro A, Taketoshi S, Seiji K, Seiko S, Osamu S and O K (2014). Evidence of Isozymes for Δ6 Fatty Acid Desaturase in Rat Hepatocytes. *Bioscience, Biotechnology, and Biochemistry*, 67, 451-454.

3. Lattka E, Eggers S, Moeller G, Heim K, Weber M, Mehta D, Prokisch H, Illig T and Adamski J. (2010). A common FADS2 promoter polymorphism increases promoter activity and facilitates binding of transcription factor ELK1. *J Lipid Res*, 51, 182-191.

4. D'andrea S, Guillou H, Jan S, Catheline D, Thibault, Bouriel M, Rioux V and Legrand P. (2002). The same rat D6-desaturase not only acts on 18- but also on 24-carbon fatty acids in very long chains polyunsaturated fatty acids biosynthesis. *Biochem J*, 364, 49-55.

5. Thomas S, Thien N T, Helge R, Bjørn O C and BH. T. (2003). Expression and Regulation of Δ5-Desaturase, Δ6-Desaturase, Stearoyl-Coenzyme A (CoA) Desaturase 1, and Stearoyl-CoA Desaturase 2 in Rat Testis. *Biology of Reproduction*, 69, 117-124.

6. Watanabe K, Ohno M, Taguchi M, Kawamoto S, Ono K and T A. (2015). Identification of amino acid residues that determine the substrate specificity of mammalian membrane-bound front-end fatty acid desaturases. *J Lipid Res*, 57, 89-99.

7. Nakamura M and Nara T. (2004). Structure, function, and dietary regulation of Δ-6, Δ-5, and Δ-9 desaturases. *Annu Rev Nutr*, 24, 345-376.

8. Je M L, Hyungjae L, SeokBeom K and P WJ. (2016). Fatty Acid Desaturases, Polyunsaturated Fatty Acid Regulation, and Biotechnological Advances. *Nutrients review*, 8, 23-36.

9. Guillou H, D'Andrea S, Rioux V, Barnouin R, Dalaine S, Pedrono F, Jan S and P. L. (2004). Distinct roles of endoplasmic reticulum cytochrome b5 and fused cytochrome b5-like domain for rat Δ-6-desaturase activity. *J Lipid Res*, 45, 32-40.

10. UniProtConsortium. (2014). UniProt: a hub for protein information. Nucleic acids research.

11. Marco B, Stefan B, Andrew W, Konstantin A, Gabriel S, Tobias S, Florian K, Tiziano G C, Martino B, Lorenza B and S T. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research*, 42, W252-W258.

12. Arnold K, Bordoli L, Kopp J and T S. (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics*, 22, 195-201.

13. Kiefer F, Arnold K, Künzli M, Bordoli L and T S. (2009). The SWISS-MODEL Repository and associated resources. *Nucleic Acids Research*, 37, D387-D392.

14. Guex N, Peitsch M C and T S. (2009). Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis*, 30, S162-S173.

15. Altschul S F, Madden T L, Schäffer A A, Zhang J, Zhang Z, MillerW and J LD. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25, 389-3402.

16. Zhang Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC bioinformatics*, 9, 40-47.

17. Sievers F , Wilm A DD, Gibson TJ , Karplus K , Li W , Lopez R , McWilliam H , Remmert M , Söding J , Thompson JD and DG H. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7, 539.

18. Felsenstein J. PHYLIP (1989). Phylogeny Inference Package. *Cladistics*, 5, 164-166.

19. Colovos C and Yeates TO. (1993). Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Science*, 2, 1511-1519.

20. Fiser A and A. S. (2003). ModLoop: automated modeling of loops in protein structures. *Bioinformatics*, 19, 2500-2501.

21. Laskowski R A, MacArthur M W, Moss D S and M TJ. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Cryst*, 26, 283-291.

22. Guex N and Peitsch M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, 18, 2714-2723.

23. Maithri S K, Ramesh K V and Mutangana D. (2013). Theoretical structure prediction of TcaA from Photorhabdus luminescens and aminopeptidase N receptor from Helicoverpa armigera. *Research Journal of Recent Sciences*, 2, 40-49.

24. Zhao FP, Fan HY, Li GH and BK Z. (2016). Complete mitochondrial genome sequence and gene organization of Chinese indigenous chickens with phylogenetic considerations. *Genet Mol Res*, 15.

25. Zhao YY, Su LN, Zhang ZM and XY W. (2016). Phylogenetic relationships of Pseudohynobius (Urodela, Hynobiidae) inferred from DNA barcoding analysis. *Genet Mol Res*, 15.

26. Wang Z X, Zhang W, Wu C, Lei H, Cieplak P and Duan Y. (2006). Strike a balance: Optimization of backbone torsion

parameters of AMBER polarizable force field for simulations of proteins and peptides. *J Comp Chem*, 27, 781-790.

27. Kleywegt GJ and Jones TA. (1996). Phi/Psi-chology: Ramachandran revisited. *Structure*, 6, 4, 1395-1400.

28. Prabhavathi M, Ashokkumar K, Geetha N and M SDK. (2011). Homology modeling and structure prediction of thioredoxin (TRX) protein in wheat (Triticum aestivum L.). *Intl J Biosci*, 1, 20-30.

29. Mutangana D and Ramesh K V. (2015). Modeling the interactions between MC2R and ACTH models from human. *Journal of Biomolecular Structure and Dynamics*, 33, 770-788.

30. Umme H, Ramesh K V, Lochana P and Mutangana D. (2014). Molecular dynamics simulation of homology modeled cytosolic hsp90 isoform from Arabidopsis thaliana. *International Journal of Analytical, Pharmaceutical and Biomedical Sciences*, 3, 63-80.