

 <p>ISSN NO. 2320-5407</p>	<p>Journal Homepage: - www.journalijar.com</p> <h2 style="text-align: center;">INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)</h2> <p style="text-align: center;">Article DOI: 10.21474/IJAR01/4811 DOI URL: http://dx.doi.org/10.21474/IJAR01/4811</p>	 <p>INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR) ISSN 2320-5407 Journal Homepage: http://www.journalijar.com Journal DOI: 10.21474/IJAR01</p>
---	--	--

RESEARCH ARTICLE

MULTI-TEMPLATE HOMOLOGY MODELING OF HUMAN MCT8 PROTEIN .

Divya Shaji.

Department of Pediatrics, Aichi Medical University, Nagakute, Aichi, Japan.

Manuscript Info

Manuscript History

Received: 11 May 2017

Final Accepted: 13 June 2017

Published: July 2017

Key words:-

MCT8 protein, Allan–Herndon–Dudley syndrome, MCT8 deficiency disease, Multi-template homology modeling, 3D structure prediction.

Abstract

Monocarboxylate transporter-8 (MCT8) is a specific thyroid hormone transporter, necessary for the uptake of thyroid hormone into target tissues. Mutations in the MCT8 gene lead to a syndrome named Allan–Herndon–Dudley syndrome (AHDS), also known as MCT8 deficiency. At present, no treatment is available for Allan–Herndon–Dudley syndrome (AHDS). Knowledge of protein three-dimensional structure (3D) is essential for understanding the mechanism of protein function, and it is crucial for the development of target-specific drugs. Unfortunately, protein structure determination using experimental methods is time consuming and not successful with all proteins, especially membrane proteins. The three-dimensional structure (3D) of MCT8 protein is unavailable in the Protein Data Bank (PDB). In the absence of 3D structures, computational methods are used for the structure prediction of proteins. Availability of 3D-model of MCT8 protein would significantly improve the development of new drugs for Allan–Herndon–Dudley syndrome (AHDS). In this study, multi-template homology modeling was used to predict the 3D structure of MCT8 protein. Picking the right set of templates for homology modeling is a difficult problem. This study identified the most suitable templates for the 3D structure prediction of MCT8. Moreover, a good quality 3D model of MCT8 was predicted based on multiple templates using MODELLERv9.17.

Copy Right, IJAR, 2017,. All rights reserved.

Introduction:-

Thyroid hormone is essential in the development of different organs, particularly the brain. Monocarboxylate transporter-8 (MCT8) is a specific transporter of triiodothyronine (T3) hormone and MCT8 gene mutations cause a rare X-linked disorder named MCT8 deficiency, also known as Allan–Herndon–Dudley syndrome (AHDS), characterized by psychomotor retardation and hypotonia. An abnormality (mutation) in the MCT8 (SLC16A2) gene (located in Xq13.2) leading to alteration in the structure and function of the MCT8 protein. The abnormal protein is unable to transport thyroid hormone produced by the thyroid gland into the brain, thus affecting its development. Identified mutations include truncations, in-frame deletions, nonsense and missense mutations. At present, no treatment is available for Allan–Herndon–Dudley syndrome (AHDS) ^{1, 2, 3, 4}.

Protein structure is important for understanding the biological function of a protein, and it is crucial for the development of target-specific drugs. Protein structure is the three-dimensional arrangement of atoms in a protein molecule. The three-dimensional structure (3D) of MCT8 protein is unavailable in the Protein Data Bank

Corresponding Author:- Divya Shaji.

Address:- Department of Pediatrics, Aichi Medical University, Nagakute, Aichi, Japan.

(PDB) <http://www.rcsb.org>⁵. In the absence of 3D structures, computational methods are used for the structure prediction of proteins. Unfortunately, experimental approaches are time consuming and expensive. Various approaches have been developed to reduce the expense and time for drug discovery. Computer aided drug design (CADD) is one of the most effective methods for the development of novel drugs^{6,7}.

Availability of 3D-model of MCT8 protein would significantly improve the development of new drugs for MCT8 deficiency disease. Homology modeling, also known as comparative modeling is a method to generate a reliable 3D model of a protein from its amino acid sequence. The first step in homology modeling is to identify the most suitable templates with high sequence similarity to build the query 3D model. The sequence similarity between templates and MCT8 protein was found to be very low. Therefore, no structure could be used as a template. Templates of MCT8 with low sequence similarity increase the complexity in straightforward homology modeling, and hence there is a need to evaluate different modeling methodologies in order to use the models for computer aided drug design. In the current study, various computational methods were used for predicting the properties and 3D structure of MCT8 protein. This study predicted a good quality 3D model of human MCT8 protein using multi-template homology modeling⁸.

Materials and Methods:-

MCT8 protein sequence:-

The sequence of the human MCT8 protein (NP_006508.2, Gene ID: 6567) consists of 539-amino acids, retrieved from the National Center for Biotechnology Information (NCBI) <http://www.ncbi.nlm.nih.gov/protein>⁹.

Primary structure prediction:-

The physico-chemical properties of the human MCT8 protein were calculated by using the ExPasy's ProtParam tool¹⁰. Molecular weight, isoelectric point (pI), total number of positive and negative residues, extinction coefficient, aliphatic index, instability index and grand average hydropathicity (GRAVY) were calculated using the default parameters.

Secondary structure prediction:-

SOPMA (Self Optimized Prediction Method with Alignment)¹¹ was used to predict the secondary structure of the MCT8 protein. Protein's secondary structural elements (SSEs) such as helices, beta sheets, turns and coils were predicted from the MCT8 sequence.

Disorder prediction:-

Intrinsically disordered regions (IDRs) of MCT8 were predicted by using IUPred¹². IUPred server recognizes disordered regions from the amino acid sequence based on the estimated pairwise energy content. The underlying assumption is that globular proteins are composed of amino acids which have the potential to form a large number of favorable interactions, whereas intrinsically disordered proteins (IDPs) adopt no stable structure because their amino acid composition does not allow sufficient favorable interaction to form¹².

3D structure prediction of MCT8 using I-TASSER:-

I-TASSER server <http://zhanglab.ccmb.med.umich.edu/I-TASSER/>¹³ is an on-line platform for protein structure and function predictions. It allows users to automatically generate high-quality predictions of 3D structure and biological function of protein molecules from their amino acid sequences. Structural templates are first identified from the PDB⁵ by multiple threading approach LOMETS (Local Meta-Threading-Server)¹⁴; full-length atomic models are then constructed by iterative template fragment assembly simulations. Finally, function insights of the target are derived by threading the 3D models through protein function database BioLiP¹⁵. LOMETS¹⁴ is an on-line web service for protein structure prediction. It generates 3D models by collecting high-scoring target-to-template alignments from 9 locally-installed threading programs (FFAS-3D, HHsearch, MUSTER, pGenTHREADER, PPAS, PRC, ROSPECT2, SP3, and SPARKS-X). BioLiP¹⁵ is a semi-manually curated database for high-quality, biologically relevant ligand-protein binding interactions. The structure data are collected primarily from the PDB⁵, with biological insights mined from literature and other specific databases. BioLiP¹⁵ aims to construct the most comprehensive and accurate database for serving the needs of ligand-protein docking, virtual ligand screening and protein function annotation. BioLiP uses a composite automated and manual procedure for examining the biological relevance of ligands in the PDB database.

The quality of the predicted model depends on the value of TM-score. TM-score is a recently proposed scale for measuring the structural similarity between two structures. A TM-score >0.5 indicates a model of correct topology and a TM-score <0.17 means a random similarity¹³.

3D structure prediction of MCT8 using phyre2:-

Phyre2 server <http://www.sbg.bio.ic.ac.uk/phyre2/> was used to predict the 3D structure for MCT8 protein. Phyre2 uses advanced remote homology detection methods to build 3D models, predict ligand binding sites and analyze the effect of amino acid variants (e.g., nonsynonymous SNPs (nsSNPs)) for a user's protein sequence¹⁶.

3D structure prediction of MCT8 using RaptorX:-

RaptorX server <http://raptorx.uchicago.edu/StructurePrediction/predict/> was used to predict the 3D structure for MCT8 protein. RaptorX predicts 3D structures for protein sequences without close homologs in the Protein Data Bank (PDB). RaptorX also predicts secondary structures, solvent accessibility and disordered regions for an input sequence¹⁷.

3D structure prediction of MCT8 using MODELLER v9.17:-

The suggested templates from the I-TASSER¹³, Phyre2¹⁶ and RaptorX¹⁷ were subjected to comparative modeling using the advanced modeling feature of the MODELLER v9.17^{18,19,20,21}. MODELLER is used for homology or comparative modeling of protein three-dimensional structures. The user provides an alignment of a sequence to be modeled with known related structures and MODELLER automatically calculates a model containing all non-hydrogen atoms. MODELLER implements comparative protein structure modeling by satisfaction of spatial restraints, and can perform many additional tasks, including de novo modeling of loops in protein structures, optimization of various models of protein structure with respect to a flexibly defined objective function, multiple alignment of protein sequences and/or structures, clustering, searching of sequence databases, comparison of protein structures, etc^{18,19,20,21}.

Protein structure evaluation:-

The stereo chemical features of the predicted MCT8 models were evaluated with the program PROCHECK²² by Ramachandran plot²³ analysis which was done through "Protein structure and model assessment tools" of SWISS-MODEL workspace²⁴. The best model was selected by Ramachandran plot analysis.

Structure refinement:-

The protein structure refinement of MCT8 best model was predicted by ModRefiner²⁵. ModRefiner is an algorithm for atomic-level, high-resolution protein structure refinement, which can start from either C-alpha trace, main-chain model or full-atomic model. Both side-chain and backbone atoms are completely flexible during structure refinement simulations, where conformational search is guided by a composite of physics- and knowledge-based force field. It has an option to allow for the assignment of a second structure which will be used as a reference to which the refinement simulations are driven. One aim of ModRefiner is to draw the initial starting models closer to their native state, in terms of hydrogen bonds, backbone topology and side-chain positioning. It also generates significant improvement in physical quality of local structures. The standalone program also supports ab initio full-atomic relaxation, where the refined model is not restrained by the initial model or the reference model²⁵.

Results and Discussion:-

Primary and secondary structure prediction:-

The physico-chemical properties of the human MCT8 protein by using the ExPasy's ProtParam tool¹⁰ are shown in Table1. The secondary structural features of the MCT8 protein predicted by SOPMA¹¹ are shown in Table2. SOPMA result showed that MCT8 protein contains 37.85% alpha helices, 17.25% beta sheets, 8.91% turns and 35.99% coils.

Table 1: Different physico-chemical properties of human Monocarboxylate transporter 8 (MCT8) by using the ExPasy's ProtParam tool ¹⁰.

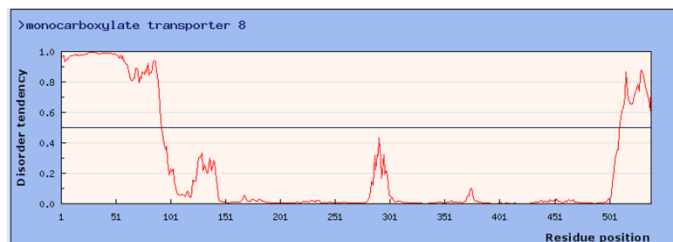
Parameters	Values
Molecular weight	59380.22
Theoretical pI	5.43
Total number of negatively charged residues (Asp + Glu)	50
Total number of positively charged residues (Arg + Lys)	39
Total number of atoms	8375
Ext. coefficient	54485
Abs 0.1% (=1 g/l) 0.918, assuming all pairs of Cys residues form cysteines	
Ext. coefficient	53860
Abs 0.1% (=1 g/l) 0.907, assuming all Cys residues are reduced	
The instability index (II)	50.21
Aliphatic index	91.54
Grand average of hydropathicity (GRAVY)	0.20

Table2: Secondary structure prediction of human Monocarboxylate Transporter 8 (MCT8) by using the SOPMA tool ¹¹.

Secondary structure	Percentage
Alpha helix	37.85%
Beta sheet	17.25%
Turn	8.91%
Coil	35.99%

Disorder prediction:-

Two Intrinsically disordered regions (IDRs) were identified in MCT8 by using IUPred ¹². The regions are from the amino acid number 1-92 and 511-539. The IUPred output is shown in Fig. 1.

**Fig. 1:-** Disordered regions of MCT8 predicted by IUPred ¹²**3D structure prediction using I-TASSER:-**

I-TASSER ¹³ server generated five 3D models for MCT8 protein. The first model was selected as the best model. The threading templates used by I-TASSER were 5l7dA, 1pw4A and 3wdoA. I-TASSER uses the TM-align structural alignment program to match the first I-TASSER model to all structures in the PDB library. I-TASSER reported the top 10 proteins from the PDB that have the closest structural similarity, i.e. the highest TM-score, to the predicted I-TASSER model. The proteins structurally close to the MCT8 protein in the PDB ⁵ were 1pw4A, 4pypA, 4ikvA, 4ldsA, 4w6vA, 2v8nA, 4cl4A, 4zowA, 4q65A and 4lepA. Due to the structural similarity, these proteins

often have similar function to the MCT8 protein. The best model for MCT8 predicted by I-TASSER is shown in Fig. 2A. The identified templates by I-TASSER were then forwarded for the MODELLERv9.17^{18,19,20,21} analysis.

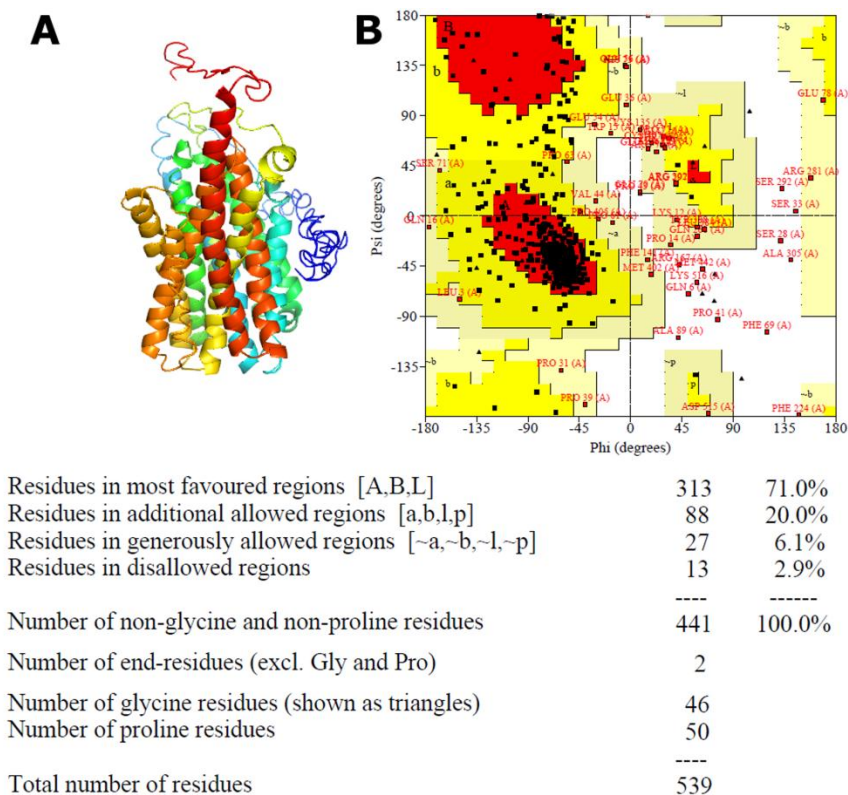


Fig. 2:- (A) 3D model of MCT8 protein predicted by I-TASSER¹³ (B) Protein structure evaluation of I-TASSER model by Ramachandran plot^{22,23}.

3D structure prediction using Phyre2:-

Phyre2¹⁶ server selected 6 different templates (1pw4A, 4cl5B, 3wdoA, 4apsB, 4j05A and 4ldsB) to model the MCT8 protein. 82% of the residues were modelled at >90% confidence. The Phyre2 final model for MCT8 is shown in Fig.3A. The identified templates by Phyre2 server were then forwarded for the MODELLERv9.17^{18,19,20,21} analysis.

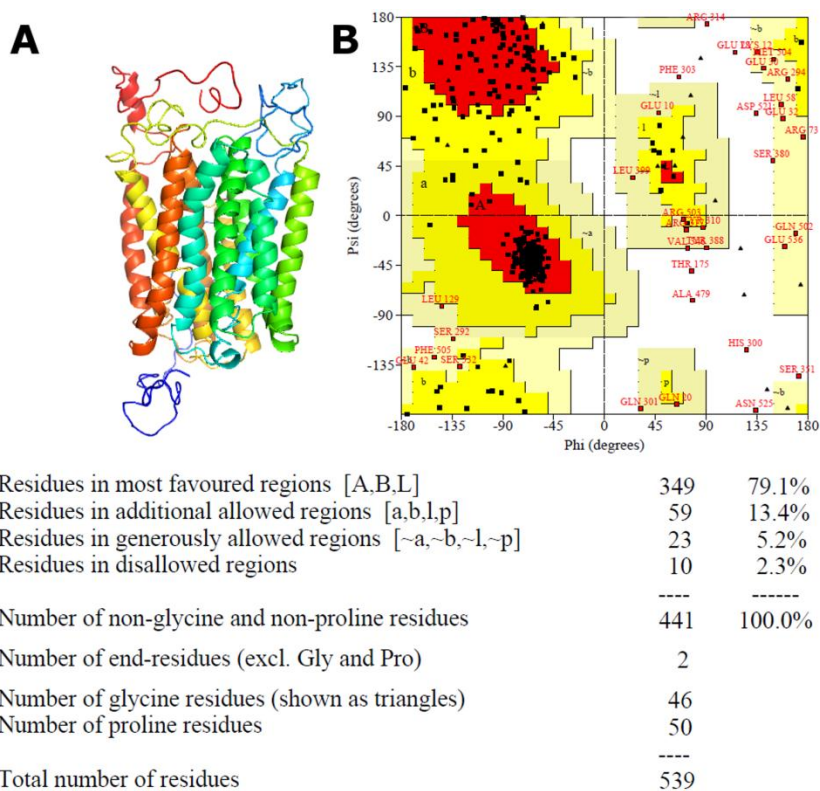


Fig. 3:- (A) 3D model of MCT8 protein predicted by Phyre2¹⁶ (B) Protein structure evaluation of Phyre2 model by Ramachandran plot^{22,23}.

3D structure prediction using RaptorX:-

446 residues (82%) of the MCT8 protein were modelled with RaptorX¹⁷ using 5 different templates (1pw4A, 4u4tA, 4ikxA, 4j05A, 4gbyA). RaptorX selected the best template as 1pw4A. The RaptorX predicted model is shown in Fig.4A which contained 65% helices, 0% strands and 34 % coils. The identified templates by RaptorX server were then forwarded for the MODELLERv9.17^{18, 19, 20, 21} analysis.

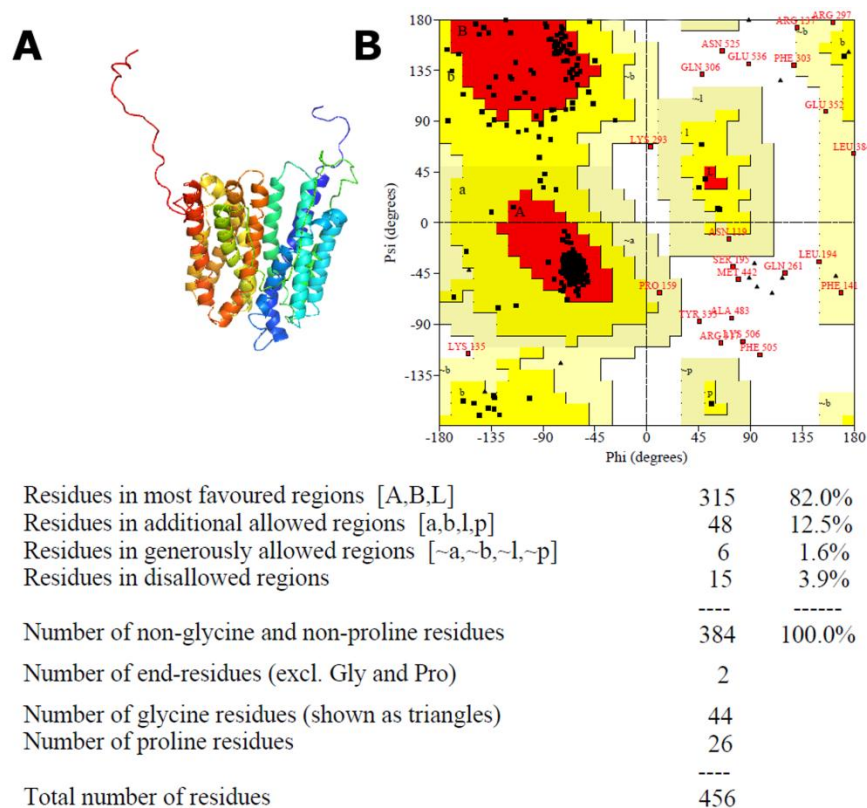


Fig 4:- (A) 3D model of MCT8 protein predicted by Raptorx¹⁷ (B) Protein structure evaluation of Raptorx model by Ramachandran plot^{22,23}.

3D structure prediction using MODELLERv9.17:-

The suggested templates from the I-TASSER¹³, Phyre2¹⁶ and RaptorX¹⁷ were subjected to comparative modeling using the advanced modeling feature of the MODELLERv9.17^{18,19,20,21}. The MODELLERv9.17 generated 5 models for MCT8 protein. The DOPE (discrete optimized protein energy) score from MODELLER was used to estimate the model quality. Out of 5 models generated by MODELLER, the best model was selected on the basis of lowest DOPE score. The best models generated by MODELLERv9.17^{18,19,20,21} for MCT8 using the suggested templates from the I-TASSER¹³, Phyre2¹⁶ and RaptorX¹⁷ are shown in Fig. 5A, 6A and 7A respectively.

Protein structure evaluation:-

The stereo chemical features of the predicted MCT8 models using MODELLERv9.17^{18,19,20,21} were evaluated with the program PROCHECK²² by Ramachandran plot²³ analysis. In the Ramachandran plot²³, a protein structure is evaluated by percentage quality measurements which include four kinds of classifications: most favored (core), additional allowed, generously allowed and disallowed regions. A good quality model should have over 90% in the most favored regions. The percentage of residues in the “most favorable/core” regions is one of the better guides to stereo chemical quality.

The MCT8 generated model from MODELLERv.9.17^{18,19,20,21} using the I-TASSER¹³ suggested templates had 80.0 % residues in the most favored region, 15.4 % in the additional allowed region, 3.6 % in the generously allowed region and 0.9 % in the disallowed region. The Ramachandran plot is shown in Fig 5B.

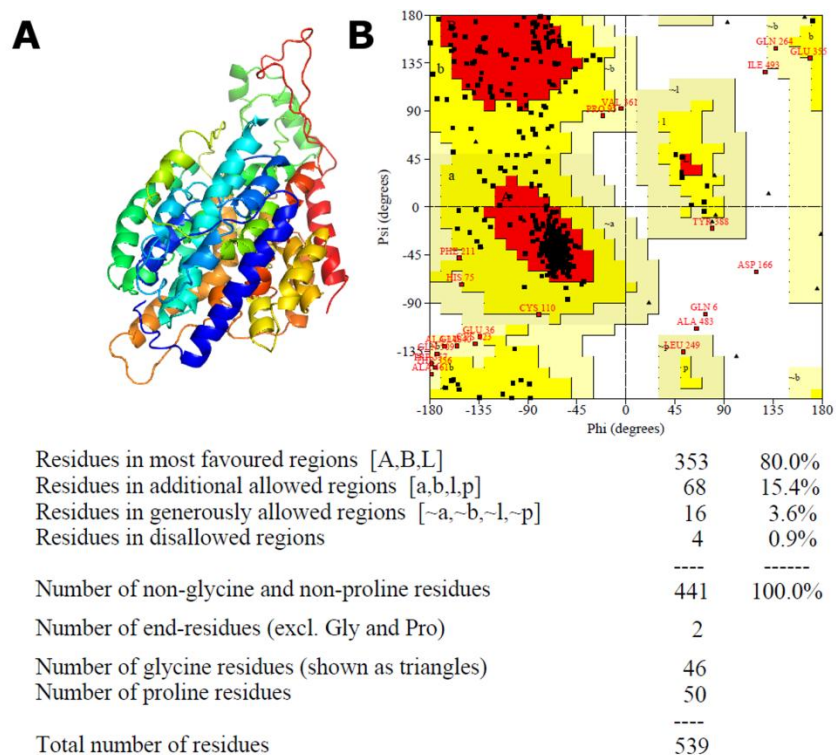


Fig. 5:- (A) Modeller9.17^{18,19,20,21} generated model of MCT8 using the suggested templates from I-TASSER¹³ (B) Protein structure evaluation of Modeller9.17 generated model using the suggested templates from I-TASSER by Ramachandran plot^{22,23}.

The MCT8 generated model from MODELLERv.9.17^{18,19,20,21} using the Phyre2¹⁶ suggested templates had 77.6 % residues in the most favored region, 16.1 % in the additional allowed region 3.2 % in the generously allowed region and 3.2 % in the disallowed region. The Ramachandran plot is shown in Fig.6B.

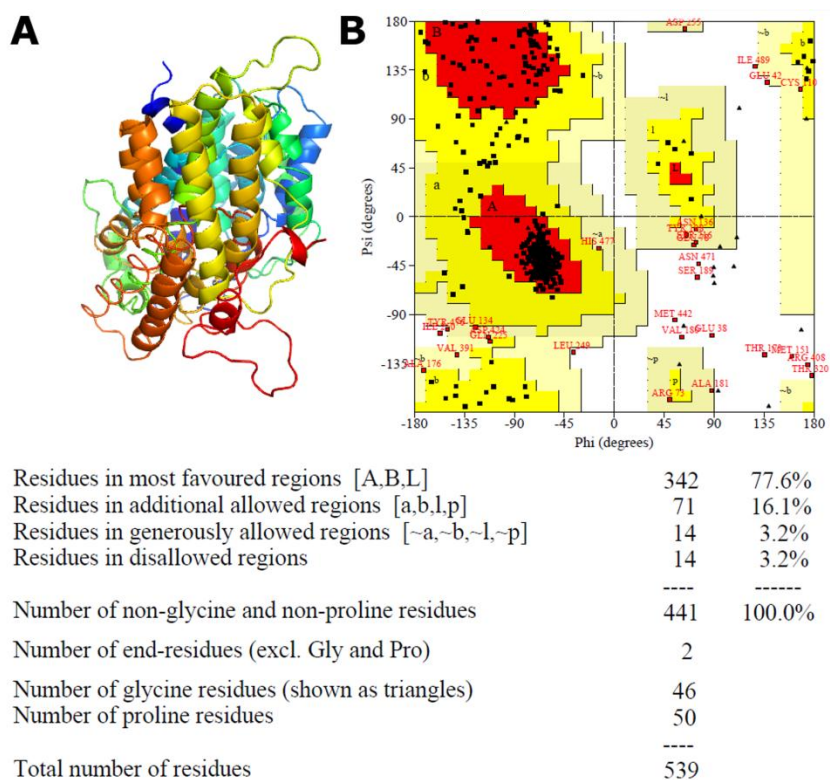


Fig. 6:- (A) Modeller9.17^{18,19,20,21} generated model Of MCT8 using the suggested templates from Phyre2¹⁶ (B) Protein structure evaluation of Modeller9.17 generated model using the suggested templates from Phyre2 by Ramachandran plot^{22,23}.

The MCT8 generated model from MODELLERv.9.17^{18,19,20,21} using the RaptorX¹⁷ suggested templates had 90.2% residues in the most favored region, 6.3 % in the additional allowed region, 2.7 % in the generously allowed region and 0.7 % in the disallowed region. The Ramachandran plot is shown in Fig. 7B.

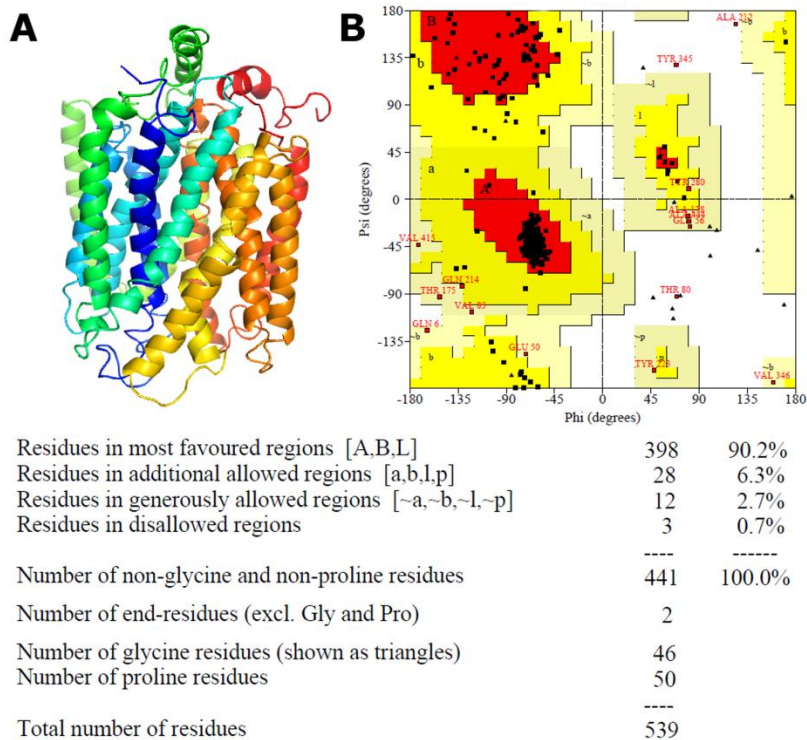


Fig. 7:- (A) Modeller9.17^{18,19,20,21} generated model of MCT8 using the suggested templates from Raptorx¹⁷ (B) Protein structure evaluation of Modeller9.17 model using the suggested templates from Raptorx by Ramachandran plot^{22, 23}.

The structure evaluation by Ramachandran plot^{22, 23} showed that MCT8 generated model from MODELLERv.9.17^{18,19,20,21} based on the RaptorX¹⁷ suggested templates (1pw4A, 4u4tA, 4ikxA, 4j05A, and 4gbyA) can be considered as the most reliable 3D protein structure for MCT8 compared with the other models.

Refinement of the best MCT8 model:-

MODELLERv9.17^{18,19,20,21} generated model for MCT8 using the suggested templates (1pw4A, 4u4tA, 4ikxA, 4j05A and 4gbyA) from RaptorX¹⁷ was considered for further refinement through ModRefiner server²⁵ to gain a better quality structure. An increase of about 2.1% residues in the favored region was seen and other parameters acquired better acceptable value. The refined model of MCT8 is shown in Fig. 8A. The Ramachandran plot of the refined model is shown in Fig. 8B.

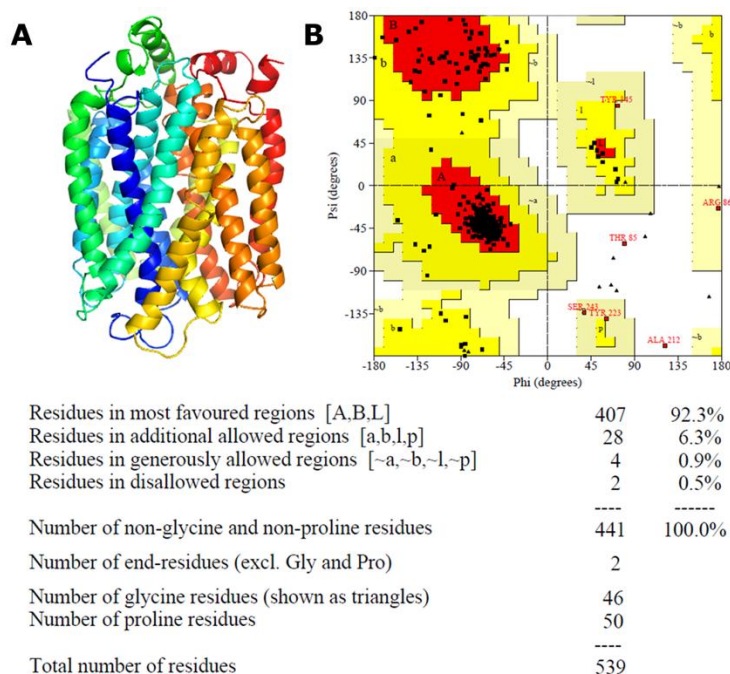


Fig. 8:- (A) Refined model of MCT8 predicted by Modrefiner²⁵. (B) Protein structure evaluation of the refined model by Ramachandran plot^{22,23}.

Conclusion:-

This study predicted a good quality 3D model of MCT8 using multi-template homology modeling. This model may have the potential to be used for computer aided active site prediction and lead discovery for the development of an effective drug against MCT8 deficiency disease.

References:-

- Dumitrescu AM, Liao XH, Best TB, Brockmann K, Refetoff S: A novel syndrome combining thyroid and neurological abnormalities is associated with mutations in a monocarboxylate transporter gene. *American Journal of Human Genetics* 2004; 74:168-175.
- Schwartz CE, May MM, Carpenter NJ, Rogers RC, Martin J, Bialer MG, Ward J, Sanabria J, Marsa S, Lewis JA, Echeverri R: Allan-Herndon-Dudley syndrome and the monocarboxylate transporter 8 (MCT8) gene. *The American Journal of Human Genetics* 2005; 77: 41:53.
- Friesema EC, Kuiper GG, Jansen J, Visser TJ, Kester MH: Thyroid hormone transport by the human monocarboxylate transporter 8 and its rate-limiting role in intracellular metabolism. *Molecular Endocrinology* 2006; 20:2761-72.
- Heuer H, Maier MK, Iden S, Mittag J, Friesema EC, Visser TJ, Bauer K: The monocarboxylate transporter 8 linked to human psychomotor retardation is highly expressed in thyroid hormone-sensitive neuron populations. *Endocrinology* 2005; 146:1701-6.
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourn PE: The Protein Data Bank. *Nucleic Acids Research* 2000; 28: 235-242.
- Pramanik S, Kutzner A, Heese K: Lead discovery and in silico 3D structure modeling of tumorigenic FAM72A (p17). *Tumor Biology* 2015; 36:239-249.
- Ou-Yang SS, Lu JY, Kong XQ, Liang ZJ, Luo C, Jiang H: Computational drug discovery. *Acta Pharmacologica Sinica* 2012; 33:1131-1140.
- Vyas VK, Ukawala RD, Ghate M, Chintha C: Homology modeling a fast tool for drug discovery: current perspectives. *Indian journal of pharmaceutical sciences* 2012; 74:1-17.
- Edgar R, Domrachev M, Lash AE: Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic acids research* 2002; 30:207-210.
- Gasteiger E, Hoogland C, Gattiker A, Duvaud SE, Wilkins MR, Appel RD, Bairoch A: Protein identification and analysis tools on the ExPASy server. *Humana Press*, 2005:571-607.

11. Geourjon C, Deleage G: SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Computer applications in the biosciences: CABIOS* 1995; 11:681-684.
12. Dosztányi Z, Csizmok V, Tompa P, Simon I: IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 2005; 21:3433-3434.
13. Zhang Y: I-TASSER server for protein 3D structure prediction. *BMC bioinformatics* 2008; 9:40.
14. Wu S, Zhang Y: LOMETS: a local meta-threading-server for protein structure prediction. *Nucleic acids research* 2007; 35:3375-82.
15. Yang J, Roy A, Zhang Y: BioLiP: a semi-manually curated database for biologically relevant ligand-protein interactions. *Nucleic acids research* 2013; 41:D1096-103.
16. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ: The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols* 2015; 10:845-858.
17. Kallberg M, Wang H, Wang S, Peng J, Wang Z, Lu H, Xu J: Template-based protein structure modeling using the RaptorX web server. *Nature protocols* 2012; 7:1511-1522.
18. Sali A, Blundell TL: Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 1993; 234:779-815.
19. Šali A, Potterton L, Yuan F, van Vlijmen H, Karplus M: Evaluation of comparative protein modeling by MODELLER. *Proteins: Structure, Function, and Bioinformatics* 1995; 23:318-326.
20. Marti-Renom MA, Stuart A, Fiser A, Sánchez R, Melo F, Sali A: Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* 2000; 29:291-325.
21. Webb B, Sali A: Comparative Protein Structure Modeling Using Modeller. *Current Protocols in Bioinformatics*. John Wiley & Sons, Inc. 2014; 5.6.1-5.6.32.
22. Laskowski RA, MacArthur MW, Moss DS, Thornton JM: PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 1993; 26:283-91.
23. Ramachandran GN, Ramakrishnan C, Sasisekharan V: Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* 1963; 7:95-99.
24. Arnold K, Bordoli L, Kopp J, Schwede T: The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* 2006; 22:195-201.
25. Dong X, Yang Z: Improving the Physical Realism and Structural Accuracy of Protein Models by a Two-step Atomic-level Energy Minimization. *Biophysical Journal* 2011; 101:2525-2534.