RESEARCH ARTICLE

MULTI-TEMPLATE HOMOLOGY MODELING OF HUMAN MCT8 PROTEIN.

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

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.

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
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 modeling.

Materials and Methods:-
MCT8 protein sequence:-

Primary structure prediction:-
The physico-chemical properties of the human MCT8 protein were calculated by using the ExPasy’s ProtParam tool. 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) 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 IUPred. 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 form.

3D structure prediction of MCT8 using I-TASSER:-
I-TASSER server http://zhanglab.ccmb.med.umich.edu/I-TASSER/ 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 PDB by multiple threading approach LOMETS (Local Meta-Threading-Server), 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 BioLiP. LOMETS is an on-line web service for protein structure prediction. It generates 3D models by collecting high-scoring target-template alignments from 9 locally-installed threading programs (FFAS-3D, HHsearch, MUSTER, pGenTHREADER, PPAS, PRC, ROSPECT2, SP3, and SPARKS-X). BioLiP is a semi-manually curated database for high-quality, biologically relevant ligand-protein binding interactions. The structure data are collected primarily from the PDB, with biological insights mined from literature and other specific databases. BioLiP 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 similarity.

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 sequence.

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 sequence.

3D structure prediction of MCT8 using MODELLER v9.17:-
The suggested templates from the I-TASSER, Phyre2 and RaptorX were subjected to comparative modeling using the advanced modeling feature of the MODELLER v9.17. 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.

Protein structure evaluation:-
The stereo chemical features of the predicted MCT8 models were evaluated with the program PROCHECK by Ramachandran plot analysis which was done through “Protein structure and model assessment tools” of SWISS-MODEL workspace. The best model was selected by Ramachandran plot analysis.

Structure refinement:-
The protein structure refinement of MCT8 best model was predicted by ModRefiner. 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 model.

Results and Discussion:-
Primary and secondary structure prediction:-
The physico-chemical properties of the human MCT8 protein by using the ExPasy’s ProtParam tool are shown in Table1. The secondary structural features of the MCT8 protein predicted by SOPMA 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.

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

**Table 2:** Secondary structure prediction of human Monocarboxylate Transporter 8 (MCT8) by using the SOPMA tool.

<table>
<thead>
<tr>
<th>Secondary structure</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha helix</td>
<td>37.85%</td>
</tr>
<tr>
<td>Beta sheet</td>
<td>17.25%</td>
</tr>
<tr>
<td>Turn</td>
<td>8.91%</td>
</tr>
<tr>
<td>Coil</td>
<td>35.99%</td>
</tr>
</tbody>
</table>

**Disorder prediction:**

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

**3D structure prediction using I-TASSER:**

I-TASSER 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 5f7dA, 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 were 1pw4A, 4ppA, 4kvA, 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 MODELLER v9.17 analysis.

**Fig. 2:** (A) 3D model of MCT8 protein predicted by I-TASSER. (B) Protein structure evaluation of I-TASSER model by Ramachandran plot.

**3D structure prediction using Phyre2:**
Phyre2 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 MODELLER v9.17 analysis.
3D structure prediction using RaptorX:-

446 residues (82%) of the MCT8 protein were modelled with RaptorX 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 analysis.
3D structure prediction using MODELLER v9.17:

The suggested templates from the I-TASSER, Phyre2 and RaptorX were subjected to comparative modeling using the advanced modeling feature of the MODELLER v9.17. The MODELLER v9.17 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 MODELLER v9.17 for MCT8 using the suggested templates from the I-TASSER, Phyre2 and RaptorX are shown in Fig. 5A, 6A and 7A respectively.

Protein structure evaluation:

The stereo chemical features of the predicted MCT8 models using MODELLER v9.17 were evaluated with the program PROCHECK by Ramachandran plot analysis. In the Ramachandran plot, 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 MODELLER v9.17 using the I-TASSER 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.
The MCT8 generated model from MODELLERv.9.17 \cite{18,19,20,21} using the Phyre2 \cite{16} 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) Modellerv9.17\textsuperscript{18,19,20,21} generated model of MCT8 using the suggested templates from Phyre2\textsuperscript{16} (B) Protein structure evaluation of Modellerv9.17 generated model using the suggested templates from Phyre2 by Ramachandran plot\textsuperscript{22,23}.

The MCT8 generated model from MODELLERv9.17\textsuperscript{18,19,20,21} using the RaptorX\textsuperscript{17} 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) Modellerv9.17 generated model of MCT8 using the suggested templates from Raptorx. (B) Protein structure evaluation of Modellerv9.17 model using the suggested templates from Raptorx by Ramachandran plot.

The structure evaluation by Ramachandran plot showed that MCT8 generated model from MODELLERv.9.17 based on the RaptorX 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 generated model for MCT8 using the suggested templates (1pw4A, 4u4tA, 4ikxA, 4j05A and 4gbyA) from RaptorX was considered for further refinement through ModRefiner server 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 Modrefiner\textsuperscript{25}. (B) Protein structure evaluation of the refined model by Ramachandran plot\textsuperscript{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:**