université PARIS DIDEROT PARIS 7

Teaching Unit: Structural Bioinformatics

M2 BI Nour Touibi

2017-2018

Plan:

- Introduction
- Residues conservation detection
- Transmembranes segments detection
- Disordered protein binding residue predictions
- Models generation:
    - The Hhpred models (Alva et al. , 2016)
    - The Phyre2 model
    - The Muster models
- Models evaluation and comparison
- Conclusion
- Bibliography

1- Introduction

The SLC16 gene family has fourteen member (Halestrap A., 2013). Solute carrier family 16 member 10 (SCL16A10) is a member of a family of plasma membrane amino acid transporters that mediate the sodium independent transport of aromatic amino acids across the plasma membrane (NCBI, 12-Nov-2017). Human SLC16A10 was localized on human chromosome 6, mapped to 6q21–q22 (Mungall A. J., 2003). It encode monocarboxylate transporters 10 (MCT10) (Halestrap A., 2013), a 515-amino-acid protein with 12 putative membrane-spanning domains.

Monocarboxylate transporter 10 (MCT10) belongs to the major facilitator superfamily (MFS). MFS is one of the largest groups of secondary active transporters conserved from bacteria to humans (Yan N., 2013). MFS proteins selectively transport a wide spectrum of substrates across biomembranes and play a pivotal role in multiple physiological processes (Yan N., 2013).

MCT10 is a Sodium-independent transporter that mediates the uptake of aromatic acids. MCT10 can function as a net efflux pathway for aromatic amino acids in the basosolateral epithelial cells (Kim DK., 2002). It is strongly expressed in kidney and skeletal muscle and at lower level in placenta and heart (Kim D. K., 2002). MCT10 is at least as active a thyroid hormone transporter as MCT8, and that both transporters facilitate iodothyronine uptake as well as efflux (Edith C. H., 2007). It was found a strong neuronal MCT10 immunocytochemical staining in a number of hypothalamic nuclei, including the PVN, IFN, and supraoptic nucleus (Friesema C. H., 2011). Intense staining was also observed in neurons of the lateral hypothalamus including the perifornical area (Friesema C. H., 2011). The overlap between neuronal distribution patterns of TR and MCT10 immunoreactivity suggests a clear role for this transporter in thyroid hormone signaling in the human hypothalamus (Alkemade A., 2005). Also, the strong expression of MCT10 and OATP1C1 in the human hypothalamus indicates a possible role in the regulation of the hypothalamus-pituitary-thyroid axis (Friesema C. H., 2011).

As mentioned above, MCT10 is a 515-amino-acid membrane protein with twelve putative membrane-spanning domains (Kim D. K., 2002 and Halestrap A., 2013). The aim of this study is the modelization of this protein. As a first step, I studied as well as possible the protein sequence. Then I searched for PDB templates. But, the highest homologous sequence identity was 14% with a very smell coverage. So threading models was generated.

In fact, protein threading is a method of protein modeling which is used to model those proteins which have the same fold as proteins of known structures, but do not have homologous proteins with known structure. It first identifies structural templates from the PDB, then thread each amino acid in the target sequence to a position in the template structure, and evaluate how well the target fits the template

In this study, first of all PSIPRED and TMHMM are used to a simple and accurate secondary structure prediction. Then, three server are used to predict the MCT10 3D structure by threading. This three server are HHpred, Phyre2 and Muster. The GalaxyRefine Web is used to refine loops. Models were then evaluate with ModEval to compare their accuracy.

2- Residues conservation detection:

To examine the conservation of residues and characterize the important amino acids for the function, I blast (blastp option) my protein sequence in the NCBI. I choose similar proteins from the same family for the multiple alignment (Fig. 1).

The blastp result show that conserved amino acids are:

	180-354		357-360		363-369		371-574	

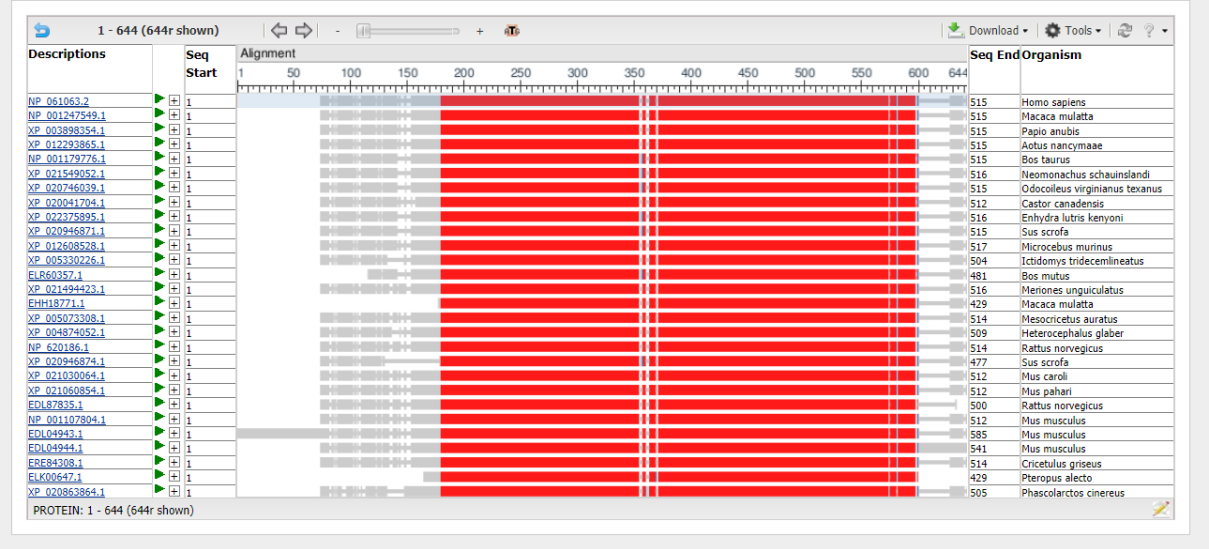

Fig. 1: Multiple alignment of similar proteins selected by blastp

3- Transmembranes segments detection:

As mentioned in the literature, Psipred and TMHMM predict twelve transmembranes segments shown in the figure 2 and 3 with their location in the protein sequence.

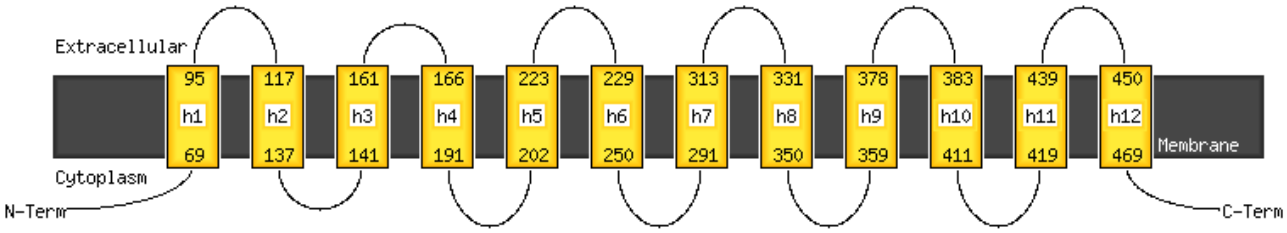

Fig. 2: Psipred result for number and location of the transmembranes segments

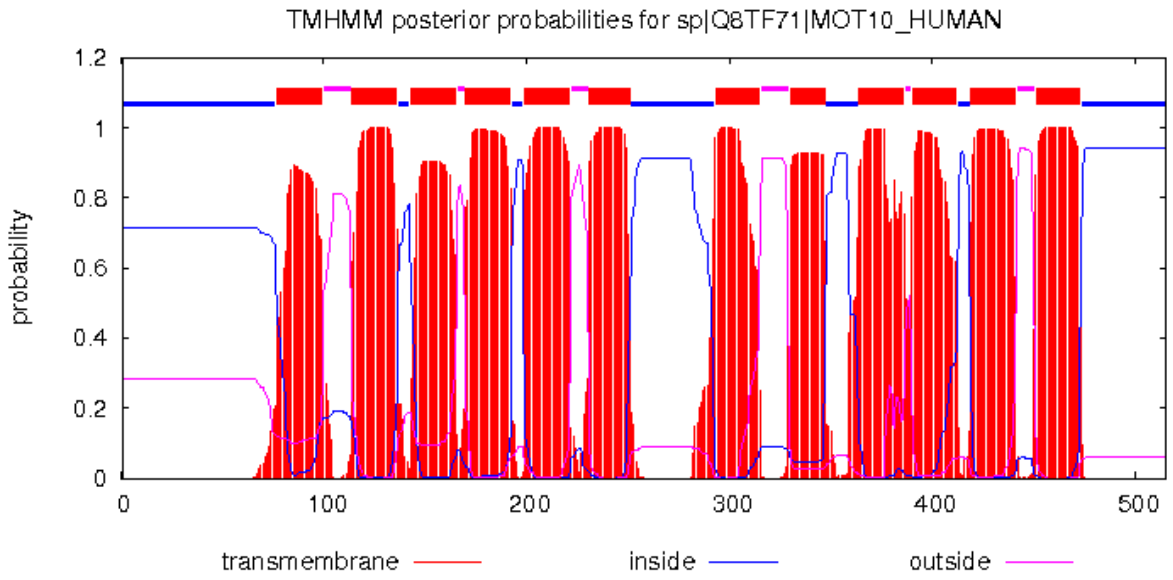

Fig. 3: TMHMM result for number and location of the transmembranes segments

4- Disordered protein binding residue predictions:

Disordered protein binding residue are residues lacks a fixed or ordered three-dimensional structure. The prediction of disordered protein binding residue is important to evaluate the future 3D structure. In our case, the disordered protein binding residues was predict with Psipred. They were localized in N- and C termini and between amino acids in 200 and 300 position (fig. 4).

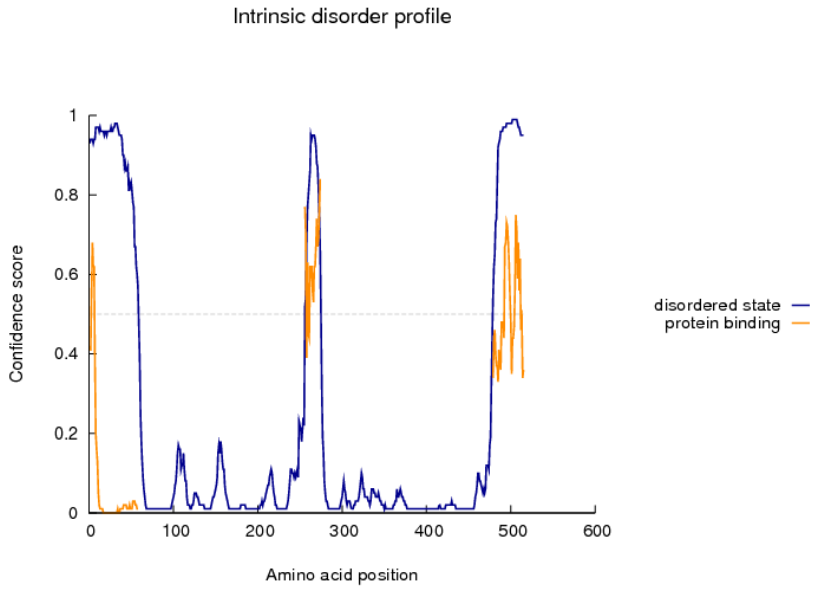

Fig. 4: Amino acids in the input sequence are considered disordered when the blue line is above the grey dashed line, that is the confidence score is higher than 0.5. The orange line shows the confidence of disordered protein binding residue predictions.

5- Models generation:

5.1- The Hhpred models (Alva et al. , 2016):

Hhpred is a free protein structure prediction server. Two models are selected from Hhpred (Table 1). The two models are refined with galaxyRaffine Web (fig.5 and fig. 6). The first hit (1PW4_A) was selected because it is the common top hit selected with the three server. The second hit (4ZW9_A) was selected because he has the most similar target length to the query length. The two models obtained with this method have twelve putative membrane-spanning domains as predicted.

Table 1: Hits selected with Hhpred

Hit	*Name*	*Probability*	*E-value*	*SS*	*Cols*	*Target Length*
1PW4_A	Glycerol-3-phosphate transporter; Glycerol-3-Phosphate, Transmembrane, Inner membrane, Transporter; 3.3A {Escherichia coli} SCOP: f.38.1.1	100	9.9e-29	45.8	431	451
4ZW9_A	Solute carrier family 2, facilitated; transporter, TRANSPORT PROTEIN; HET: BGC, GLC, OLC; 1.502A {Homo sapiens}	99.95	5.6e-29	33.9	463	518

Fig. 5: 1PW4_A model obtained with Hhpred and visualized with Pymol. Conserved helix are colored in blue and conserved loops are colored in purple.

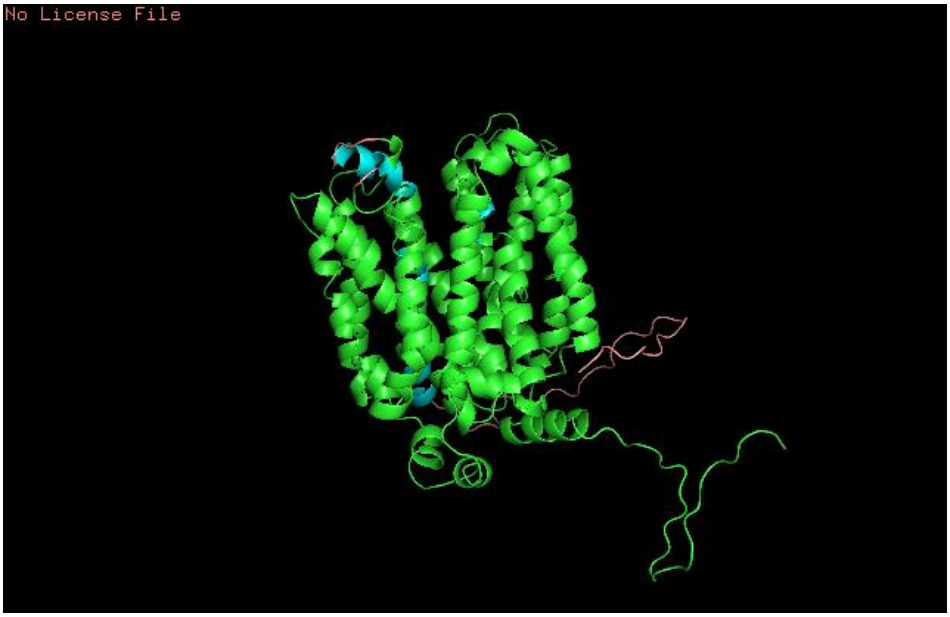

Fig. 6: 4ZW9_A model obtained with Hhpred and visualized with Pymol. Conserved helix and loops are colored in green.

5.2- The Phyre2 model:

The top model based on template selected with Phyre2 is 1PW4_A from the Superfamily MFS of general substrate transporter which is member of Glycerol-3-phosphate transporter family. This model (Fig. 8) is a 100% confident model with 83% coverage (426 residues have been modelled with 100.0% confidence by the single highest scoring template).

Transmembrane helices have been predicted to adopt the topology shown below (Fig. 7). This prediction confirm the last one selected with Psipred.

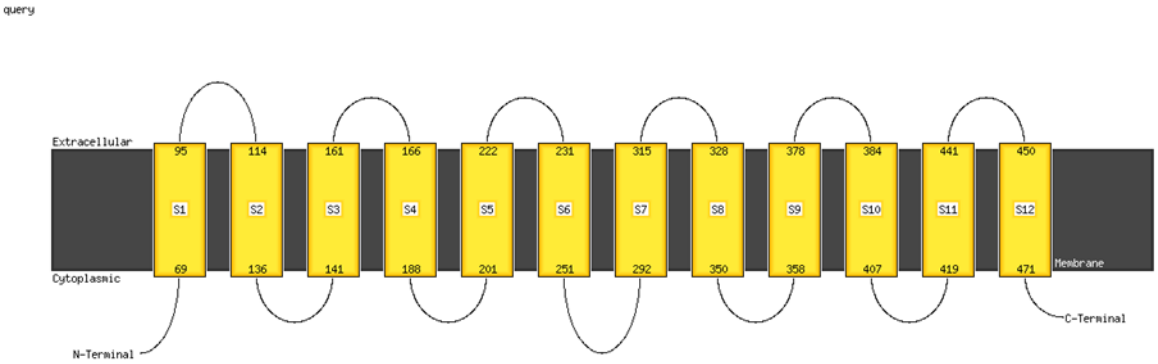

Fig. 7: Phyre2 result for number and location of the transmembranes segments

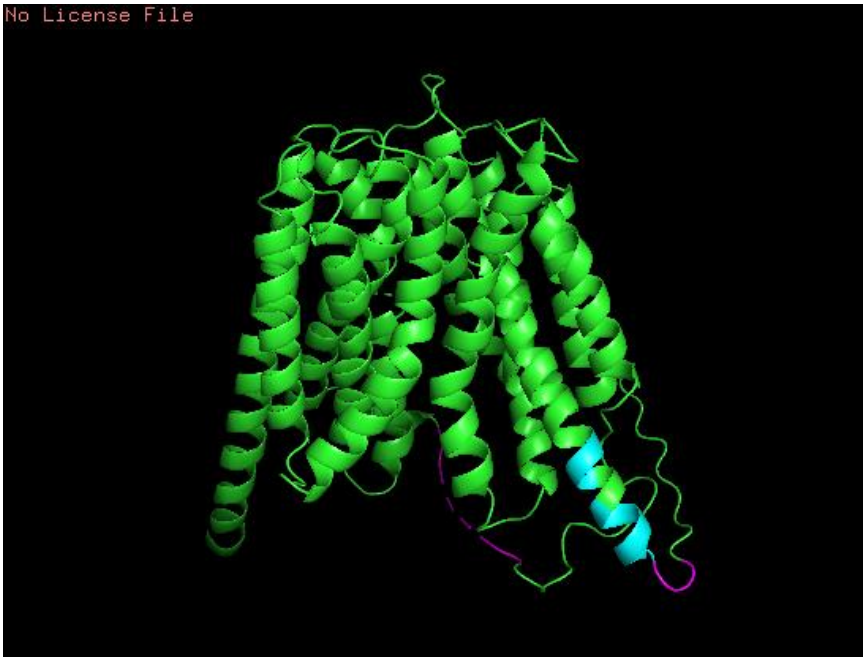

Fig. 8: 1PW9_A model obtained with Phyre2 and visualized with Pymol. Conserved helix and loops are colored in green.

5.3- The Muster models :

Muster generate 10 templates how are evaluate by according a type (good/bad) (Table 2). The second good template is our common template 1PW4_A. I visualize the first and the second template (Figs. 8 and 9) with Pymol to compare the 3D structure.

Table 2: Muster results

Rank	Template	Align_length	Coverage	Zscore	Seq_id	Type	Target-template alignments	3-D models from threading alignments	Full-length models by MODELLER
1	3wdoA	429	0.833	9.831	0.110	Good	alignment_1	threading_1	model_1
2	1pw4A	423	0.821	9.811	0.111	Good	alignment_2	threading_2	model_2
3	4zowA	386	0.749	8.459	0.093	Good	alignment_3	threading_3	model_3
4	4j05A	399	0.774	8.172	0.098	Good	alignment_4	threading_4	model_4
5	4iu8A	385	0.747	7.864	0.132	Good	alignment_5	threading_5	model_5
6	4ikvA	425	0.825	7.438	0.108	Bad	alignment_6	threading_6	model_6
7	4m64A	412	0.8	7.232	0.083	Bad	alignment_7	threading_7	model_7
8	4apsA	397	0.770	7.171	0.136	Bad	alignment_8	threading_8	model_8
9	4w6vA	419	0.813	7.155	0.138	Bad	alignment_9	threading_9	model_9
10	4q65A	400	0.776	7.135	0.110	Bad	alignment_10	threading_10	model_10

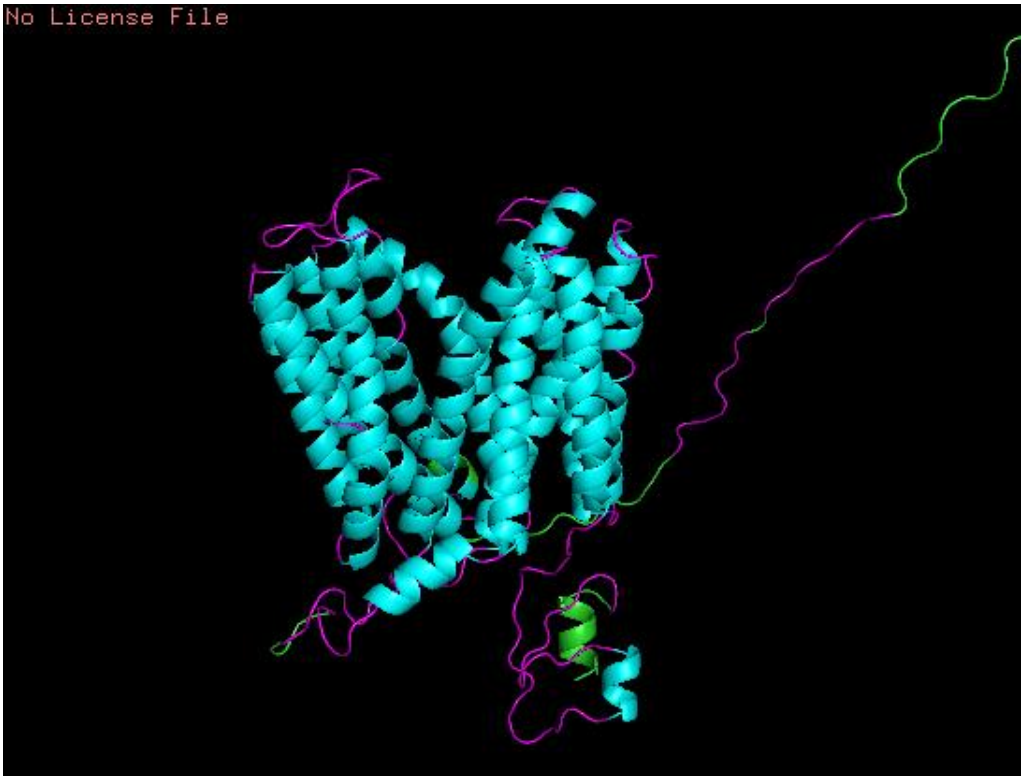

Fig. 8: 3WDO_A model obtained with Muster and visualized with Pymol. Conserved helix are colored in blue and conserved loops are colored in purple.

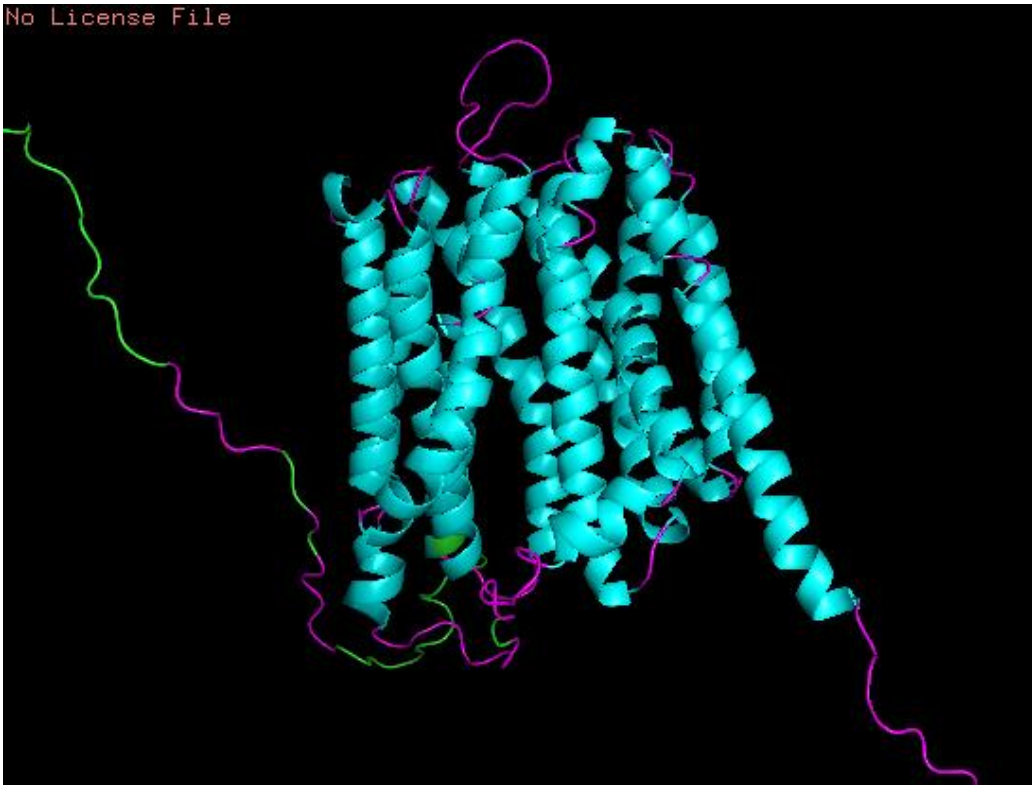

Fig. 9: 1PW4_A model obtained with Muster and visualized with Pymol. Conserved helix are colored in blue and conserved loops are colored in purple.

6- Models evaluation and comparison:

All selected model present twelve putative membrane-spanning domains as mentioned in the literature. Models from Hhpred and Phyre2 present missing amino acids from the end of the sequence. But, that match with the disordered protein binding residue obtained with Psipred.
Besides, the common model selected by the three servers is 1PW4_A. It is a protein membrane with twelve putative membrane-spanning domains and belongs to the MFS. 4ZW9_A is also a protein membrane with twelve putative membrane-spanning domains. So, models from different servers will be evaluate to select the best model (Table 3).
According to the literature, the best model has the lowest value of z-DOPE. However, the predicted root-mean-squared deviation (RMSD) shows that the models are not utile.
The DOPE profile (Figs. 10, 11, 12) can be used to compare the fit of the 3D model across the regions of high and low sequence conservation. The plotted DOPE score profiles (Figs. 10, 11, 12) shows how much the model is following the template.

Table 3: ModEval results

	Hhpred 1PW4_A model	*Phyre2 1PW4_A model*	*Muster 1PW4_A model*
TSVMod Results			
Predicted RMSD	10.311 Â	22.079 Â	16.175 Â
Predicted Native Overlap (3.5 Â)	0.086	0.066	0.088
Modeller Scoring Results			
Sequence Identity	30 %	30 %	30 %
z-DOPE	-0.697	0.996	-0.159

	Hhpred 4ZW9_A model	*Muster 3WDO_A model*
TSVMod Results		
Predicted RMSD	6.585 Â	14.238 Â
Predicted Native Overlap (3.5 Â)	0.526	0.192
Modeller Scoring Results		
Sequence Identity	30 %	30 %
z-DOPE	-0.990	-0.369

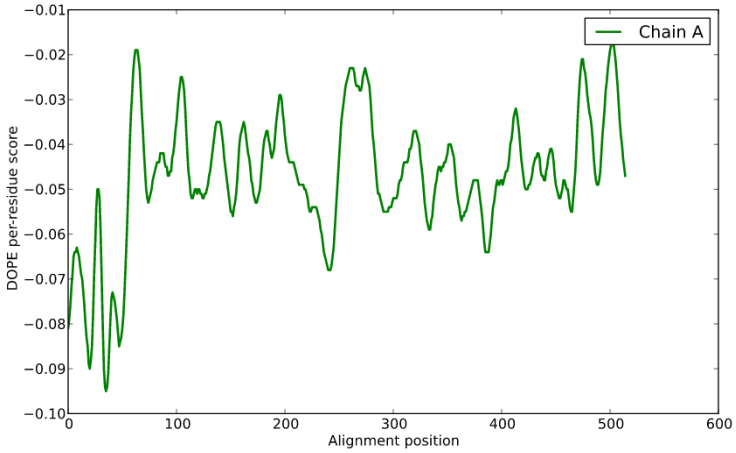

Fig. 10: Muster 3WDO_A model DOPE profile

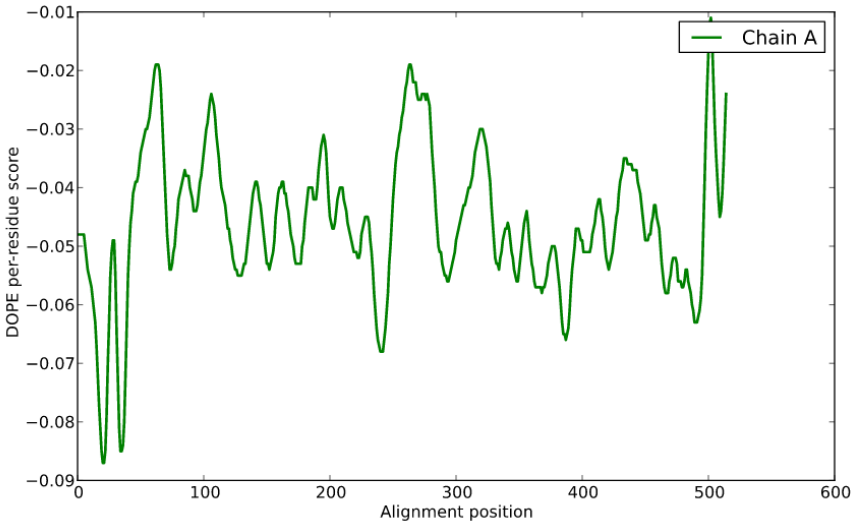

Fig. 11: Muster 1PW4_A model DOPE profile

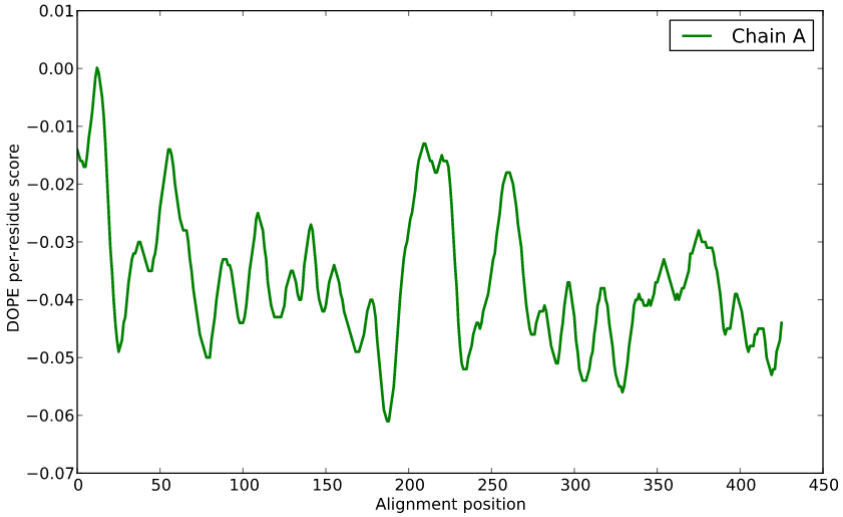

Fig. 12: Phyre2 1PW4_A model DOPE profile

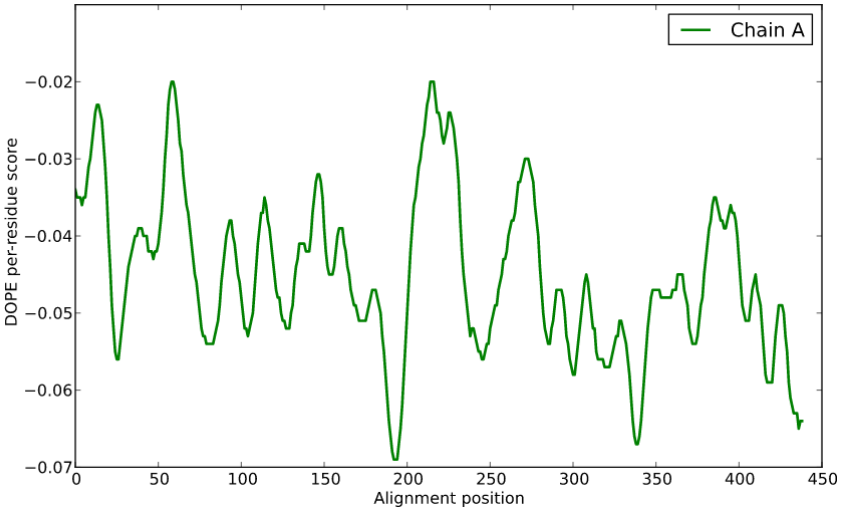

Fig. 13: Hhpred 1PW4_A model DOPE profile

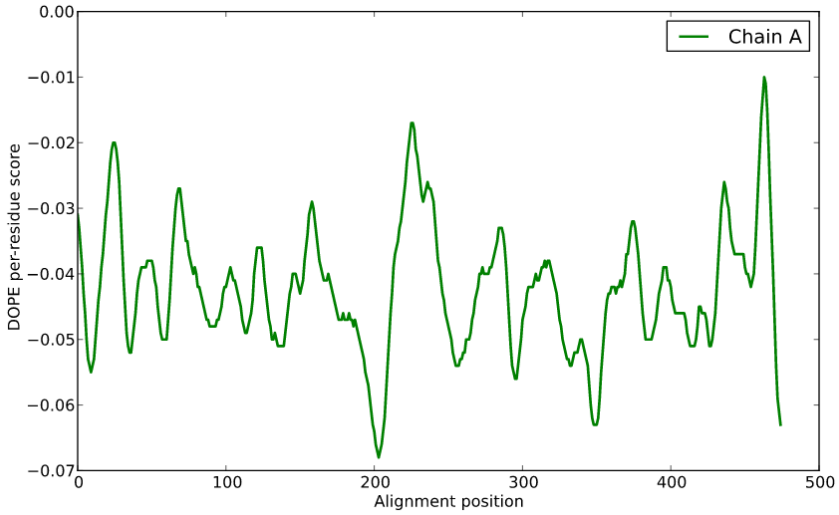

Fig. 14: Hhpred 4ZW9_A model DOPE profile

7- Conclusion:

According to the evaluation results, it was difficult to choose a model for my protein. Finally, based on biological characteristics such as protein membrane, the number of transmembranes domains and the different scores selected from the evaluation method, the most accurate model is Hhpred 4ZW9_A model (Fig. 6 bis). Unless, 1PW4 model was common to the three servers, I personally think that the 4ZW9_A model evaluation scores are significantly more accurate.

Fig. 6 bis: 4ZW9_A model obtained with Hhpred and visualized with Pymol. Conserved helix and loops are colored in green.

Bibliography

1- Nieng Yan. "Structural advances for the major facilitator superfamily (MFS) transporters". Biochemical Sciences, 2013, Vol. 38, No. 3.
2- A. J. Mungall and al. "The DNA sequence and analysis of human chromosome 6". Nature 2003.
3- Halestrap A.P. "The SLC16 gene family - structure, role and regulation in health and disease". Mol Aspects Med. 2013. doi: 10.1016/j.mam.2012.05.003.
4- Kim DK, Kanai Y, Matsuo H, Kim JY, Chairoungdua A, Kobayashi Y, Enomoto A, Cha SH, Goya T, Endou H. "The Human T-Type Amino Acid Transporter-1: Characterization, Gene Organization, and Chromosomal Location". Genomics Volume 79, Issue 1, January 2002, Pages 95-103 DOI: 10.1006/geno.2001.6678
5- Edith C. H. Friesema, Jurgen Jansen, Jan-willem Jachtenberg, W. Edward Visser, Monique H. A. Kester, and Theo J. Visser. "Effective Cellular Uptake and Efflux of Thyroid Hormone by Human Monocarboxylate Transporter 10". Molecular Endocrinology, doi: 10.1210/me.2007-0112
6- Edith C. H. Friesema, Andries Kalsbeek, Dick F. Swaab, Theo J. Visser, and Eric Fliers, J Clin. "Expression of Thyroid Hormone Transporters in the Human Hypothalamus". Endocrinol Metab, June 2011, 96(6):E967–E971
7- Alkemade A, Vuijst CL, Unmehopa UA, Bakker O, Vennstro¨m B, WiersingaWM, Swaab DF. "Fliers E2005 Thyroid hormone receptor expression in the human hypothalamus and anterior pituitary". J Clin Endocrinol Metab 90:904 –912)