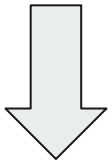
Prediction and modeling of the protein W0K4U9

EL HARTY Hager & MESSAK Imane

Université de Paris

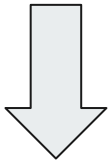
Problematic

A large number of proteins available on Uniprot with an unknown structure and function



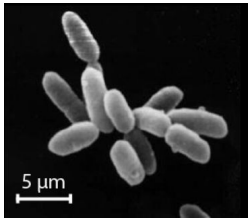
Biological function

Implicated in the development of a disease

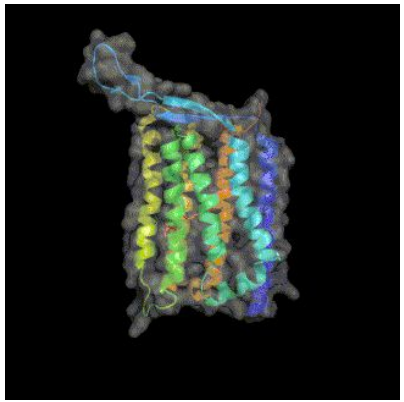


Predict using different tools the structure of the membrane protein

About our protein



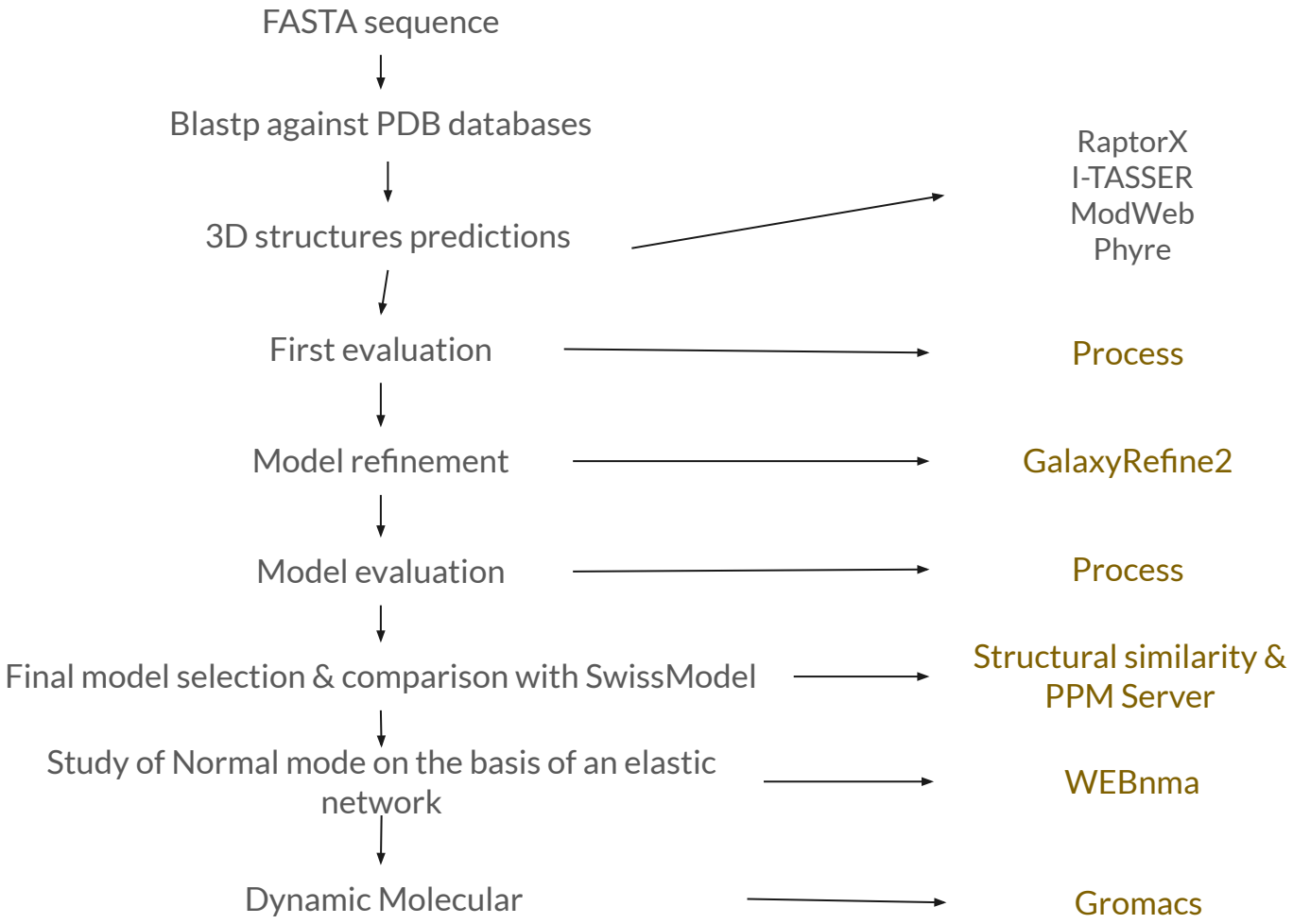
- ➔ Transmembrane protein
- ➔ Noted as "*unreviewed*" in Uniprot
- ➔ 268 amino acids
- ➔ Found in the organism *Halobacterium sp. DL1*
- ➔ Coded by the gene *HALDL1_13775*
- ➔ Belongs to the *Heliorhodopsin* family
- ➔ No known bibliography about this protein (function, expression, pathology..)



https://www.uniprot.org/uniprot/W0K4U9 https://en.wikipedia.org/wiki/Heliorhodopsin

Materials & Methods

Pipeline


FASTA sequence
Blastp against PDB databases
3D structures predictions
RaptorX
I-TASSER
ModWeb
Phyre
First evaluation
Process
Model refinement
GalaxyRefine2
Model evaluation
Process
Final model selection & comparison with SwissModel
Structural similarity &
PPM Server
Study of Normal mode on the basis of an elastic network
WEBnma
Dynamic Molecular
Gromacs

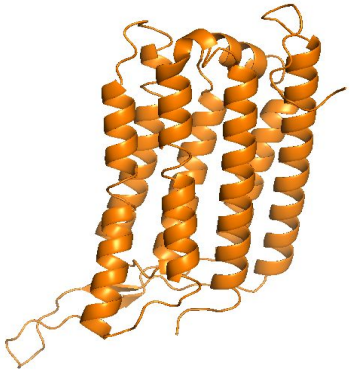
Results

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☑	Crystal structure of Thermoplasmatales archaeon heliorhodopsin [Thermoplasmatales archaeon SG8-52-1]	Thermoplasmatales arch…	258	258	94%	3e-86	51.38%	259	6IS6_A
☑	Crystal structure of Thermoplasmatales archaeon heliorhodopsin E108D mutant [Thermoplasmatales archaeo…	Thermoplasmatales arch…	257	257	94%	1e-85	50.99%	259	7CLJ_A
☑	Crystal structure of bacterial heliorhodopsin 48C12 [Actinobacteria bacterium]	Actinobacteria bacterium	170	170	91%	1e-51	41.70%	252	6UH3_A
☑	Crystal structure of the 48C12 heliorhodopsin in the violet form at pH 8.8 [Actinobacteria bacterium]	Actinobacteria bacterium	170	170	91%	2e-51	41.70%	264	6SU3_A

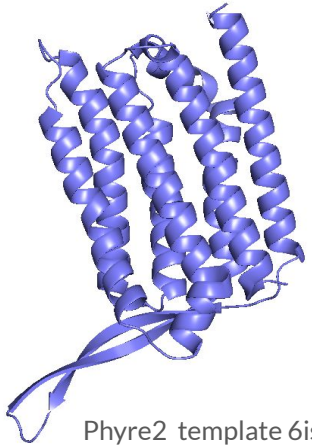
Results of the Blastp against PDB

https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome

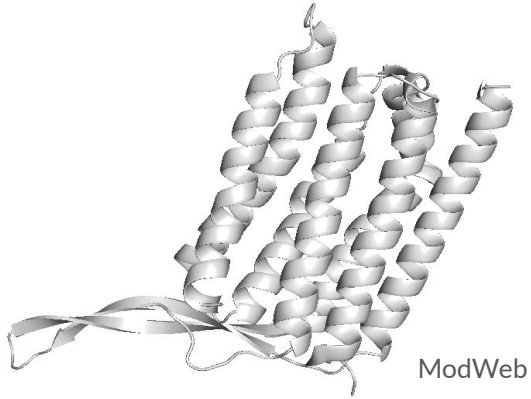
Some Resulting Models



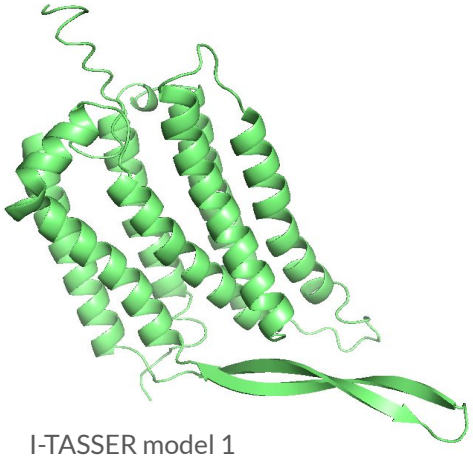
RaptorX Model1



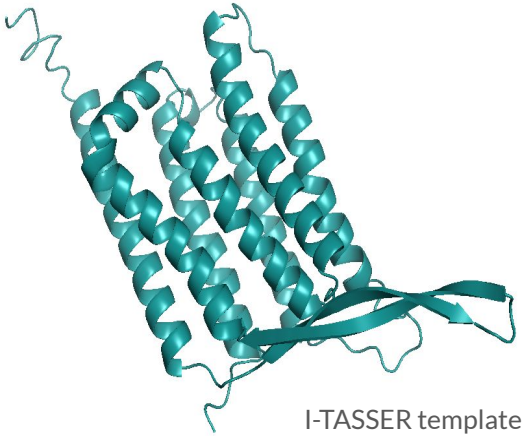
Phyre2 template 6is6



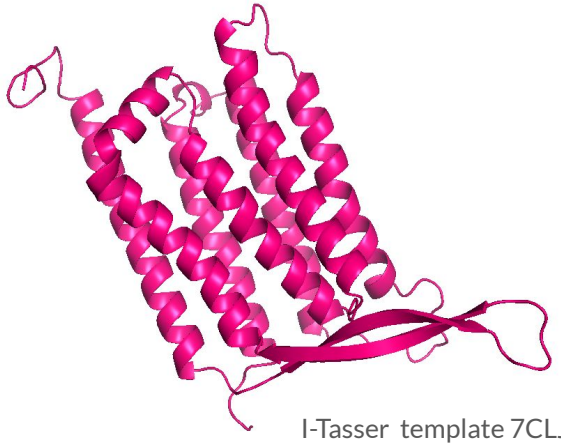
ModWeb template 6is6



I-TASSER model 1



I-TASSER template 6is6



I-Tasser template 7CLJA

Model Refinement

	Phyre2	ModWeb	RaptorX	I-Tasser
Number of models	2	3	1	6
Template	6is6 6uh3	6is6 6su3x 6uh3	-	6is6 3abw 6is6 *2 7clja*2

Table: Summary of our pre-selection of prediction models.

	Modèle initial	Modèle 1 raffiner
Overall Quality	4.5	5.5
Covalent Bond Quality	6.5	7.5
Non-Covalent/Packing Quality	3.5	4.5
Torsion Angle Quality	5.5	6.5

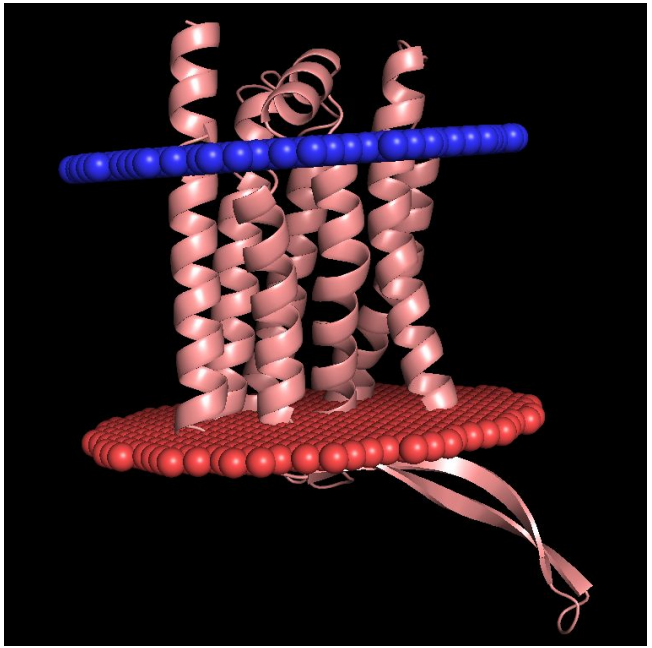
Table 1: Comparison table of the model 2 (6uh3_A) predicted by Pyre2

	Modèle initial	Modèle 1 raffiner
Overall Quality	4.5	5.5
Covalent Bond Quality	6.5	7.5
Non-Covalent/Packing Quality	3.5	4.5
Torsion Angle Quality	5.5	6.5

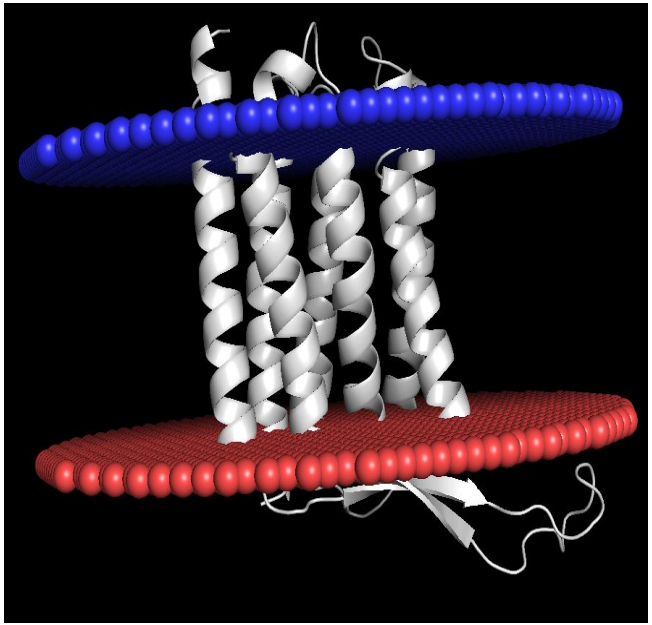
Table 2: Comparison table of the model 6is6A predicted by ModWeb

	Modèle initial	Modèle 1 raffiner
Overall Quality	3.5	5.5
Covalent Bond Quality	6.5	7.5
Non-Covalent/Packing Quality	3.5	4.5
Torsion Angle Quality	1.5	6.5

Table 3: Comparison table of the model 6is6_A predicted by I-TASSER

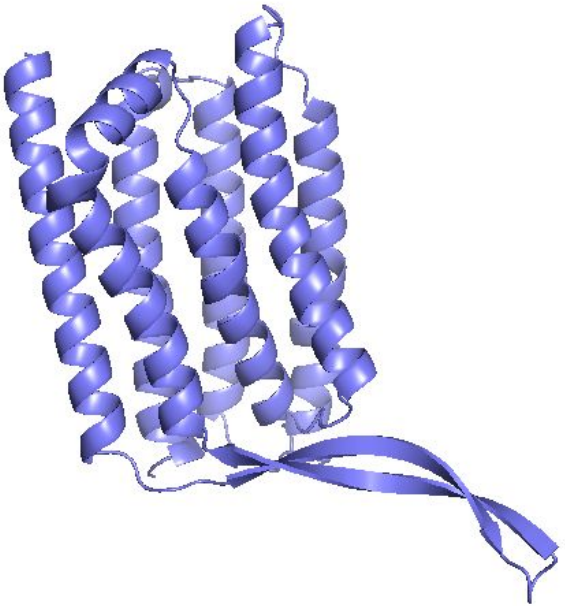


Visualisation of the model chosen by ModWeb



Visualisation of the model chosen by Pyre2

orientations of proteins in membranes (OPM) database



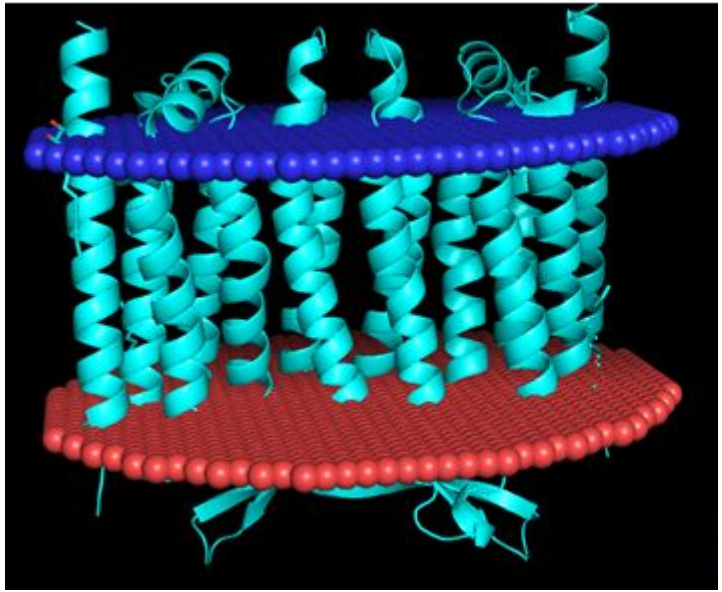
Structure the template 6is6_A done by ModWeb

- ModWeb does a comparative Modeling
- Uses the 6IS6_A template
- 248 residue modeled by ModWeb

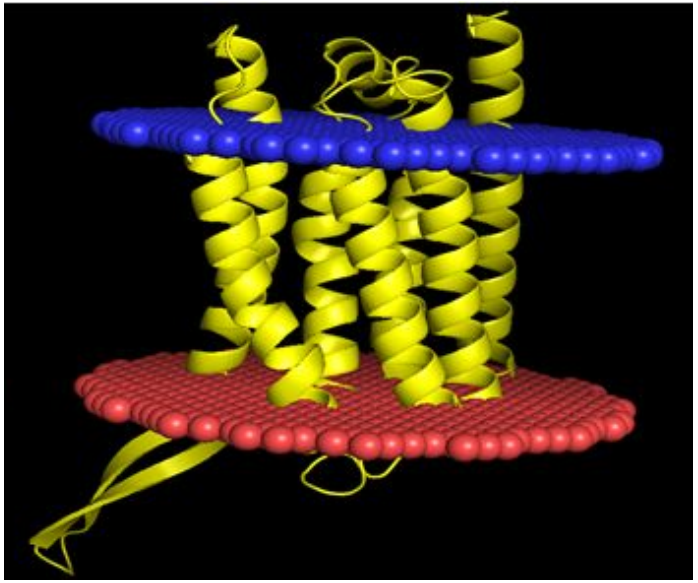
	Template PDB Code	PDB Segment	PDB Comment	CATH coverage or Link	Dataset	Seq Ident	E-value	Model Score	Quality	Modeled Segment
☐	6is6A	(4-252)		CATH	MW-ModWeb20201227_D2	52	0	1		13-266

Template chosen by ModWEb

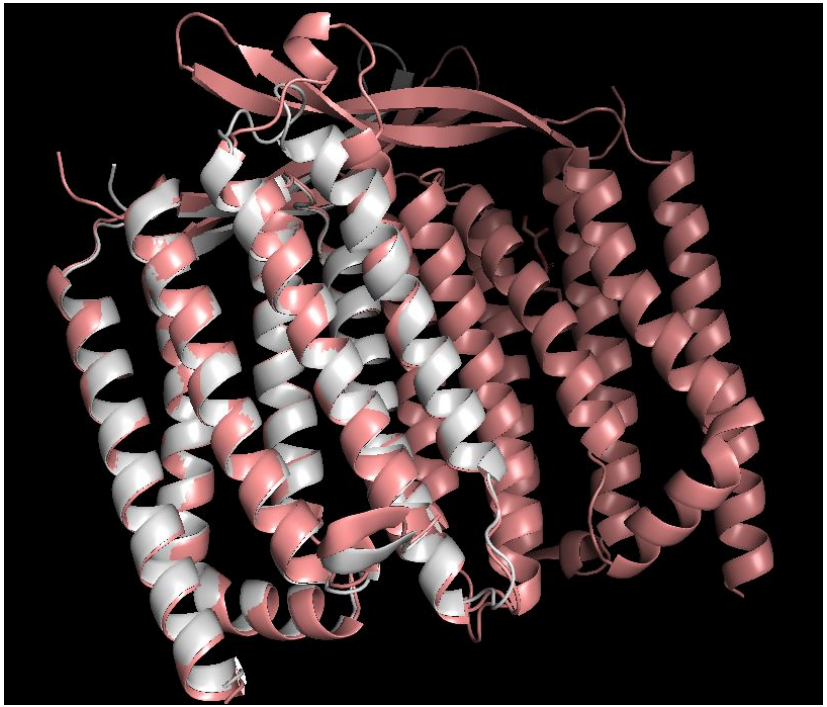
SwissModel VS Our model



Predicted with SwissModel using the 7CLJ template



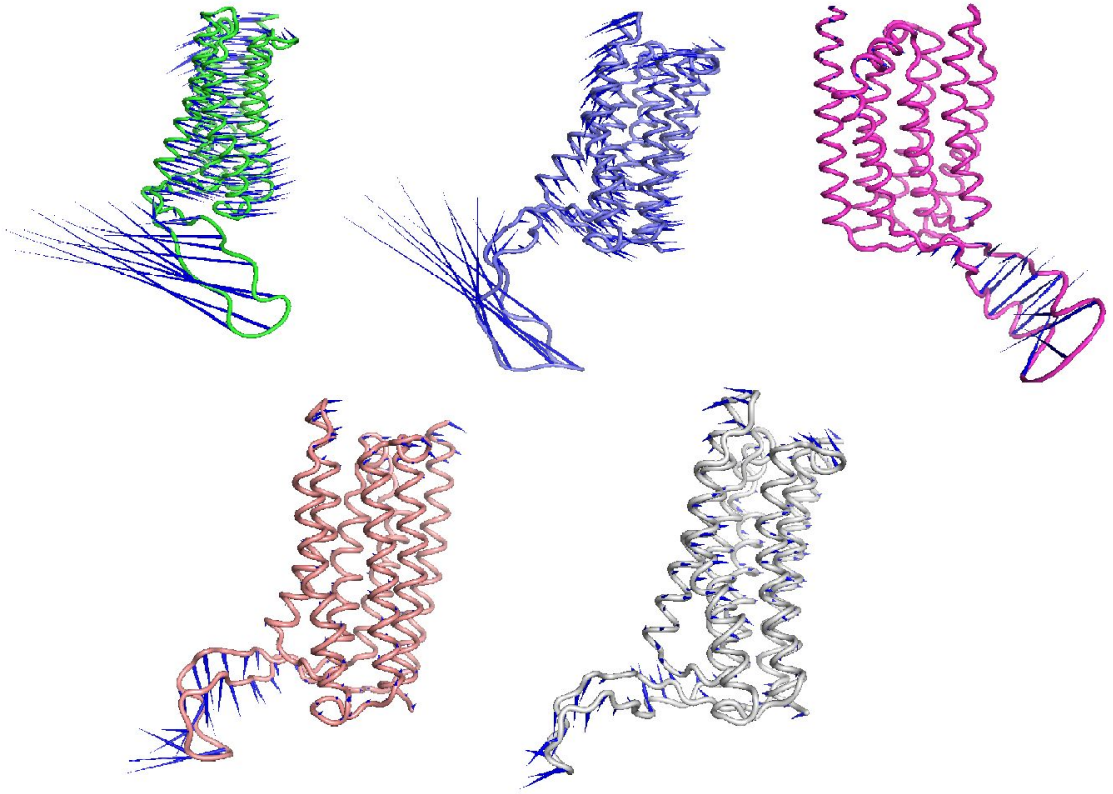
Predicted with ModWeb using the 6IS6 template



RMSD = 0.609 Å

Alignment of the predicted model by SwissModel and our model

Normal modes on the basis of an elastic network

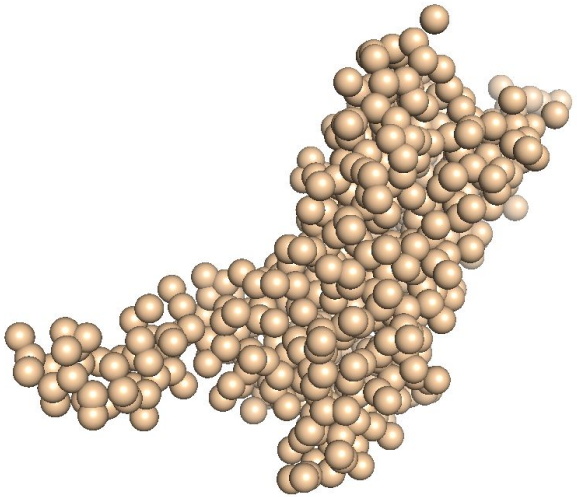


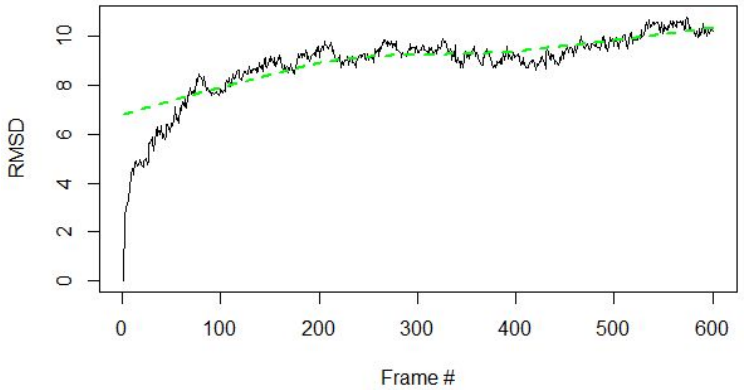
6 normal modes

Two different movements of our structure (7,8,9 and 11, 12)

Molecular Dynamics

- Coarse grain simulation with the Martini force field of our protein solved in water during 600 ps



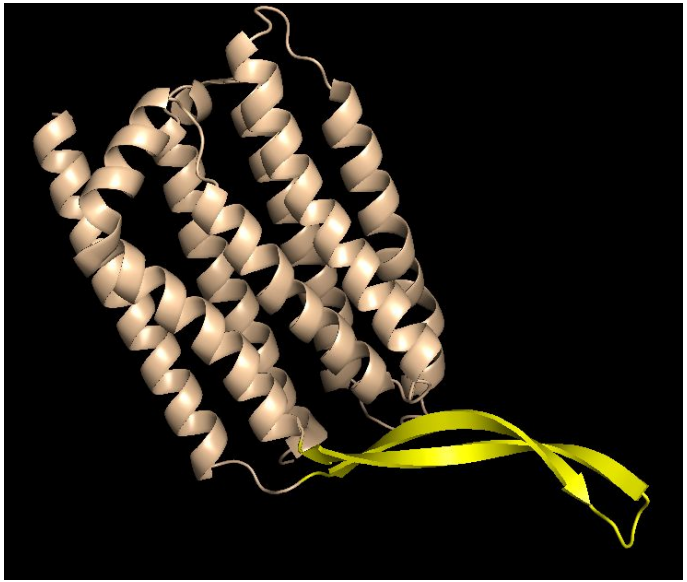


Evolution of the RMSD over time of our model

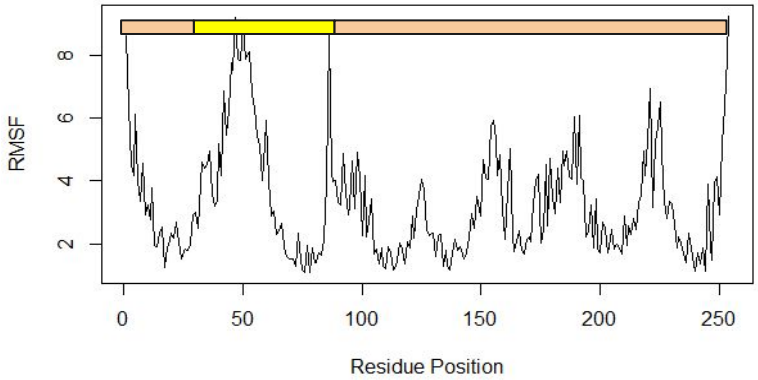
GROMACS FAST. FLEXIBLE. FREE.

Martini Coarse Grain Forcefield for Biomolecules

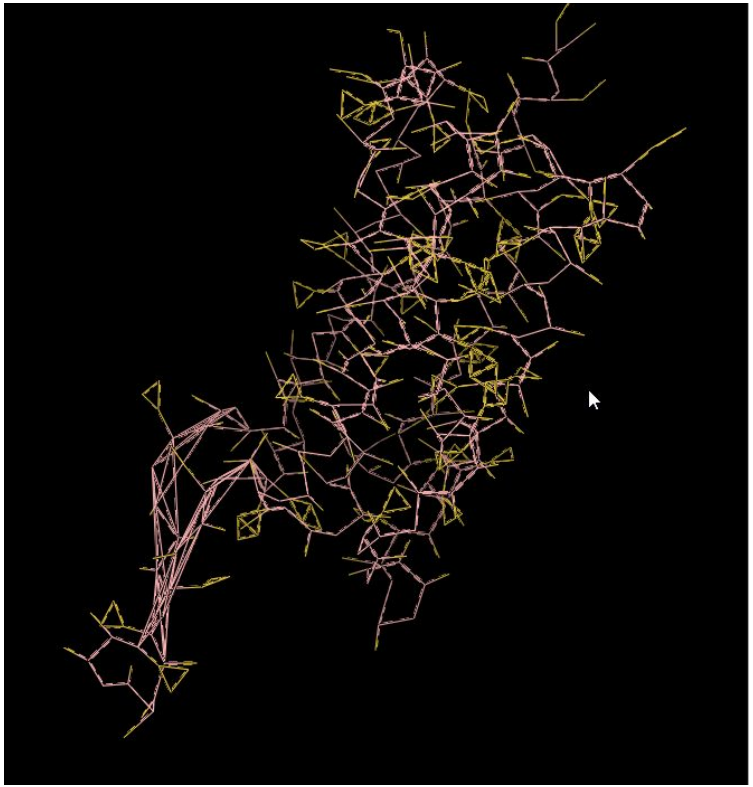
BIO3D



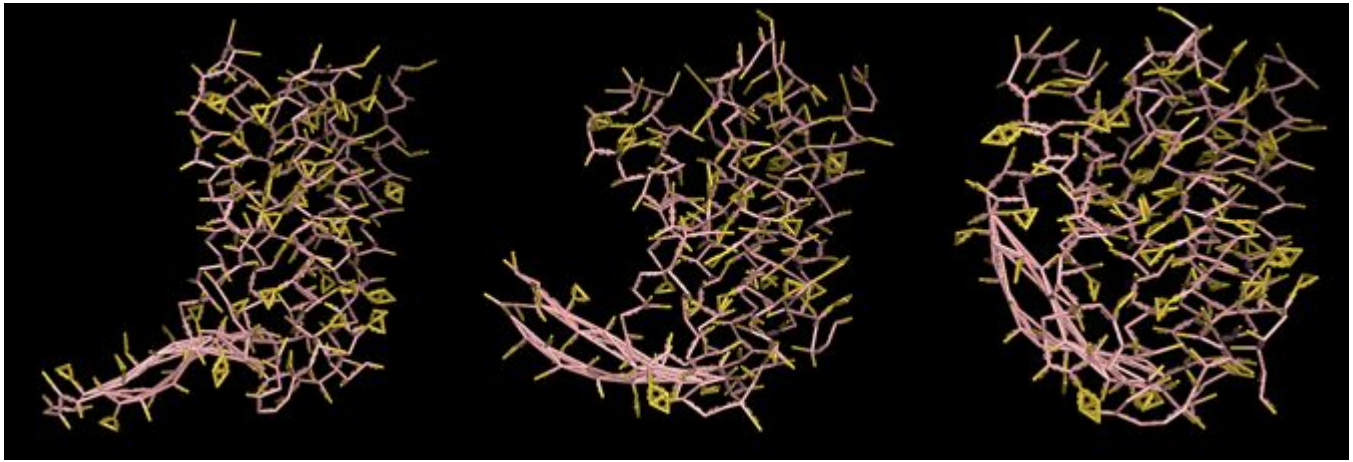
Structure of our protein with the beta sheet colored yellow



Evolution of the RMSF over time of our model



Different states of our protein by applying molecular dynamics



Different states of our protein by applying molecular dynamics

Conclusion

- We succeeded into Modelling our protein
- Predicting the 3D structure allowed us to observe the folded structure of the 6 alpha helices and the beta sheet of the transmembrane protein
- The study of conformational changes using normal modes, allowed us to identify different movements such as the folding of the beta sheet and its sweeping from left to right.
- The objective now would be to study the movements of the protein in its biological environment,
- In the future we should do the same study to the SwissModel to view the different movements, and compare them to those of our model.

Thank you for your attention

ANNEX: Materials

	Template search	SS prediction	Model evaluation
ModWeb (User -friendly interface Structure modeling by satisfaction of spatial restraints)	Yes	No	Identity score MPQS
RaptorX (User -friendly interface for protein structure and function prediction)	Alignement threading	No	P-value Score uGDT RMSD
I-TASSER (Integrated platform for automated protein structure and function prediction)	LOMETS consists of multiple threading algorithms	Yes	C_score TM_score RMSD Ident/Cover
Phyre2 (User -friendly interface to cutting-edge bioinformatics methods)	HMM-HMM matching	Yes	Identity