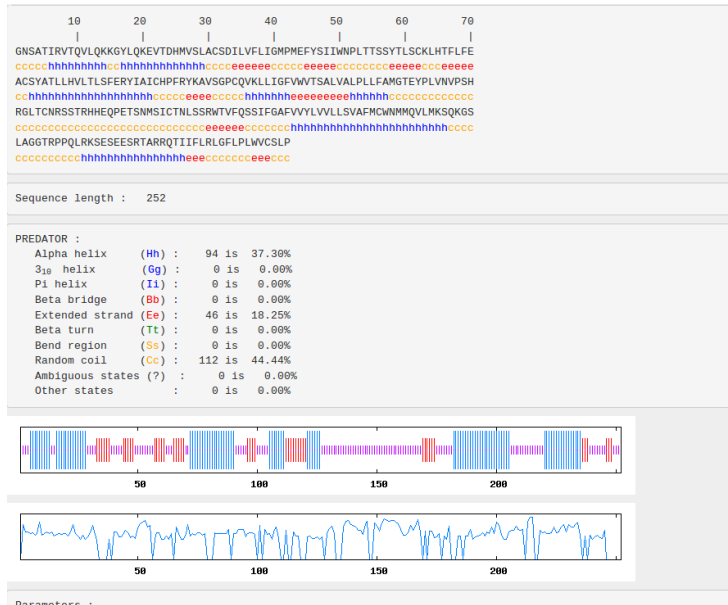


Lab: Secret sequence

1. What Pfam domain does this sequence you got match?

7 transmembrane receptor (rhodopsin familiy)

2. What is the secondary structure of your sequeffence ?



The secondary structure consists of 7 alpha helices, which is about 37% and random coil (44%) and extended strand (18%).

3. Identify possible interaction partners of your protein domain. What is the evidence for that ?

There are 12 interactions according to Pfam to the searched protein sequence:

- Cytochrome b562** Arnesano F, Banci L, Bertini I, Faraone-Mennella J, Rosato A, Barker PD, Fersht AR; , Biochemistry 1999;38:8657-8670.: The solution structure of oxidized Escherichia coli cytochrome b562
- G-protein alpha subunit** https://en.wikipedia.org/wiki/G_alpha_subunit
- Amino terminal of the G-protein receptor rhodopsin** Yeagle PL, Salloum A, Chopra A, Bhawsar N, Ali L, Kuzmanovski G, Alderfer JL, Albert AD; , J Pept Res. 2000;55:455-465.: Structures of the intradiskal loops and amino terminus of the G-protein receptor, rhodopsin
- Phage lysozyme** https://en.wikipedia.org/w/index.php?title=Glycoside_hydrolase_family_24
- CXCR4 Chemokine receptor N terminal** Veldkamp CT, Seibert C, Peterson FC, De la Cruz NB, Haugner JC 3rd, Basnet H, Sakmar TP, Volkman BF;; Sci Signal. 2008;1:ra4.: Structural basis of CXCR4 sulfotyrosine recognition by the chemokine SDF-1/CXCL12
- Signal peptide binding domain** Clemons WM Jr, Gowda K, Black SD, Zwieb C, Ramakrishnan V; , J Mol Biol 1999;292:697-705.: Crystal structure of the conserved subdomain of human protein SRP54M at 2.1 A resolution: evidence for the mechanism of signal peptide binding.

7. **WD40 repeat** https://en.wikipedia.org/w/index.php?title=WD40_repeat

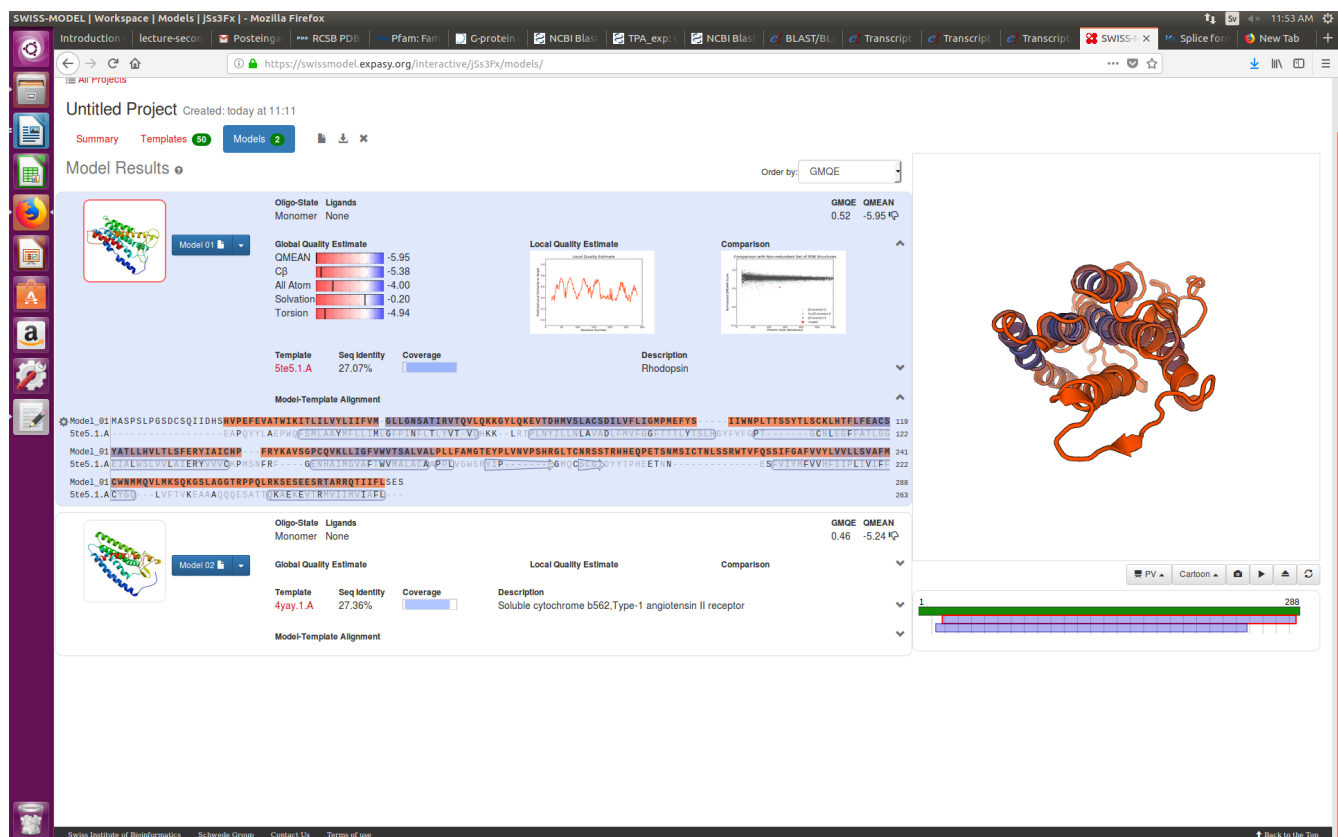
8. **SRP19 protein** Lingelbach K, Zwieb C, Webb JR, Marshallsay C, Hoben PJ, Walter P, Dobberstein B; , Nucleic Acids Res 1988;16:9431-9442.: Isolation and characterization of a cDNA clone encoding the 19 kDa protein of signal recognition particle (SRP): expression and binding to 7SL RNA

4. Identify the full length protein. What is the name and UniProt ID of the protein ? Which organism does it belong to?

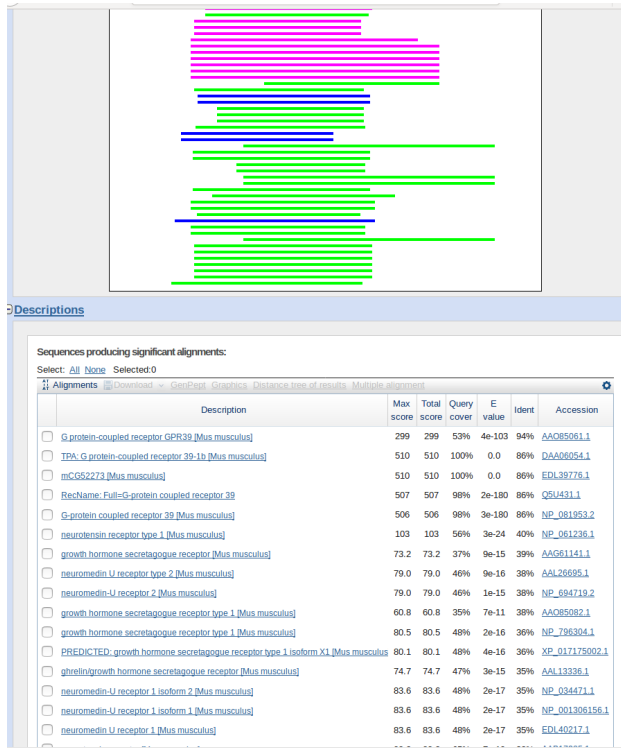
A6QR72, Organism Homo sapiens

5. What PDB files are homologous to your protein. Make a model of as long part as possible of your protein. Upload a picture of the model as a part your answer.

Checking the PDB database, the output is 211 homologous models. Nevertheless, the identity is very low for all of them (max. = 39%). Because of that I produced my own model on SWISS-MODEL.



6. Find the closest protein in mouse. What is the sequence identity ?



The closest protein is G-protein-coupled receptor GPR39 and the sequence identity is 94%.

7. What is the annotation score and evidence for the full length protein ?

The annotation score is 1 out of 5.

Evidence:

GPR39 splice variants versus antisense gene LYPD1: expression and regulation in gastrointestinal tract, endocrine pancreas, liver, and white adipose tissue." [Egerod K.L., Holst B., Petersen P.S., Hansen J.B., Mulder J., Hokfelt T., Schwartz T.W. Mol. Endocrinol. 21:1685-1698\(2007\)](#)

"Two alternatively spliced GPR39 transcripts in seabream: molecular cloning, genomic organization, and regulation of gene expression by metabolic signals." [Zhang Y., Liu Y., Huang X., Liu X., Jiao B., Meng Z., Zhu P., Li S., Lin H., Cheng C.H. J. Endocrinol. 199:457-470\(2008\)](#)

8. What is the sub-cellular location of your full length protein?

Integral component of membrane.

9. What is the molecular function of your full length protein ? Please provide the relevant GO-terms.

G-protein coupled receptor activity, GO:0004930

10. Use Ensembl to find possible splice forms of your protein. How many splice forms exist ?

There is one splice form, which consists of 2 exons and one intron.