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

    7Tm_1 (Rhodopsin-like family)

2. What is the secondary structure of your sequence ?

    Predicted 7 a-helices, 6 topological domains

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

    Pfam suggests no interactions

    Full-length protein interactions could be valid since our fragment is 250 residues out of 360 and is a transmembrane domain. Site of interaction is not specified or investigated in many experiments

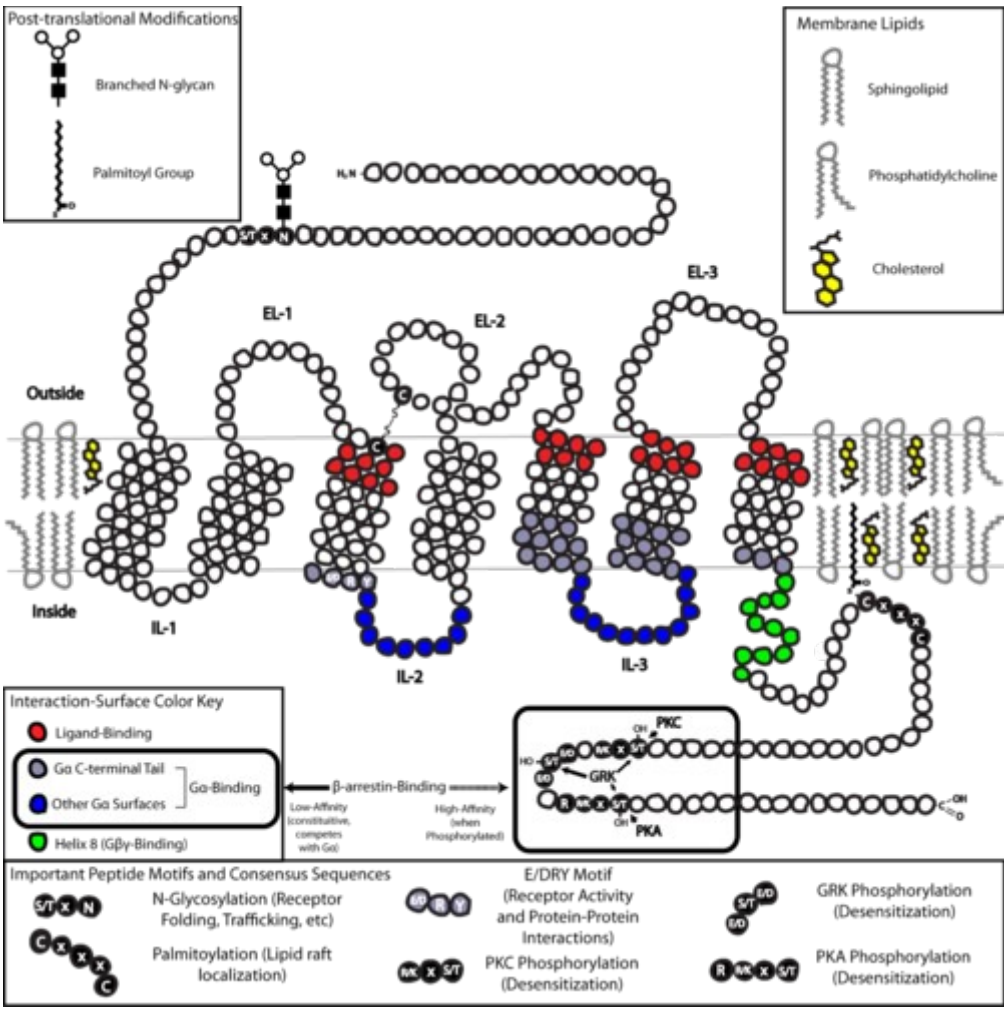
    IntAct search(full protein):
    mtad_bacan, q81pt2_bacan, cnr2_human, psmd2_human
    Evidence: experimental evidence from two-hybrid pooling approach, scintillation proximity assay, luminescence-based mammalian interactome mapping, all published results.

    String database
    Experimental evidence, medium confidence:
    SRP14(Alu RNA-binding), GNA15 (G-protein)

    BindingDB: over 8000 hits of chemical ligands with evidence of interaction, not sure how many are naturally or commonly occurring in vicinity.

    GO – process and function:
    There is a lot of evidence of interaction, and interaction partners can be inferred from processes and biological significance of GO terms. Inflammatory regulation, immune response, signal transduction have been experimentally verified(in some cases on mouse homologous protein), suggesting that CNR2 is commonly found on surface of B-cells, NK-cells and macrophages, where it can have cannabinoid-mediated suppressive effect on downstream inflammatory processes. An alternative isoform of the receptor is also found on peripheral neurons where it inhibits, through downstream signaling, signal transfuction. Evidently from these experiments, the receptor may interact with cannabinoid compounds.

    The illustration below demonstrates which parts of GPCR may interact with ligand, and this suggests that ligand-binding region is a part of my domain.



Since it's a trans-membrane domain, its integration into the lipid bilayer would require that at some point it interacts with a translocon, and chaperone proteins which assist membrane integration.

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

P34972
CNR2, Human

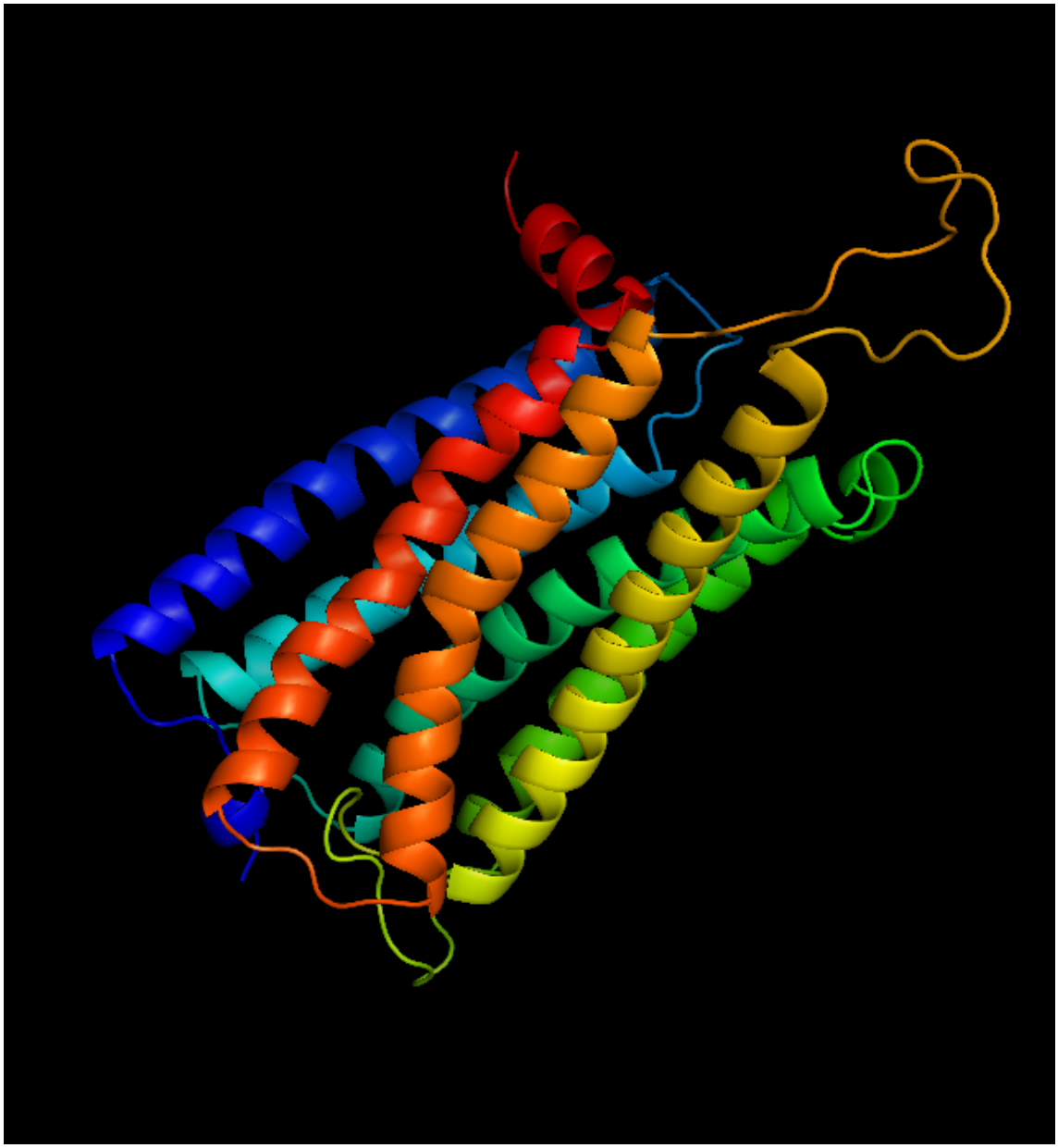
5. What PDB files are homologous to your protein. Make a model of as long part as possible of your protein. Upload a picture of the model as a part your answer.

Homologous structures
https://prosite.expasy.org/cgi-bin/pdb/pdb_structure_list.cgi?src=PS50262

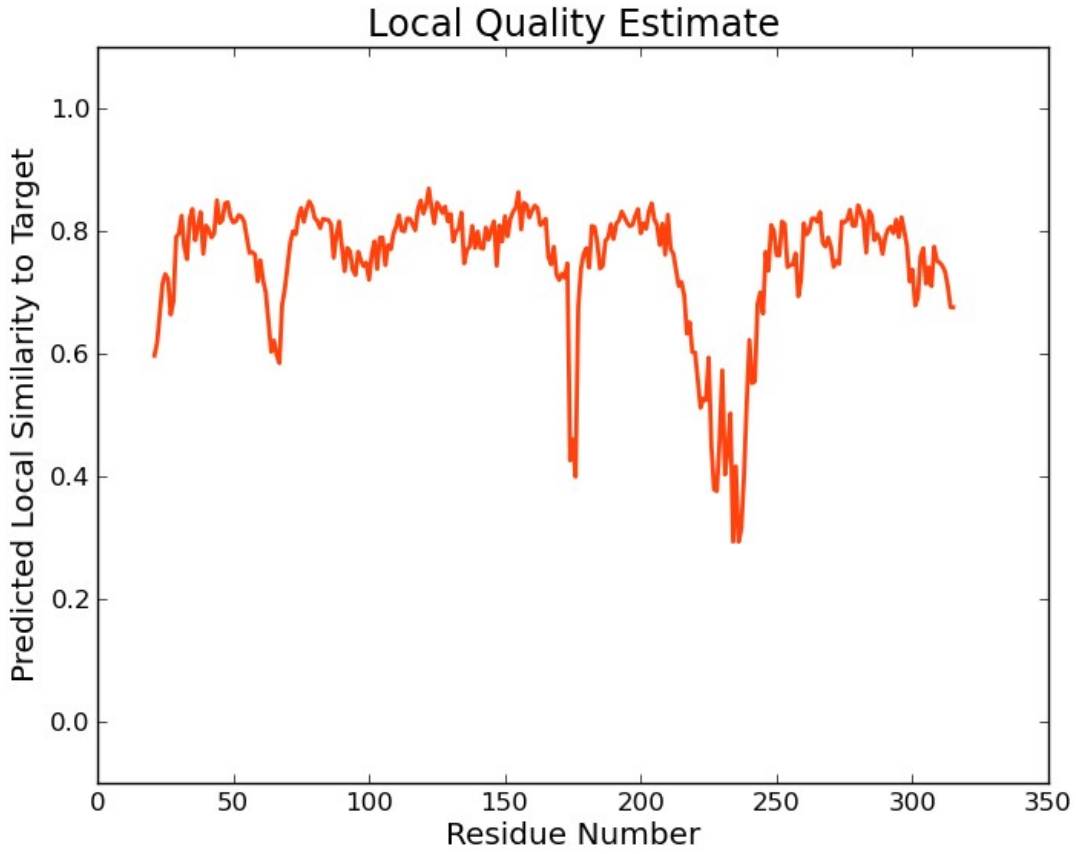
HMMER: 88 hits

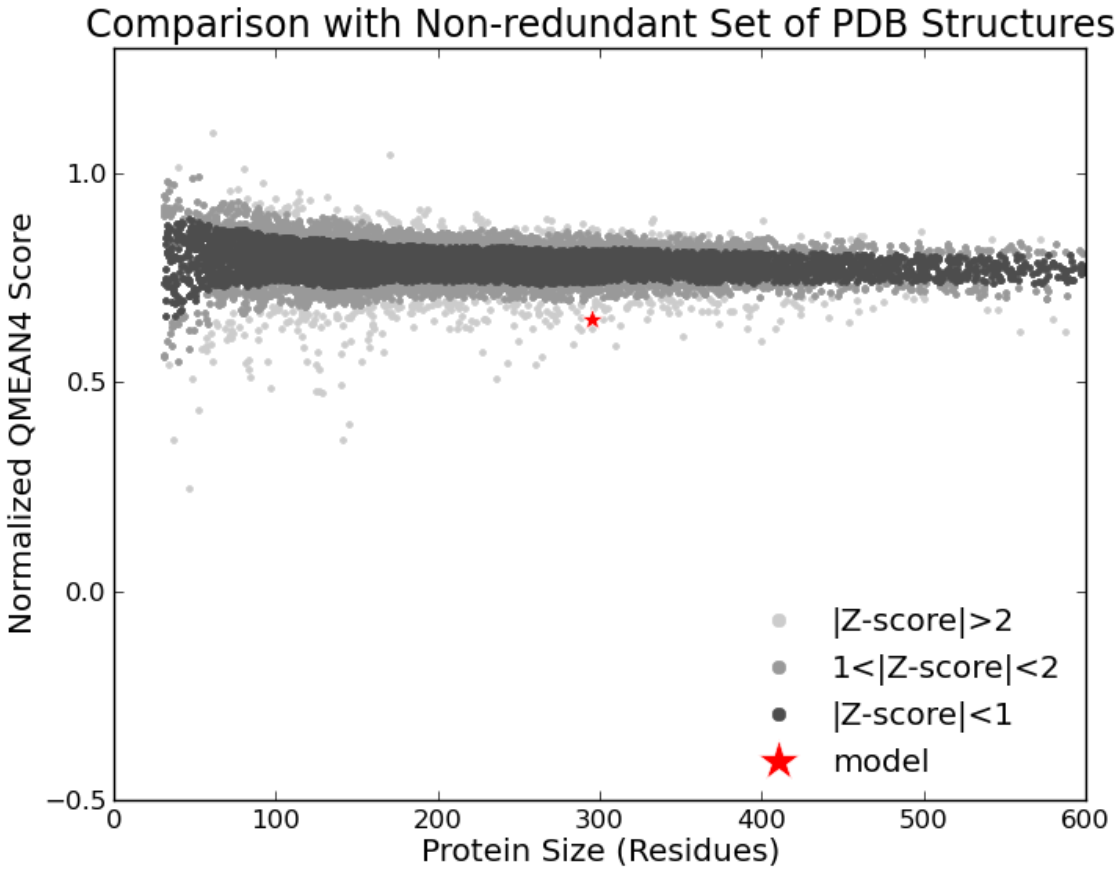
25 hits in PDB with some similarity, inclusing the high sequence similarity (46%) CNR1 with known full structure

1.



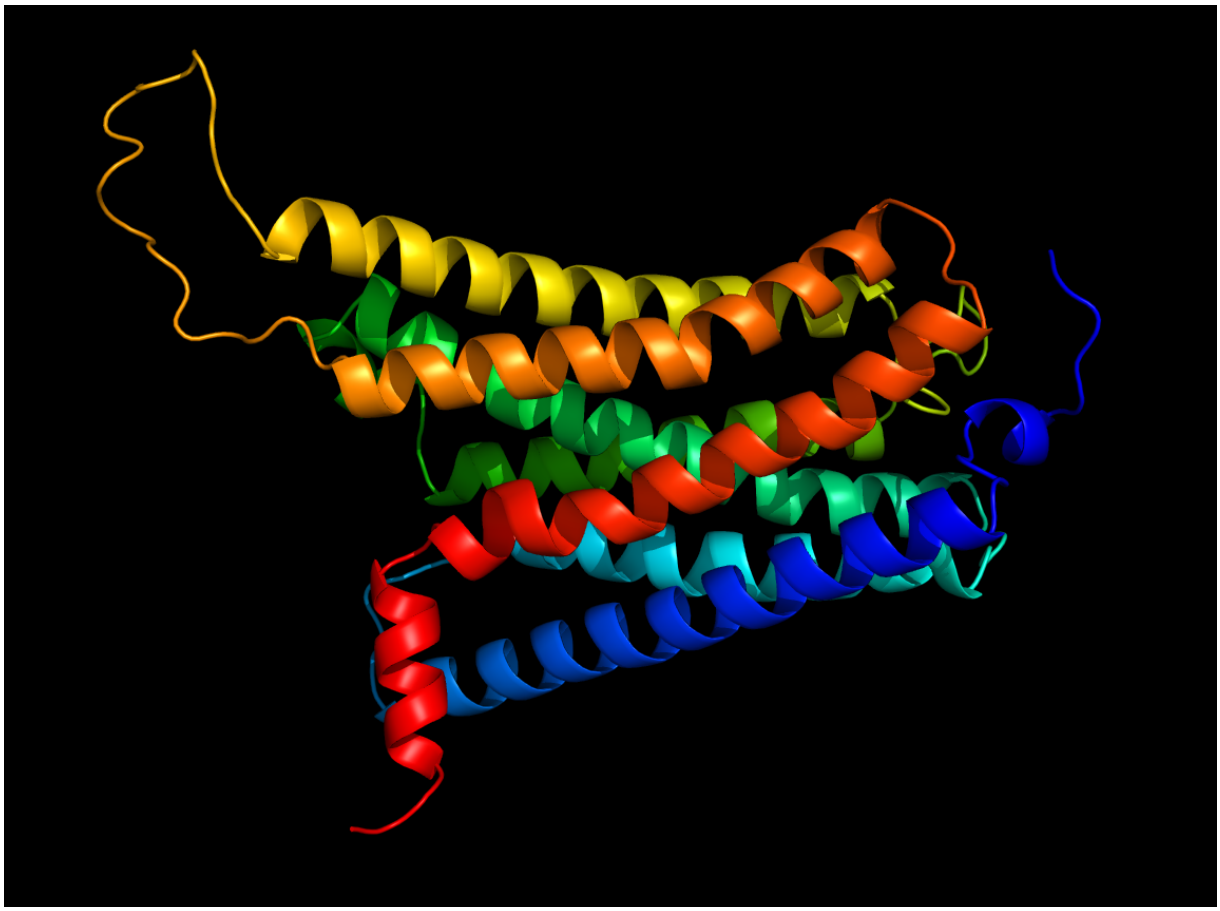
Best model built with Swissmodel, based on Cannabinoid receptor 1
PDB template: 5tgz
Model has Qmean of -3.17 which is the lowest(best) of all models generated. The fitness of model is quite poor, with -4 being a cutoff for a bad model.
Sequence similarity is 46%


Local Quality Estimate
Predicted Local Similarity to Target
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
200
250
300
350
Residue Number


Comparison with Non-redundant Set of PDB Structures
Normalized QMEAN4 Score
1.0
0.5
0.0
−0.5
0
100
200
300
400
500
600
Protein Size (Residues)
|Z-score|>2
1<|Z-score|<2
|Z-score|<1
model

From this assessment, the model is bordering a bad quality fit. The orange helix corresponds to the lowest quality fit.

2. Best model built with Modeller



PDB template: 5xra
Sequence similarity 37%
According to quality control attributes Verify3d, Solvx, the model is pretty poor. Anolea gives the best score of all models attempted, only 36.54% high energy pairs.

Both models look very similar when overlapped, with same amount of helices, their positions and number of turns

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

83%

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

5/5, Experimental evidence

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

Cell membrane/ Plasma membrane

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

GO:0004871 signal transducer activity
GO:0004930 G-protein coupled receptor activity
GO:0004949 cannabinoid receptor activity

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

1 Splice form