UGWU PASCHAL

+2348166207095 | ugwupaschal@gmail.com

https://www.linkedin.com/in/paschal-ugwu-52abb6229/

Omics Logic Project

Topic: The use of Pymol and UCSF Chimera to visualize BRCA1 protein alignment.

September, 2020

Abstract

In structure-function research, related structures are visualized and compared. I started working on this project to learn more about and gain ease in utilizing Pymol and UCSF Chimera to see how two (2) distinct BRCA1 proteins aligned to one another. AlphaFold, BlastP, EMBL-EBI, UniProtKB/Swiss-Prot, PDB database, UCSF Chimera, and Pymol were the tools utilized for this study. For the purpose of developing a breast cancer cure, researchers studying BRCA1 proteins will find this project useful.

Background

According to Meng and his colleagues in 2006, comparing related structures and viewing the structures in the context of sequence alignments are important tasks in protein structure-function research. While many programs exist for individual aspects of such work, there is a need for interactive visualization tools that: (a) provide a deep integration of sequence and structure, far beyond mapping where a sequence region falls in the structure and vice versa; (b) facilitate changing data of one type based on the other (for example, using only sequence-conserved residues to match structures, or adjusting a sequence alignment based on spatial fit); (c) can be used with a researcher's own data, including arbitrary sequence alignments and annotations, closely or distantly related sets of proteins, etc.; and (d) interoperate with each other and with a full complement of molecular graphics features.

Aim of project

To increase knowledge about, and confidence using Pymol and UCSF Chimera to visualize how two different proteins from BRCA1 gene align to each other.

Method

1. Opened AlphaFold Protein Structure Database.
2. Searched for the protein of interest.
3. Downloaded protein in PDB format (with accession number: Q3B890).
4. Copied the protein sequence and pasted in BlastP NCBI website.
5. Opened EMBL-EBI.
6. Under Blast, clicked on protein

a. Checked UniProtKB/Swiss-Prot (the manually annotated section of UniProtKB).

b. Under structure, I checked AlphaFold DB.

c. Then entered copied protein sequence.

d. Finally, I hit the "submit" button.

7. Opened the downloaded PDB file on Chimera.
8. Went back to PDB file website, entered similar protein in FASTA format and downloaded PDB (accession number: Q3B891) to compare visualization in Chimera with the previous one.
9. From tools on Chimera, I selected structure composition then match maker.
10. I selected reference structure as the "similar protein" which I downloaded from PDB site, while "structure to match" was the structure from AlphaFold database.
11. I left other structures on default then clicked "apply".
12. From favorite, I selected "model panel".
13. Again, I clicked on "favorite" then selected "Reply Log" to get information on the RMSD report.
14. Using Pymol, I clicked "A" in front of the "reference structure", this led to "action", then I selected "align".
15. I selected "align to entire structure to match".
16. I compared the RMSD value gotten from "Pymol" with that gotten from "UCSF Chimera".

Results

Proteins used for Project:
Accession number1: Q3B890
Accession number2: Q3B891

PYMOL
Match: read scoring matrix.
Match: assigning 259 x 480 pairwise scores.
MatchAlign: aligning residues (259 vs 480)...
MatchAlign: score 29.000
ExecutiveAlign: 26 atoms aligned.
ExecutiveRMS: 1 atoms rejected during cycle 1 (RMSD=11.37).
Executive: RMSD =   6.093 (25 to 25 atoms)
Executive: object "aln_AF-Q3B890-F1-model_v3_to_4ole" created.

UCSF CHIMERA
Computing secondary structure assignments for model(s) #1, #0
using ksdssp (Kabsch and Sander Define Secondary Structure
of Proteins) with the parameters:
  energy cutoff -0.5
  minimum helix length 3
  minimum strand length 3
Matchmaker 4ole.pdb, chain A (#1) with AF-Q3B890-F1-model_v3.pdb, chain A (#0), sequence alignment score = 15.4
with these parameters:
        chain pairing: bb
        Needleman-Wunsch using BLOSUM-62
        ss fraction: 0.3
        gap open (HH/SS/other) 18/18/6, extend 1
        ss matrix: (O, S): -6 (H, O): -6 (H, H): 6 (S, S): 6 (H, S): -9 (O, O): 4
        iteration cutoff: 2
RMSD between 6 pruned atom pairs is 1.285 angstroms; (across all 7 pairs: 1.957)

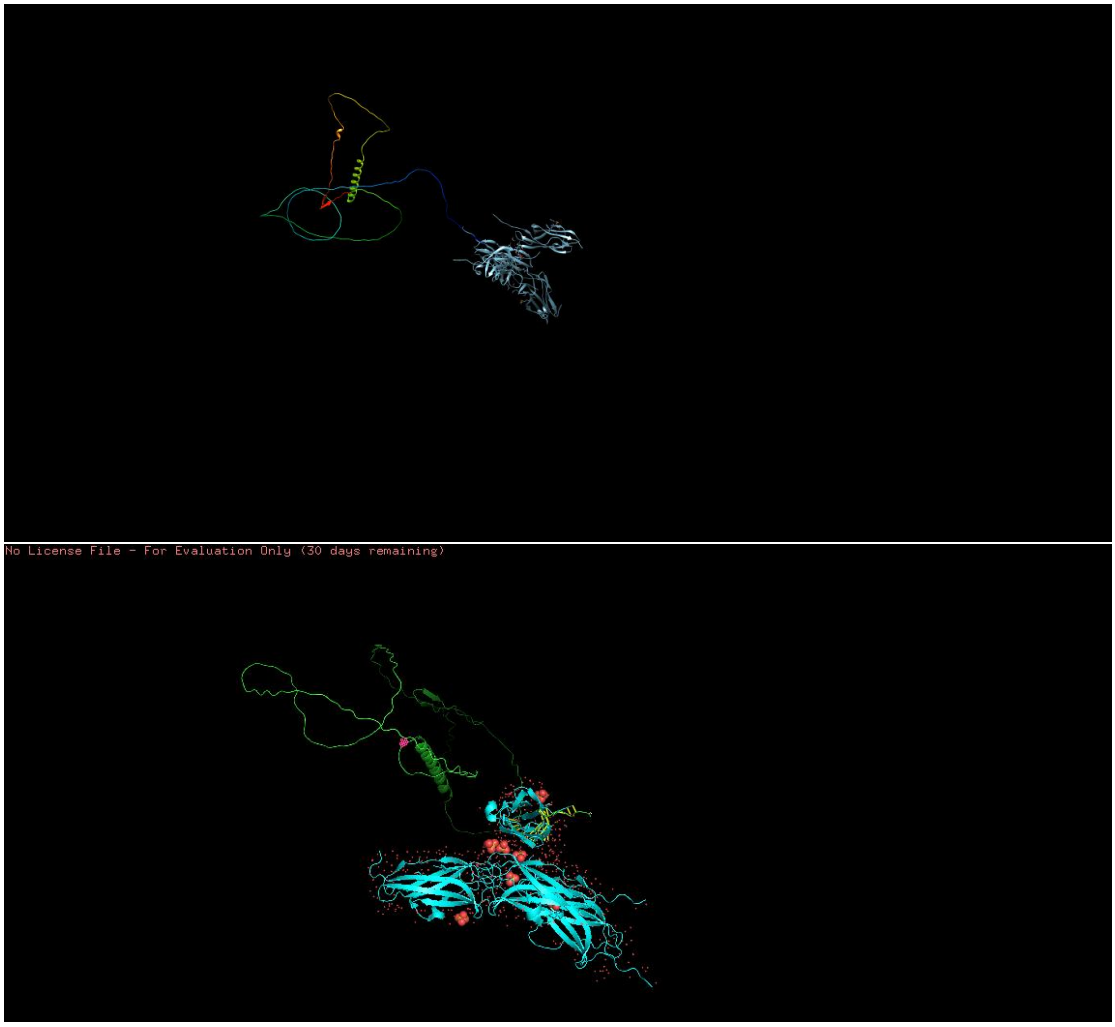

Figure 1: Alignment resuls from UCSF Chimera and Pymol respectively.

Conclusion

Pymol and UCSF Chimera tools were successfully used to visualize how two different proteins aligned to each other. This project has shown that these tools will be (and are already) greatly relevant in developing drugs to arrest several diseases that are caused by protein sequence mutation. This very project will be relevant to researchers working on BRCA1 proteins, to develop cure for breast cancer.

References

Meng, E. C., Pettersen, E. F., Couch, G. S., Huang, C. C., and Ferrin, T. E. (2006). Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC bioinformatics*, *7*(1), 1-10.