Biochemical society – Summer report – Rateep Nasim

One of the fundamental aims of this summer project was familiarizing myself with some popular molecular modelling and molecular viewing software such as Pymol and ChimeraX. Browsing through different protein structures obtained from the Protein Data Bank (PDB), I was able to view these structures primarily on ChimeraX (https://www.cgl.ucsf.edu/chimerax/) and also Pymol https://pymol.org/2/. The proper version of the latter was not available for free usage, hence I mainly decided to familiarize myself with Chimera X which remains completely free for academic use. With repetition, I was able to learn about a range of available tools and features implemented on Chimera X. This involved overlaying one protein 3D structure onto another using the matchmaker, highlighting the different domains and structures of interest and looking at specific residues for which different interactions were taking place. Earlier in the summer internship, I was also introduced to Uniprot (https://www.uniprot.org/)- a database of protein sequences, the PDB (https://www.rcsb.org/) – Protein Data Bank, as well as different databases such as the STRING and the STITCH database. Once accustomed to this as well as ChimeraX, I began looking at protein structure modelling using AlphaFold2.0. A major aim of my work was to learn the basic protein structure viewing, some relevant databases mentioned above and then apply the skill and knowledge gain to model some protein-protein interaction that is relevant from physiological and pathophysiological point of view. For this, we chose to predict the interaction between Neuropilin-1 (NRP-1) and some mutant forms of glycyl-transfer RNA (tRNA) synthetase (GlyRS) that underlie Charcot–Marie–Tooth (CMT) type 2 Diseases ($GlyRS^{CMT2D}$) .

As NRP-1 binding with $GLyRS^{CMT2D}$ interferes with VEGF binding to NRP-1, it was essential to look at this binding in further detail. VEGF-A binds with NRP1'S B1 domain and to see this interaction we focused on the B1 and B2 domain as a comparison to see the boundaries. Uniprot sequences and previous structures of B1B2 (of NRP-1 protein) interactions with VEGFA from the PDB (2QQI,1KEX) were first gathered. Next, the sequences were inserted into Alphafold and the session was run. From this we were able to conclude that the B1B2 structure is structurally similar to the 2QQI complex. Furthermore, the B1 domain generated by alphafold was structurally similar to the 1KEX complex on the Protein data bank, this was done by matchmaker tool on chimeraX.

The first main problem was seen as with my VEGFA165 + NRP1B1B2 structure. The CRCDKPRR peptide should be into the B1 domain as seen in another model. Therefore, it was possible that alphafold needed a VEGF receptor to get the model. The B2 structure was accurate but as the peptide needed to be into the B1 domain it was not structurally accurate. We then tried Cluspro (https://cluspro.bu.edu/login.php)instead of AlphaFold. The same result was gathered with Cluspro balance as once again, the CRCDKPRR peptide was not into the B1 domain. Cluspro Vdw+et worked however as the peptide was into the B1 domain.

To investigate more into detail about NRPB1 domain with VEGFA, the VEGFA protein was truncated and tested MiniVEGFA (the peptide ERTCRCDKPRR) with B1 binding using Rosettafold, EsmFold and Alphafold. Out of Rosettafold, EsmFold and Alphafold, Alphafold was the most accurate in modelling mini VEGFA with NRP1B1 domain. In a minimalistic matter, if NRP1 is modelled with mini VEGFA, then Alphafold can make a successful model and prediction for the complex.

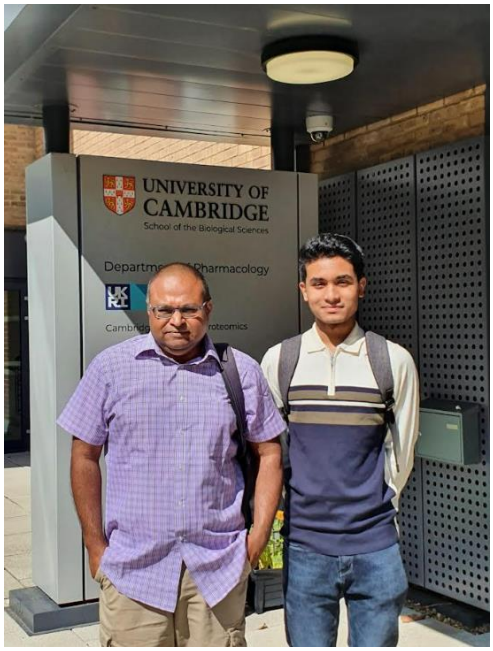
The 7P5U structure from the PDB was used as the DRAATPHHRPQPR peptide showed to interact with the B1 domain. From this the structure was constantly remodeled adding 1 more residue each time (PQR to RPQPR to HRPQPR) to determine the specific point in which it stops working. From these last 2 experiments we can conclude that Alphafold was unable to predict a suitable structure for the smaller peptide compared with the B1 domain, however Alphafold was able to successfully predict the complex of mini VEGFA with the B1 domain.

Another task was to predict NRP1-B1 complexes with MiniWARS, T2-WARS and human tryptophan – Trna ligase using ColabFold based AlphaFold multimer. MiniWARS binding with NRPB1 stops VEGFA binding with B1 as shown by Alphafold miniwars B1 domain. NRP1 B1 interacts differently with mini WARS than with full WARS.

Finally I predicted the surface electrostatics of neuropilin 1 using APBS and predicted the conformational flexibilitiy of proteins using Md simulation – CABSflex.

This project was important for me as I was provided with an opportunity to learn the basics of molecular modelling and molecular viewing using ChimeraX and Pymol. I was further able to learn how to predict different protein protein complexes and do protein structure modelling using Alpha Fold, Rosetta Fold, ESM fold. Finally, I was shown protein-protein docking using Cluspro and ZDOCK. This project provides societal benefit as different types of drugs such as peptidomimetic drugs could be developed to stop the unwanted binding and interaction of NRP-1 with mutant $GlyRS^{CMT2D}$ and therefore not preventing the binding of VEGF to NRP-1. This will, in turn, prevent peripheral neuronal degenerations associated with CMT2D. Finally, this summer internship helps the biochemical society achieve its strategies by offering flexible working models, this allowed me to complete a lot of the summer work from my remote work location at home. Furthermore, the internship provides inclusivity and diversity and finally an engagement experience with the community.

With the skills generated through this internship, I am able to create and predict protein protein interactions and generate different models for them which could be used in a scientific research paper or journal. Through this internship, I am currently working with the BMPR2 protein and trying to look at the interaction between BMPR2 and other receptors and proteins to potentially work on a research paper. Furthermore, I can potentially work with Dr Rahman to publish a paper based on our findings in the summer internship which will further elevate my research experience and lead to more opportunities in the future.


UNIVERSITY OF
CAMBRIDGE
School of the Biological Sciences