

Building 3D structure of protein from predicted distances

As the Minifold model predicts the distances and angles from primary sequences of proteins, the predicted distances can be utilised to create a 3D model of the protein once it is converted into the suitable format.

To model the 3D structure of a single protein(for example:- 1GOU)

- 1) create the PSSM(.pssm) file for the protein of interest such as 1GOU using POSSUM webserver:- possum.erc.monash.edu
- 2) create a notebook file and copy the required libraries and function utilities from `pretrain_model_pssm_l*_l.ipynb`
- 3) load the protein sequence of interest
- 4) load its pssm file and copy the piece of code from `modify_pssm.py` with necessary editings
- 5) load the trained model
- 6) predict the distances of the protein using `model.predict()`
- 7) copy the piece of code from `RR_format.py` to the same notebook file
- 8) edit the code using the resultant predictions
- 9) save the RR formatted file as "protein_sequence_ofinterest.rr/.txt"
- 10) use the RR formatted file as input for 3D structure modelling software GDFuzz3D:- <http://www.genesilico.pl/gdserver/GDFuzz3D/>

Results obtained from GDFuzz3D for the protein sequence 1GOU

