Protein shape similary score

The shape of a protein is critical to its function and plays an essential role in drug discovery. By finding similarities in shape between proteins, scientists can infer function and relationships between proteins. Since the available protein dataset is increasing significantly every year, it is necessary to find a method that allows rapid scanning to retrieve those similar proteins given a desired target. Therefore, we propose a method in which an ANN model receives one “target” and one "candidate" protein 3D shape, and predicts a "similarity score" between the two. By this way, given a desired shape, we could quickly scan the available dataset of protein structures (up to 500,000) and rank them based on the predicted score. The developed program will be part of a pipeline for design of new proteins.

We would use different neural network architectures, data analysis techniques, and use widely used Python frameworks like PyTorch.

keywords: computational structural biology, deep learning, data analysis, artificial neural networks, PyTorch.

Contact:
- Daniel Varela (daniel.varela@biochemistry.lu.se)
- Ingemar André (ingemar.andre@biochemistry.lu.se)

More Info:
- https://andrelab.biochemistry.lu.se/