Alpha Fold Notes:

Intro to Alpha Fold:

AlphaFold is a major advancement with the aim to predict a protein’s structure from its amino acid sequence alone. In nature, proteins reliably fold into precise 3D conformations that is critical for its function based on nothing more than the sequence of amino acids that it is composed of. In fact, mutations in proteins that lead to misfolding are often associated with disease states, for example, Alzheimer’s and Parkinson’s. However, we have not been able to understand this folding process nor predict the 3D shape of a protein based on its sequence alone [1].

The final output of AlphaFold is a file containing the 3D coordinates for every non-hydrogen atom in the protein. It also outputs a graph showing the confidence levels for every amino acid residue, which allows users to assess the reliability of the predicted structure [1].

The AlphaFold network directly predicts the 3D coordinates of all heavy atoms for a given protein using the primary amino acid sequence and aligned sequences of homologues as inputs (Fig. 1e; see Methods for details of inputs including databases, MSA construction and use of templates) [2].

Predicting the 3D structure of protein chains from their primary sequence of amino acids is a fundamental open problem in computational molecular biology. Any approach to this problem must deal with the basic fact that protein structures are invariant under translations and rotations. AlphaFold is a step towards to solve this issue [3].

ALPHAFOLD2

Recently, in the CASP14 experiment, AlphaFold2 (AF2) reached an unprecedented performance level in structure prediction of single-chain proteins16. Thanks to an advanced deep learning model that efficiently utilises evolutionary and structural information, this method consistently outperformed all competitors, reaching an average GDT\_TS score of 9016. Recently, RoseTTAFold was developed, trying to implement similar principles17. Since then, other end-to-end structure predictors have emerged using different principles such as fast multiple sequence alignment (MSA) processing in DMPFold218 and language model representations19 [4].

[1]

Felix, ‘A brief introduction to AlphaFold | Science | Felix Online’. <https://felixonline.co.uk/issue/1778/science/a-brief-introduction-to-alphafold> (accessed Nov. 09, 2022).

[2]

J. Jumper *et al.*, ‘Highly accurate protein structure prediction with AlphaFold’, *Nature*, vol. 596, no. 7873, Art. no. 7873, Aug. 2021, doi: [10.1038/s41586-021-03819-2](https://doi.org/10.1038/s41586-021-03819-2).

[3]

P. Baldi and G. Pollastri, ‘The Principled Design of Large-Scale Recursive Neural Network Architectures–DAG-RNNs and the Protein Structure Prediction Problem’, p. 28.