Poster

AlphaFold Database and its impact on the crystallographic world

Sudakshina Ganguly

EMBl-EBI, UK
sganguly@ebi.ac.uk

The recent surge in accuracy of predicted protein structures, particularly those generated by AlphaFold, offers a glimmer of hope in bridging the ever-growing gap between known protein sequences and experimentally determined structures. This newfound potential has not gone unnoticed by the structural biology community, with researchers swiftly recognizing the ability of AlphaFold predictions to aid in X-ray crystallography and cryo-EM structure determination.

The potential of AlphaFold in aiding X-ray crystallography and cryo-EM, two core techniques for structure determination, has been rapidly recognized. This impact is particularly felt in areas like structure determination itself, structure-based drug discovery, and structural bioinformatics. Notably, AlphaFold tackles a major hurdle in X-ray crystallography, the phase problem. By providing a predicted initial model of the protein, AlphaFold helps researchers calculate electron-density maps from diffraction data.

However, it's crucial to interpret predicted structures cautiously, considering their confidence scores. AlphaFold provides confidence scores (pLDDT and PAE) to highlight uncertainties. These scores pinpoint regions of the structure with lower reliability and help assess the relative positions of different parts.

The AlphaFold Protein Structure Database has revolutionised structural biology by amassing a staggering collection of over 214 million predicted protein structures. This number dwarfs the initial 300,000 structures released in 2021. Furthermore, the database has continuously evolved, incorporating new tools such as the structure similarity cluster. This tool, based on sequence and structural comparisons, allows researchers to navigate the vast amount of data and study protein function and evolution across diverse life forms.

This presentation delves deeper into the AlphaFold Database, exploring its potential and applications in biological research. We will examine the various approaches to selecting optimal predicted structures as initial models, along with the different scores and metrics used in this process, we will also highlight the new developments on the AlphaFold Database and how they can be used to tackle research questions.