Most Human Protein Structure Predictions are now Freely Available

George Arthur*

Department of Developmental Biology, University of Cambridge, Cambridge, United Kingdom

Proteins are made up of strings of amino acids that determine the function of proteins in cells when folded into 3D forms. Protein structures have been determined using experimental techniques such as X-ray crystallography and cryo-electron microscopy for decades. However, such approaches can be time consuming and expensive, and some proteins are not receptive to this type of research [1].

For the laypersons, it's protein structure prediction. Scientists will soon have access to software that precisely detects the 3D form of proteins. A comprehensive understanding of a protein's structure can provide essential insight into the mechanism of certain biological processes or serve as a springboard for medication development. Alpha Fold, a programme developed by Deep Mind, a UK-based artificial intelligence start up, has made substantial progress in lowering the time it takes to predict the structure of a protein from months to minutes with near-perfect accuracy. Now, according to a research published in Nature on July 22, Alpha Fold and the European Molecular Biology Laboratory (EMBL) have collaborated to create a publicly accessible database with over 350,000 protein structures. "With this knowledge, we will be better able to unravel the molecular mechanisms of life and accelerate our efforts to protect and treat human and environmental health [2].

The human proteome or all proteins known to be coded for by human DNA, numbers around 20,000. Only about 17% of those compounds have had their structures confirmed by laboratory investigation. Computational predictions of structures required a long time and were often wrong before the introduction of neural networks and current computer processors. According to Deep Mind, the new database has structures for 98.5 percent of the human proteome that are accurate to a high degree of confidence. The database also includes proteins from 20 model organisms, including Caenorhabditis elegans and Drosophila melanogaster, for a total of 350,000 structures [3].

Deep Mind estimates that by the end of the year, the database would have grown from 350,000 to 130 million structures. Access to this treasure could make it easier to build synthetic proteins, according to Nature, because it could be more consistently anticipated how they will interact with other proteins. Alpha Fold isn't the only programme for folding proteins. For example, Rose TT AFold, which was inspired by Alpha Fold, uses the same technology to calculate data in a variety of ways. It was only recently made available to the general public, and its designers expect it to gain from the new database. David Baker, one of the RoseTT AFold architects, told Science, "It's amazing they've made this available." [4]

"Until today, experimental methods have clearly determined the structures of only 17 percent of human proteins. Deep Mind's artificial intelligence technology, Alpha Fold, can take a protein's genetic sequence and predict its shape directly. Deep Mind researchers stated in a report published in the scientific journal Nature on Thursday that Alpha Fold had "very high confidence" in 36 percent of the novel human protein predictions and "some level of confidence" in 58 percent of them [5].

REFERENCES

1. Tunyasuvunakool K, Adler J, Wu Z, Green T, Zielinski M, Zidek A, et al. Highly accurate protein structure prediction for the human proteome. Nat. 2021; 596:590-6.

2. Hayashi T, Matsuzaki Y, Yanagisawa K, Ohue M, Akiyama Y. MEGADOCK-Web: an integrated database of high-throughput structure-based protein-protein interaction predictions. BMC bioinfo. 2018; 61-72.

3. Han GS, Yu ZG, Anh V, Krishnajith AP, Tian YC. An ensemble method for predicting subnuclear localizations from primary protein structures. PLoS One. 2013; 8:e57225.

4. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC bioinfo. 2008; 9:1-8.

5. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with Alpha Fold. Nat. 2021; 596:583-9.

***Correspondence to:** George Arthur, Department of Developmental Biology, University of Cambridge, Cambridge, United Kingdom; E-mail: arthur@georg.edu.uk

Citation: Arthur G (2021). Most Human Protein Structure Predictions are now Freely Available. Cell Dev Biol.10:239.