www.frontiersofknowledgeawards-fbbva.es

Acceptance speech

20 June 2023

John Michael Jumper, awardee in the Biology and Biomedicine category (15th edition)

It is a great honor to accept the BBVA Frontiers of Knowledge Award on behalf of myself and Demis Hassabis for our work on protein structure prediction. This award represents the achievement of the whole AlphaFold team at DeepMind whose creativity and ingenuity over several years are an enormous reason for AlphaFold's success.

Determining protein structures – the repeatable, intricate three-dimensional pattern of its atoms in its natural environment – has been a central problem in biology for more than sixty years. By determining protein structures, we are able to better understand how proteins function, how genetic variations alter their function, and aid in the design of new proteins to serve our own purposes. Through painstaking development of experimental techniques, scientists worldwide have solved around 200,000 protein structures experimentally and deposited them in the Protein Data Bank. The amount of work and time these structures represent is astounding, and determining a single experimental structure still takes months to years for skilled practitioners. Since nearly the beginning of the field, scientists had searched for theoretical or computational methods to predict a protein's structure without having to use experimental methods.

By the time that our team's work started on protein structure prediction, there had been a long history of methods developed to predict protein structure and, importantly, a very rigorous biennial assessment of the field in the CASP competition. These methods worked when the structure of a closely related protein was already known, but progress was slow in predicting the structure of proteins without known homologues. The fundamental challenge is that both the physical and evolutionary signals of protein structure are typically very weak and methods must be extremely accurate to make experimentally useful predictions. While integrating machine learning tools into physical and evolutionary approaches had delivered accuracy improvements, the results were still far short of the atomic accuracy needed for most applications.

The key innovations of our AlphaFold work were in showing how to build machine learning architectures and training procedures that were dramatically better adapted to understanding protein physics and evolution than standard machine learning techniques. The convolutional and standard transformer network architectures that had been previously used to try to learn protein structure had been developed in the context of learning images and text respectively. It would have required many more structures than are available in the Protein Data Bank to learn protein structure with the unadapted techniques we had previously. By fusing our intuitions on protein biology into the design of the network, we were able to build architectures that learned protein structure about one hundred times more efficiently, in terms of data required to achieve the same accuracy. The result is a deep learning system that predicts protein structures with a median error of less than an angstrom – less than the diameter of an atom.

The release of AlphaFold two years ago created an explosion of applications throughout the structural biology community, from understanding of basic biology and evolution to therapeutics and vaccines. For example, scientists have been able to use AlphaFold in conjunction with other experimental techniques to understand the structure and function of some of the largest and most complex molecular machines in the cell, such as the nuclear pore which controls what can come in and out of the nucleus of the cell. Another incredible development is how AlphaFold and related techniques have been used to aid protein engineering to develop new proteins for therapeutic purposes, both for small proteins and large systems such as a "molecular syringe" for targeted delivery of protein therapeutics with potential applications to cancer therapies. It is exciting to see every day the breadth of science that has been enabled or accelerated by accurate structure prediction.

On behalf of Dr. Hassabis and myself, we are deeply honored to accept this award. I want to close by saying that we think we have taken a significant step into the era in which the study of molecular biology will become computational and ultimately more predictive, enabling us to find new cures for diseases and engineer new solutions to some of life's most pressing problems.