AI continues to build the foundation for a remarkable future in biology

Published August 19, 2022

Christopher Gannatti, CFA
Global Head of Research

28 July 2022 was an historic day in both biology and artificial intelligence (AI). DeepMind, a firm specialising in AI research owned by Alphabet, made freely available the structural data on more than 200 million proteins from its AlphaFold tool. This represents data on roughly 1 million species and covers the vast majority of known proteins on earth[1].

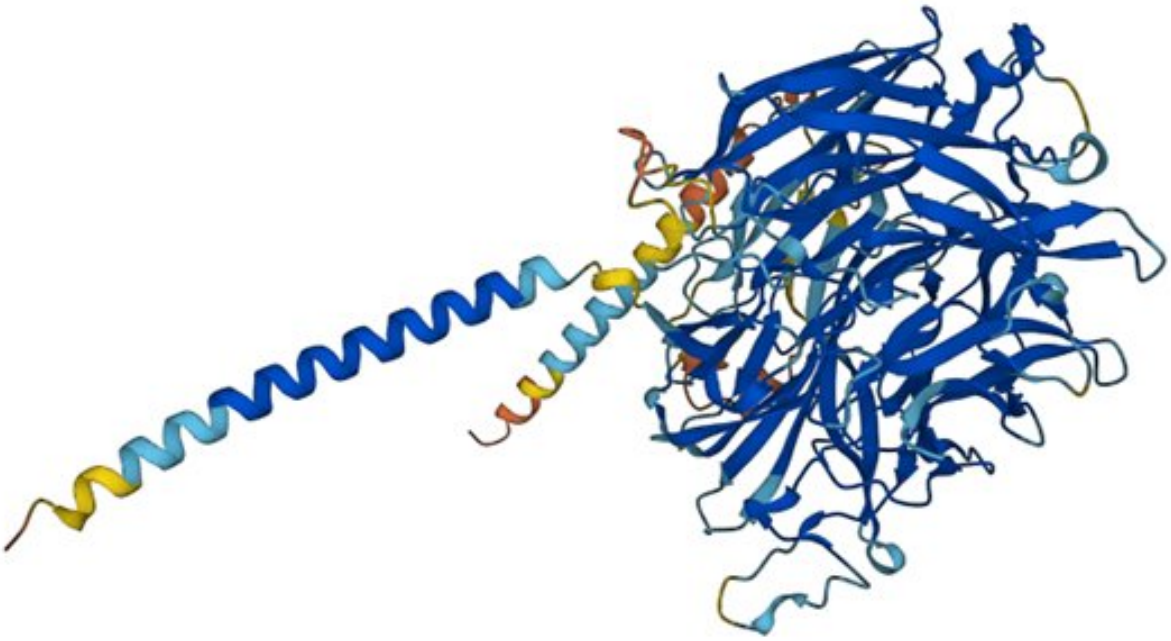
In proteins, shape can determine function

In the late 1990's into the early 2000's, the scientific community was awash with news of the race to sequence the human genome. This genome contains the instructions embedded in DNA about how cells should build certain structures, typically through the formation of proteins that are made from different combinations of amino acids.

In a sense, DNA is the instruction manual, amino acids are the building blocks, and proteins are the product. Knowing the code, however, is not the full story.

Looking at figure 1 is instructive on the point. This is the image of a protein that may protect the organism responsible for malaria from an attack by the human immune system. Even if you knew the list and the order of all the amino acids, it would be difficult to go from that list to something that looks like figure 1 in three dimensions.

Figure 1: Protein associated with the malaria parasite



Source: https://alphafold.ebi.ac.uk/entry/Q8I3H7

The importance of the shape of the protein could not be overstated:

- It can correspond to the way in which it might react in the presence of different molecules, like those associated with different drug therapies
- Variations on the shape—sometimes termed mutations—could be instructive in determining the causal factors of certain symptoms or diseases
- Parts of the shape could be used as targets—think of the 'spike protein' associated with the virus behind the Covid-19 pandemic, specifically targeted within the mRNA vaccines.

AlphaFold Represents a Leap Forward on the Journey

Scientific breakthroughs are difficult, in that in many cases one builds on another and another and another…a process that can take decades prior to widespread results that impact the lives of the general public. For instance—we sequenced the human genome, but that did not necessarily lead to immediate cures of all sorts of diseases or conditions. mRNA[2] research had been occurring for decades, but the Covid-19 pandemic was somewhat of a catalyst to supercharge the process to use it as a case for vaccines.

AlphaFold's new database is therefore unlikely to lead to immediate cures for difficult conditions. The critical element with regards to how researchers that would have formerly had to undertake a cumbersome process of X-ray crystallography to determine the shape of a given protein could instead go to the database. Experimental techniques would still have their place, but less time would have to be spent on the equivalent of the 'blank page.'

What's also incredible is that AlphaFold's database is, in conjunction with the European Molecular Biology Laboratory's European Bioinformatics Institute (EMBL-EBI) freely available with a simple interface. It also provides an estimate of the accuracy, recognising that predictions based on AI do not yield perfect results all of the time. Roughly 35% of the 214 million predictions are deemed highly accurate—roughly as good as experimental results. A further 45% are deemed to be accurate enough for many applications[3].

Drug discovery—Better therapeutics developed more efficiently

Even before the onset of inflation at the levels we see in the summer of 2022, it was widely recognised that drug development was time consuming and expensive, and as a result many different medications carried with them exorbitant price tags. Any process that could mitigate this pressure without degrading the quality of the therapies would be valuable.

Considering the following could be instructive as the space continues to progress[4]:

- **Pipeline growth:** 20 smaller companies focused on AI drug discovery, typically with a focus on smaller molecules, over a period from 2010 to 2021, had development pipelines that were roughly 50% as robust as those of 20 of the largest 'big pharma' companies. We recognise that the reporting of pipelines may not be perfect and that a molecule in a pipeline is not a finished product, but activity is the first step on the path
- **Pipeline composition:** This is not always disclosed, but from the information available it does indicate that the AI-focused companies will tend to focus on well established biological targets for their therapies, around which much is known. Data is the fuel for AI, and these companies will also want higher chances of success. Bigger pharma companies will be more likely to venture into more emerging areas of drug discovery
- **Chemical structures and properties:** It is too early to be able to draw any robust conclusions regarding AI drug discovery efforts versus big pharma efforts at this point
- **Discovery Timelines:** Preliminary data would appear to indicate that, if traditional approaches would tend to take 5 or 6 years in preclinical phases, AI-focused drug discovery might be able to, in certain cases, take this timeline down to 4 years

We'd note that currently it's a story of more 'progress' than 'perfection', in that we would appear to be some distance away from AI being able to fully create new drugs, but AI is representing an entirely new set of tools that could have beneficial impacts. AlphaFold's database, for example, may provide drug researchers with important inputs and catalysts for different ideas, even if it doesn't have the immediate answer or cure right there in its system.

Focus on the AI & BioRevolution megatrends

At WisdomTree, we focus on both the AI and the BioRevolution megatrends (click to find out more). What we see here with the case of AlphaFold is an important case study in the fact that AI is a tool that can have the potential to supercharge other megatrends, in this case the BioRevolution. It is no accident that the BioRevolution is ramping up at the same time there are massive amounts of data, massive amounts

of processing power and other things like cloud computing readily available. It is very exciting to consider what the coming decades can bring within these areas.

[1] Source: Callaway, Ewen. "'The Entire Protein Universe': AI Predicts Shape of Nearly Every Known Protein." Nature. Volume 608. 4 August 2022.

[2] mRNA – Messenger ribonucleic acid

[3] Source: Callaway, 4 August 2022.

[4] Source: Jayatunga et al. "AI in small-molecule drug discovery: a coming wave?" Nature Review: Drug Discovery. Volume 21. March 2022.

Related blogs

+ Can AI replace people? The truck-driving case study

+ The nearly unlimited demand for better energy storage

+ Part 1: A realistic framing of the progress in Artificial Intelligence

+ Part 2 AI: Easy to Discuss the Hype, but what about Substance

Important Risks Related to this Article

Important Information

Marketing communications issued in the European Economic Area ("EEA"): This document has been issued and approved by WisdomTree Ireland Limited, which is authorised and regulated by the Central Bank of Ireland.

Marketing communications issued in jurisdictions outside of the EEA: This document has been issued and approved by WisdomTree UK Limited, which is authorised and regulated by the United Kingdom Financial Conduct Authority.

WisdomTree Ireland Limited and WisdomTree UK Limited are each referred to as "WisdomTree" (as applicable). Our Conflicts of Interest Policy and Inventory are available on request.

For professional clients only. The information contained in this document is for your general information only and is neither an offer for sale nor a solicitation of an offer to buy securities or shares. This document should not be used as the basis for any investment decision. Investments may go up or down in value and you may lose some or all of the amount invested. Past performance is not necessarily a guide to future performance. Any decision to invest should be based on the information contained in the appropriate prospectus and affer seeking independent investment, tax and legal advice.

The application of regulations and tax laws can offen lead to a number of different interpretations. Any views or opinions expressed in this communication represent the views of WisdomTree and should not be construed as regulatory, tax or legal advice. WisdomTree makes no warranty or representation as to the accuracy of any of the views or opinions expressed in this communication. Any decision to invest should be based on the information contained in the appropriate prospectus and affer seeking independent investment, tax and legal advice.

This document is not, and under no circumstances is to be construed as, an advertisement or any other step in furtherance of a public offering of shares or securities in the United States or any province or territory thereof. Neither this document nor any copy hereof should be taken, transmitted or distributed (directly or indirectly) into the United States.

Although WisdomTree endeavours to ensure the accuracy of the content in this document, WisdomTree does not warrant or guarantee its accuracy or correctness. Where WisdomTree has expressed its own opinions related to product or market activity, these views may change. Neither WisdomTree, nor any affiliate, nor any of their respective officers, directors, partners, or employees accepts any liability whatsoever for any direct or consequential loss arising from any use of this document or its contents.