2024 Superconvergence biorevolution series: Premonitions of AlphaFold 3

Published 27 August 2024

Christopher Gannatti, CFA
Global Head of Research

Key Takeaways

- ChatGPT brought a lot of attention to the idea of using AI to predict text, and predicting text naturally lead toward predicting images or even video based on different prompting. Can we use these concepts to predict different ways proteins and molecules might interact within cells?
- Protein folding and structure is a fascinating problem to consider, with a single protein having potentially something like 10^{300} different ways it might fold—far to many for any system to try to attempt solutions one at a time.
- It's remarkable to consider that the protein folding question has persisted for something like 50 years and we have started to see a steady stream of advances, such as AlphaFold, AlphaFold 2, AlphaFold 3… It's amazing how the convergence of AI, compute capability, and biotech understanding is pushing these questions forward quite quickly.

When constructing our Biorevolution strategy, we worked alongside futurist Dr Jamie Metzl, who is a member of the World Health Organization's expert committee on human genome editing. We believe that we are on the precipice of a remarkable period, which could last a few decades, where we challenge and ultimately evolve how we do things, such as:

- How we handle human health care
- How we grow food for an expanding global population
- How we generate novel materials, chemicals and energy from biological sources
- How we think about storing massive amounts of data with higher density and fidelity than we have in the past

Dr Metzl recently published the book *Superconvergence: How the Genetics, Biotech, and AI Revolutions will Transform our Lives, Work and World*. Over the summer, we will publish a series of blogs that draw attention to some of the ideas presented in the book.

The bottom line

Thematic investing, in a sense, is about storytelling. Superconvergence does a great job conveying the narrative behind the WisdomTree BioRevolution ESG Screened Index.

Thank you ChatGPT

In my opinion, the best thing about ChatGPT is how it made the concept of generative artificial intelligence (AI) accessible to almost anyone. In its early phases, the popularity of generative AI focused on generating text, but this has since evolved into generating images, sounds, and videos. It makes sense to consider that anything with a system and a structure, where a system can be trained on the different rules and relationships, could lend itself to being predicted.

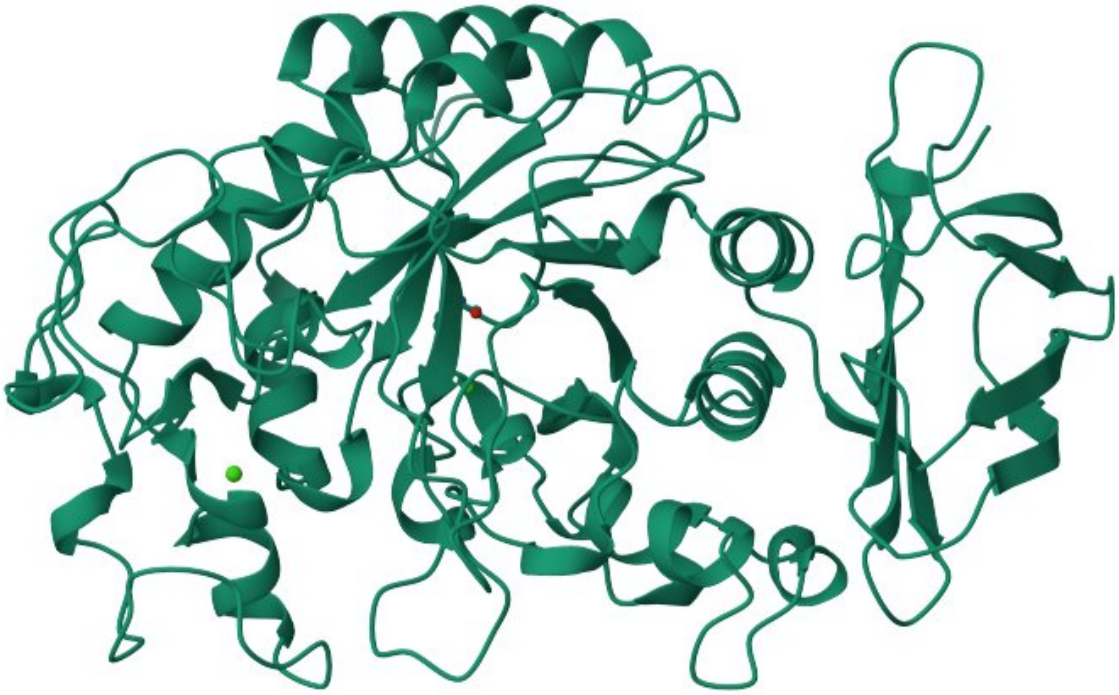
The power of these systems lies in the fact that they can continually make predictions that might be thought-provoking and less about whether any of the predictions are always 100% accurate. When we shift our thinking into biotechnology, it is not about 'predicting a cure' but instead 'predicting a novel path of research' that may lead a researcher towards an interesting therapeutic result.

Can generative AI systems predict useful protein structures?

Predicting different series of words that make sense against a prompt or the position of pixels within an image is one thing—predicting interrelationships between molecules in a biological system is entirely different.

We show Figure 1 to give anyone who does not have a background in molecular biology (I include myself in that group) an appreciation for the visual complexity of a single protein—in this case, human salivary amylase (saliva) from the Protein Data Bank in Europe. It's impossible not to be impressed with how scientists physically performing X-ray crystallography established the foundation of determining protein structure.

Figure 1: Human saliva protein structure



Source: https://www.ebi.ac.uk/pdbe/entry/pdb/1smd

Announcement of AlphaFold in 2020

Professor John Moult, Co-Founder and Chair of Critical Assessment of Protein Structure Prediction (CASP) from the University of Maryland, said[1]:

We have been stuck on this one problem—how do proteins fold up—for nearly 50 years. To see DeepMind produce a solution for this, having worked personally on this problem for so long and after so many stops and starts, wondering if we'd ever get there, is a very special moment.

CASP experiments aim to establish the current 'state of the art' in protein structure prediction, identify progress and highlight where future efforts may be productively focused. They occur every two years, with the first having taken place in 1994[2].

In 1969, when considering the challenge of predicting a protein's three-dimensional structure, it was estimated that a typical protein could have something like 10,000 possible conformations. This tells us that a 'brute force' approach where a system looks at every possibility would not be feasible[3].

Simulating a cell

We can learn a lot from better, higher-resolution simulations. On 18 March 2024, Nvidia announced an Earth Climate Digital Twin—the concept being that if we can simulate climate and weather at an increasing resolution, we could use that to understand better and ultimately predict changes in weather and climate[4].

While the idea seems simple, collecting and processing enough data to have even a reasonable shot at being accurate enough to matter is not easy.

Given the difficulty and cost of organising clinical trials to test different possible therapies, it's not much of a leap to think that if we could only simulate the human body and its underlying systems, we could learn a lot and depend less on organising human clinical trials.

It's interesting to step back and recognise that throughout history, we have analysed DNA, rRNA, amino acids, and proteins. We are building our understanding piece by piece. The complexity involved in each of these steps is staggering.

From Superconvergence:

DeepMind founder Demis Hassabis told Eric Topol in 2022 that:

One of my dreams in the next 10 years is to produce a virtual cell. What I mean by virtual cell is you model the whole function of the cell with an AI system. You could do virtual experiments on that cell, and the predictions that come out of that would hold when you check them in the wet lab. Can you imagine, if you had something like that, how much faster and more efficient that would make the whole drug discovery and clinical trials process? . . . You can think of what we've done with AlphaFold as the first step of the ladder. . . . Then you build up slowly, maybe to pathways and eventually to cells and then ultimately perhaps the whole organism. That's the dream[5].

Introducing AlphaFold 3

Something funny about publishing writing on the internet or speaking on different podcasts is that, if you know how to search for it, you can find all sorts of predictions that people have made or aspirations that they have had frozen in that moment in time. It's interesting to read that quote from Demis Hassabis in July 2024, when we know that AlphaFold 3 has recently been released.

On 8 May 2024, the following was published[6]:

Introducing AlphaFold 3, a new AI model developed by Google DeepMind and Isomorphic Labs. By accurately predicting the structure of proteins, DNA, RNA, ligands and more, and how they interact, we hope it will transform our understanding of the biological world and drug discovery.

My admittedly amateur read is that if AlphaFold 3 was simulating an entire cell, they'd have written that, so my interpretation is that this represents an important step on that path and that we should expect more and more versions that can incorporate a greater and greater resolution to cells and then ultimately living organisms. It's fascinating that even as the book Superconvergence was being published, the state of the art in seeking to predict protein structure moving toward cell simulation continued to advance. We sequenced the genome, we predicted how different proteins would fold, and now we are starting to predict how different molecules and proteins will interact. It appears that these advancements are coming faster and faster. Picturing a timeline in your mind's eye makes it very clear, in that the Human Genome Project was completed in 2003[7], AlphaFold was announced in 2020[8] and AlphaFold 3 was announced in 2024[9].

As computational hardware and AI advance, the types of things that can be accomplished in biotechnology should also have the same advancement potential. The convergence of these different megatrends could make the coming years quite exciting.

[1] Source: https://deepmind.google/discover/blog/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology/
[2] Source: https://predictioncenter.org/index.cgi
[3] Source: Levinthal, Cyrus (notes by A. Rawitch). 1969. How to fold graciously. In Mössbauer Spectroscopy in Biological Systems, edited by P. Debrunner, J. C. M. Tsibris, and E. M?nck. Proceedings of a meeting held at Allerton House, March 17 and 18, 1969, Monticello, Illinois. Urbana, Ill.: University of Illinois Press.
[4] Source: https://nvidianews.nvidia.com/news/nvidia-announces-earth-climate-digital-twin#:~:text=GTC%E2%80%94To%20accelerate%20efforts%20to,and%20climate%20at%20unprecedented%20scale.
[5] Source: Metzl, Jamie. Superconvergence: How the Genetics, Biotech, and AI Revolutions will Transform our Lives, Work and World. Timber Press: 2024.
[6] Source: https://blog.google/technology/ai/google-deepmind-isomorphic-alphafold-3-ai-model/#life-molecules
[7] Source: https://www.genome.gov/human-genome-project/timeline#:~:text=More%20%2B-,2003,two%20years%20ahead%20of%20schedule.
[8] Source: https://deepmind.google/discover/blog/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology/
[9] Source: https://blog.google/technology/ai/google-deepmind-isomorphic-alphafold-3-ai-model/#life-molecules

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.