

# Série/2024 sur la superconvergence en matière de biotechnologie/: les prémonitions d'AlphaFold/3

Publié le 27 août 2024

**Christopher Gannatti, CFA**

Global Head of Research

- ChatGPT a suscité une grande attention autour de l'idée d'utiliser l'IA pour prédire du texte, ce qui conduit naturellement à prédire des images et même des vidéos sur la base de différentes invites. Pouvons-nous utiliser ces concepts pour prédire différentes manières dont les protéines et les molécules pourraient interagir au sein des cellules ?
- Le repliement et la structure des protéines constituent un sujet fascinant, dans la mesure où une seule protéine peut présenter environ  $10^{300}$  manières différentes de se replier, un chiffre bien trop élevé pour que des solutions soient tentées une par une.
- Aspect incroyable, la question du repliement des protéines persiste depuis une cinquantaine d'années, et nous commençons à observer un flux constant d'avancées dans ce domaine, parmi lesquelles AlphaFold, AlphaFold 2, AlphaFold 3... Il est fascinant de constater comment la convergence de l'IA, des capacités de calcul et de la biotechnologie permet d'avancer relativement rapidement sur ces questions.

Lors de l'élaboration de notre [stratégie Biorevolution](#), nous avons travaillé aux côtés du futurologue [Dr Jamie Metzl](#), qui est membre du comité d'experts de l'Organisation mondiale de la santé sur l'édition du génome humain. Nous sommes convaincus d'être sur le point de connaître une formidable période, qui pourrait durer plusieurs décennies, au cours de laquelle nous remettrons en question et feront en fin de compte évoluer notre mode de fonctionnement, notamment :

- La manière dont nous gérons les soins de santé pour les individus
- La manière dont nous adaptons la production alimentaire à une population mondiale croissante
- La manière dont nous générons des énergies, matériaux et produits chimiques novateurs à partir de sources biologiques
- La manière dont nous appréhendons le stockage de quantités massives de données avec une densité et une fidélité plus élevées que par le passé

Le Dr Metzl a récemment publié l'ouvrage intitulé [Superconvergence: How the Genetics, Biotech, and AI Revolutions will Transform our Lives, Work and World](#) (Superconvergence : comment les révolutions de la génétique, de la biotechnologie et de l'IA transformeront notre existence, notre travail et notre monde).

Au cours de l'été, nous publierons une série d'articles qui mettront en lumière quelques-unes des idées présentées dans cet ouvrage.

L'investissement thématique consiste avant tout à construire un récit cohérent. L'ouvrage Superconvergence réussit parfaitement à présenter le récit qui se cache derrière l'[indice WisdomTree BioRevolution ESG Screened](#).

### **Merci ChatGPT**

Le meilleur aspect de ChatGPT réside selon moi dans le fait qu'il ait rendu le concept d'intelligence artificielle (IA) générative accessible à quasiment tous les êtres humains. Dans ses premières phases, l'IA générative était surtout populaire pour ses capacités de génération de texte, puis les possibilités ont évolué avec la création d'images, de sons et de vidéos. Il est logique de considérer que tout ce qui possède un système et une structure, avec un système capable de faire l'apprentissage des différentes règles et relations, peut se prêter à la prédiction.

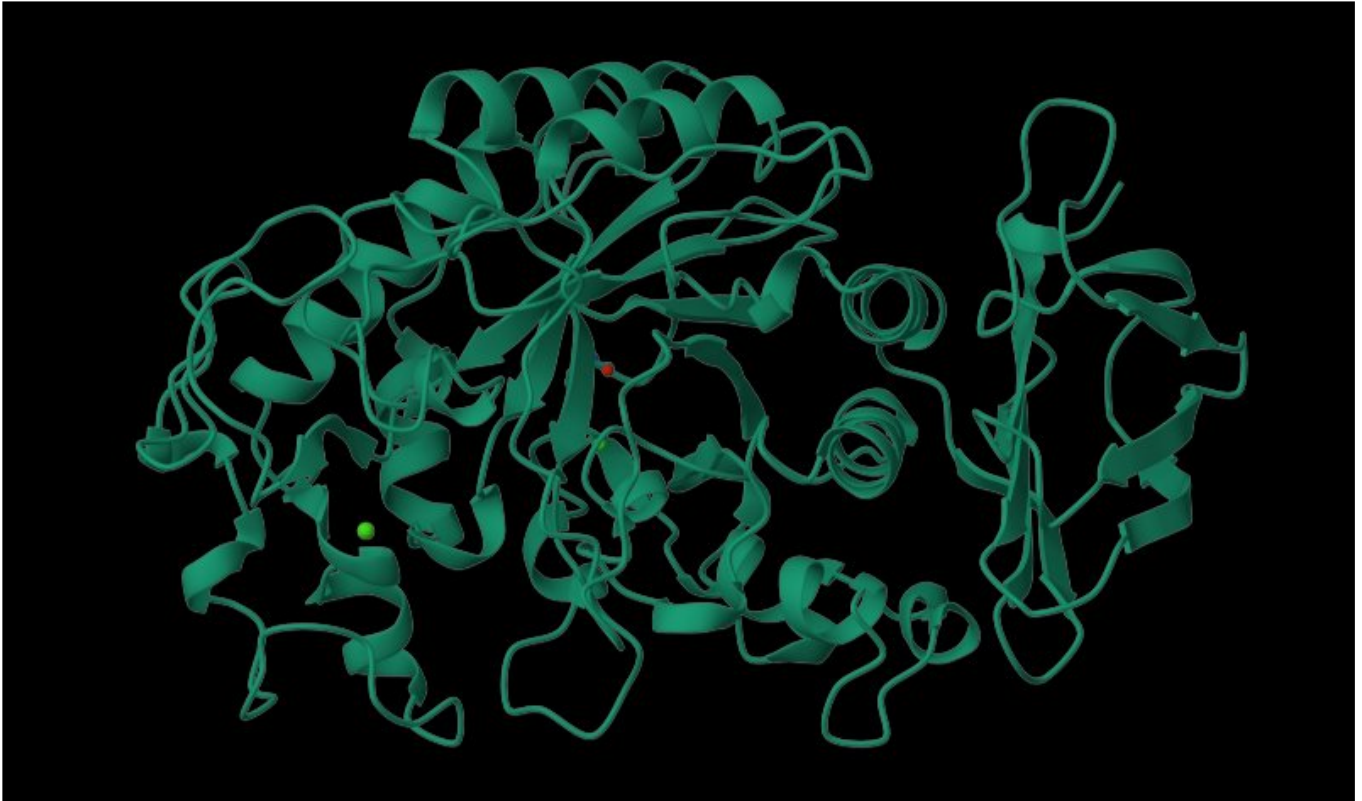
La puissance d'un tel système réside plus dans le fait qu'il peut, en continu, effectuer des prédictions qui suscitent la réflexion, et moins dans la possibilité que ces prédictions soient toujours exactes à 100 %. Lorsque nous transformons notre réflexion en biotechnologie, il ne s'agit pas de « prédire un remède », mais plutôt de « prédire un nouveau chemin de recherche », susceptible de guider un chercheur vers un résultat thérapeutique intéressant.

### **Les systèmes d'IA générative peuvent-ils prédire des structures protéiques utiles ?**

C'est une chose de prédire des séries de mots revêtant un sens à partir d'une invite, ou la position des pixels sur une image ; c'en est une autre, tout à fait différente, de prédire des interactions entre molécules dans un système biologique.

L'illustration 1 permet à tous ceux qui ne connaissent pas la biologie moléculaire (et j'en fais partie) de découvrir la complexité visuelle d'une seule protéine, en l'occurrence une amylase salivaire humaine (salive) issue de la banque de données européenne sur les protéines. Impossible de ne pas être impressionné par la manière dont les scientifiques réalisent physiquement une cristallographie aux rayons X pour déterminer les bases de la structure d'une protéine.

### **Illustration 1 : Structure protéique de la salive humaine**



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

### **Annnonce d'AlphaFold en 2020**

Le professeur John Moult, cofondateur et président du département d'évaluation critique de la prédiction de la structure des protéines (Critical Assessment of Protein Structure Prediction, CASP) de l'Université du Maryland, a déclaré<sup>1</sup>:

*Depuis près de 50 ans, nous ne parvenons pas à résoudre ce problème : comment les protéines se replient-elles ? Ayant personnellement travaillé pendant de si nombreuses années sur ce point critique, et après tant d'interruptions et de reprises de nos efforts, tant d'interrogations sur la possibilité de le résoudre un jour, je trouve absolument formidable que DeepMind apporte une solution pour le surmonter.*

Les expériences CASP visent à établir « l'état de l'art » actuel en matière de prédiction de la structure des protéines, à identifier les avancées, ainsi qu'à déterminer les aspects sur lesquels les futurs efforts pourraient être axés de manière productive. Elles sont effectuées tous les deux ans, la première ayant eu lieu en 1994<sup>2</sup>.

En 1969, face au défi consistant à prédire la structure tridimensionnelle d'une protéine, il a été estimé qu'une protéine typique pouvait présenter environ 10 000 conformations possibles. On comprend dès lors qu'une approche approximative consistant pour un système à examiner toutes les possibilités ne soit pas réalisable<sup>3</sup>.

### **Simuler une cellule**

Nous pouvons apprendre beaucoup de choses grâce à des simulations de meilleure résolution. Le 18 mars 2024, Nvidia a annoncé le lancement d'un jumeau numérique du climat terrestre reposant sur le concept suivant : si nous parvenions à simuler le climat et la météorologie avec une résolution croissante, nous pourrions utiliser cette simulation pour mieux comprendre et en fin de compte pour prédire les changements climatiques et météorologiques<sup>4</sup>. Bien que l'idée semble simple, il n'est pas facile de collecter et de traiter suffisamment de données pour aboutir à des chances raisonnables de précision satisfaisante.

Compte tenu de la difficulté et du coût que représente l'organisation d'essais cliniques visant à tester différentes thérapies possibles, il n'est pas absurde de penser que si nous parvenions à simuler le corps humain et son fonctionnement, nous pourrions en apprendre beaucoup plus, sans dépendre pour cela de l'organisation d'essais cliniques chez l'homme.

Il est intéressant de prendre un certain recul pour saisir que nous avons au cours de l'histoire analysé l'ADN, l'ARN, les acides aminés et les protéines. Nous développons notre compréhension étape par étape. La complexité qui caractérise chacune de ces étapes est vertigineuse.

Extrait de l'ouvrage Superconvergence :

Demis Hassabis, fondateur de DeepMind, a confié à Eric Topol en 2022 :

*L'un de mes rêves au cours des dix prochaines années serait de produire une cellule virtuelle. J'entends par là le fait de modéliser l'ensemble du fonctionnement d'une cellule dans un système d'IA. Nous pourrions alors réaliser des expériences virtuelles sur cette cellule, et les prédictions qui en résulteraient seraient toujours valables au moment de leur vérification en « wet lab ». Si vous disposiez d'une telle solution, imaginez à quel point seraient améliorées la rapidité et l'efficacité de l'ensemble du processus de découverte de médicaments et d'essais cliniques. ... On peut considérer ce que nous avons accompli avec AlphaFold comme le premier barreau de l'échelle... Nous progresserons ensuite lentement, peut-être en direction de voies biologiques, puis de cellules, et finalement de l'ensemble de l'organisme. C'est le rêve que nous poursuivons<sup>5</sup>.*

### **Présentation d'AlphaFold 3**

Ce qui est amusant dans la publication de travaux sur Internet ou au travers de podcasts, c'est que si vous savez comment les rechercher, vous pouvez obtenir toutes sortes de prédictions que des personnes ont formulées, ou d'aspirations qui étaient les leurs à ce moment-là. Il est intéressant de lire cette citation de Demis Hassabis de juillet 2024, lorsque nous savons qu'AlphaFold 3 a récemment été lancé.

Le 8 mai 2024, le document suivant a été publié<sup>6</sup>:

*Présentation d'AlphaFold 3, un nouveau modèle d'IA développé par Google DeepMind et Isomorphic Labs. En prédisant avec précision la structure des protéines, de l'ADN, de l'ARN, des ligands et plus encore, ainsi que la manière dont ils interagissent, nous espérons transformer notre compréhension du monde biologique et de la découverte de médicaments.*

La lecture que j'en fais en tant que profane, c'est que si AlphaFold 3 simulait une cellule entière, la présentation l'indiquerait. Mon interprétation est par conséquent la suivante : il s'agit d'une étape importante dans cette direction, et nous devons nous attendre à toujours plus de versions, incorporant une résolution sans cesse croissante des cellules, puis en fin de compte des organismes vivants. Il est fascinant de constater que même après la publication de l'ouvrage Superconvergence, l'état de l'art dans la recherche qui entend parvenir à prédire la structure des protéines via la simulation d'une cellule a continué de progresser. Nous sommes parvenus à séquencer le génome, à prédire le repliement de différentes protéines, et nous commençons désormais à prédire comment différentes molécules et protéines interagiront. Ces avancées semblent aujourd'hui de plus en plus rapides. Cela est évident si l'on se figure la chronologie : le projet sur le génome humain a été achevé en 2003, AlphaFold a été annoncé en 2020, et AlphaFold 3 a été lancé en 2024. À mesure que progressent les matériels informatiques et l'IA, les découvertes susceptibles d'être atteintes en matière biotechnologique présenteront certainement le même potentiel d'avancée. La convergence de ces différentes mégatendances pourrait rendre les années à venir passionnantes.

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 de A. Rawitch). 1969. How to fold graciously. Dans Mössbauer Spectroscopy in Biological Systems, publié par P. Debrunner, J. C. M. Tsibris et E. Münck. Conclusions d'une réunion organisée à Allerton House, les 17 et 18 mars 1969, Monticello, Illinois. Urbana, Illinois : Presse de l'Université de l'Illinois.

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 (Superconvergence : comment les révolutions de la génétique, de la biotechnologie et de l'IA transformeront notre existence, notre travail et notre monde). 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/time-line#:~: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

### Informations importantes

**Communications commerciales publiées dans l'EEE** Ce document est publié et approuvé par WisdomTree Ireland Limited, une société autorisée et réglementée par la Central Bank of Ireland.

**Communications commerciales émises dans des juridictions en dehors de l'EEE** Ce document est publié et approuvé par WisdomTree UK Limited, une société autorisée et réglementée par la Financial Conduct Authority du Royaume-Uni.

WisdomTree Ireland Limited et WisdomTree UK Limited sont toutes les deux désignées comme « WisdomTree » (le cas échéant). Notre Politique sur les conflits d'intérêts et notre Inventaire sont disponibles sur demande.

**Réservé aux clients professionnels uniquement. Les informations figurant dans ce document sont fournies à titre informatif et ne constituent pas une ore de vente, ou une sollicitation d'ore d'achat de titres ou d'actions. Ce document ne doit pas être utilisé comme fondement d'une décision d'investissement. La valeur des investissements peut fluctuer et vous êtes susceptible de perte tout ou partie du montant investi. La performance passée ne constitue pas nécessairement une indication des performances futures. Toute décision d'investissement doit être fondée sur les informations figurant dans le prospectus approprié et sur des conseils indépendants en matière d'investissement, fiscaux et juridiques.**

L'application des réglementations et lois fiscales peut souvent conduire à des interprétations diérentes. Tous les points de vue ou opinions exprimés dans cette communication représentent les points de vue de WisdomTree et ne doivent pas être interprétés comme des conseils réglementaires, fiscaux ou juridiques. WisdomTree ne donne aucune garantie ou représentation quant à l'exactitude des vues ou opinions exprimées dans cette communication. Toute décision d'investissement doit être fondée sur les informations contenues dans le prospectus approprié et après avoir sollicité des conseils indépendants en matière d'investissement, fiscaux et juridiques. Ce document n'est pas et ne doit en aucun cas être interprété comme une publicité ou une ore publique d'actions ou de titres aux États-Unis ou dans toute province ou tout territoire des États-Unis. L'introduction, la transmission et la distribution (directes ou indirectes) de l'original ou d'une copie de ce document sont interdites aux États-Unis.

Bien que WisdomTree s'efforce d'assurer l'exactitude du contenu de ce document, WisdomTree ne peut en garantir l'exactitude. Les fournisseurs de données tiers sollicités pour obtenir les informations contenues dans le présent document ne donnent aucune garantie ou représentation de quelque sorte en rapport avec ces données. Lorsque WisdomTree exprime ses propres opinions concernant le produit ou l'activité du marché, ces opinions sont susceptibles de changer. WisdomTree, ses alliés et leurs dirigeants, directeurs, partenaires ou employés respectifs déclinent toute responsabilité pour toute perte directe ou indirecte découlant de l'utilisation de ce document ou de son contenu.