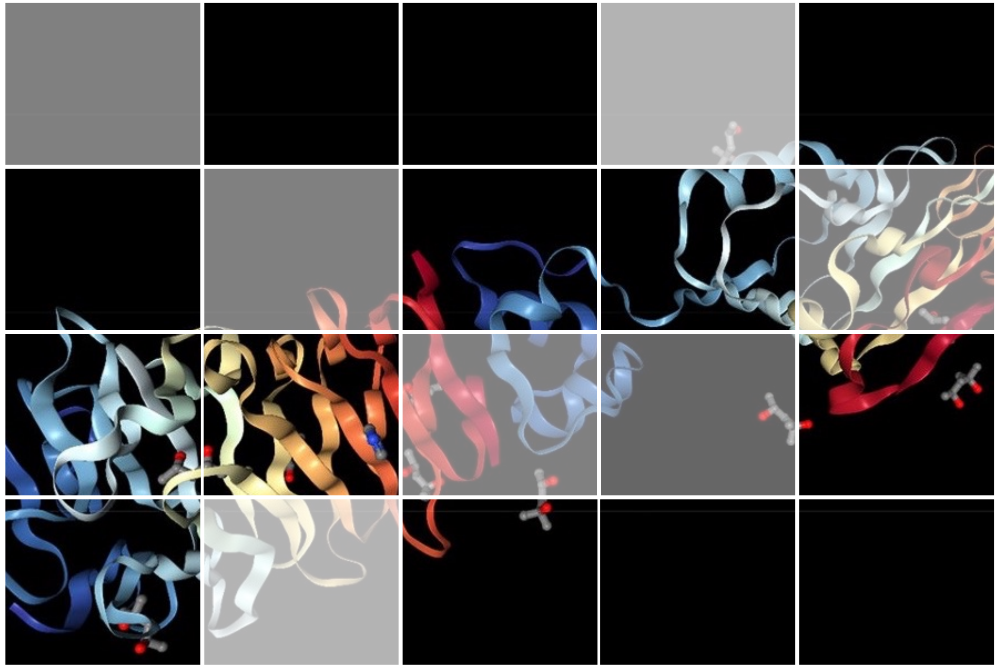
Hipolit's Biotech Breakdown: **Introduction to AlphaFold**



Hipolit Cichocki **2022.04**
Dragon Gate Investment Partners LLC

Disclaimer

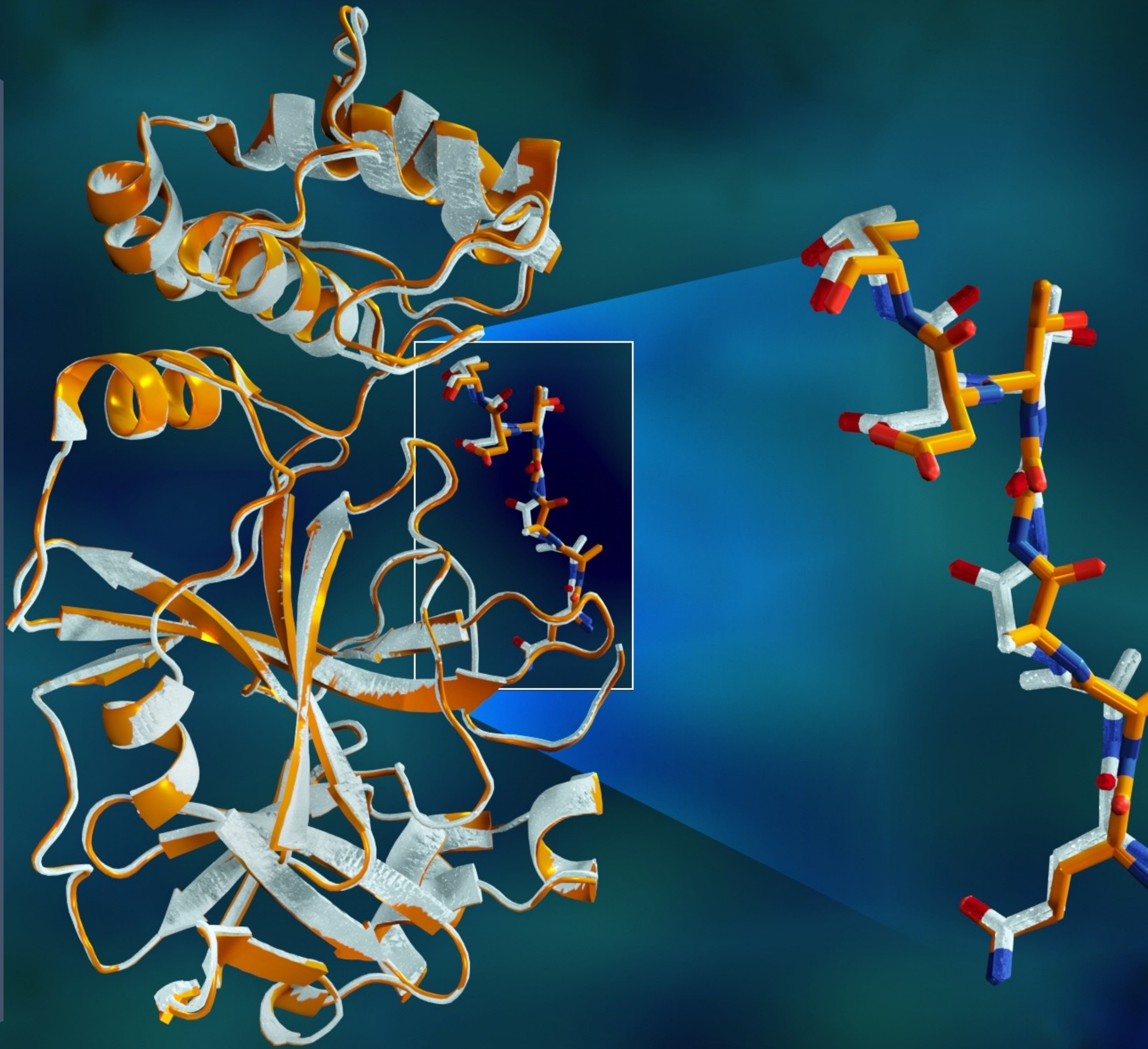
This report is only for informational purposes and does not purport to make any forecasts or predictions and nothing in this report should be construed as doing so. It is merely intended to help investors better understand the company in a research report format.

Dragon Gate Investment Partners prepared the information in this report. Dragon Gate Investment Partners has no obligation to inform you when information in this report changes.

This report is for information purposes only. Under no circumstances is it to be used or considered as a solicitation to buy or sell any securities. While the information contained herein has been obtained from sources we believe to be reliable, Dragon Gate Investment Partners does not represent that it is accurate or complete, and accordingly, should not be relied upon as such. Risk factors and actual results may differ significantly from the information contained herein. This report or any portion hereof may not be reprinted, sold, or redistributed without the written consent of Dragon Gate Investment Partners.

This report is prepared for Institutional Consideration Only. Estimates of future performance are based on assumptions that may not be realized. Past performance is not necessarily a guide to future performance.

Why Protein Folding Matters



Building Blocks of Life

Proteins don’t just **underpin** the biological processes in your body but every **biological process in every living thing.** They’re the building blocks of life.

Currently, there are **around 100 million known distinct proteins**, with many more found every year. **Each one has a unique 3D shape that determines how it works and what it does**.

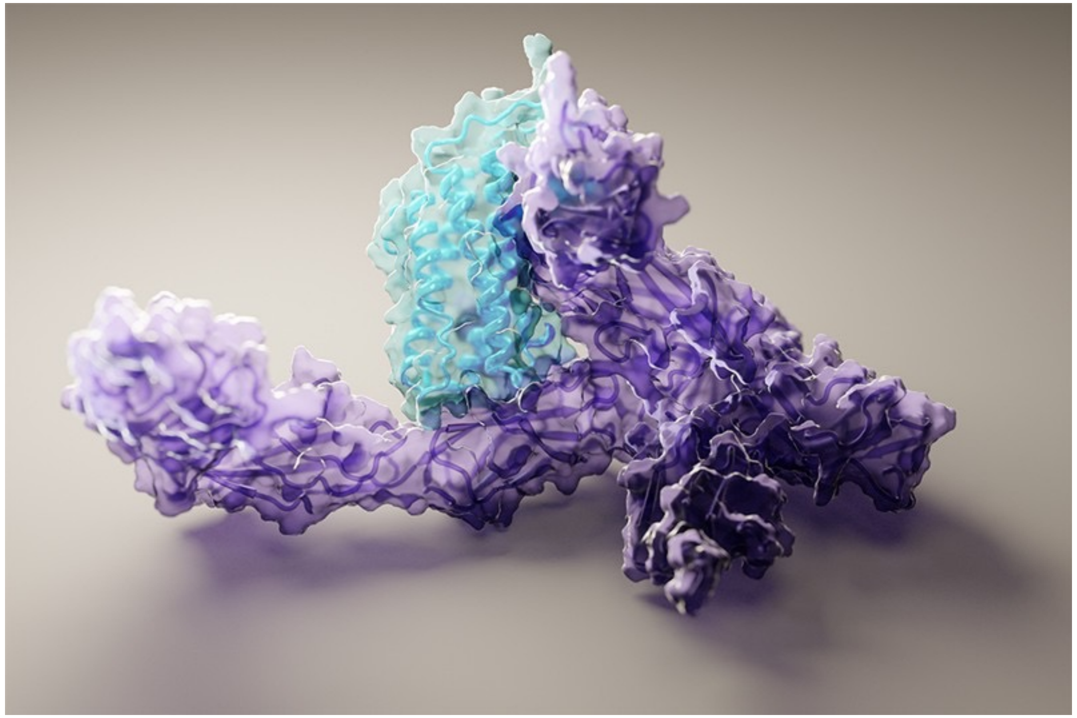
Figuring out the exact structure of a protein remains an expensive and often time-consuming process, meaning **we only know the exact 3D structure of a tiny fraction of the proteins known to science**.

Finding a way to close this rapidly expanding gap and predict the structure of millions of unknown proteins could not only help us tackle disease and more quickly find new medicines but perhaps also unlock the mysteries of how life itself works.

Source: “AlphaFold”, https://deepmind.com/research/case-studies/alphafold. Accessed Mar 22, 2022.

Why Protein's 3D Structure Matter

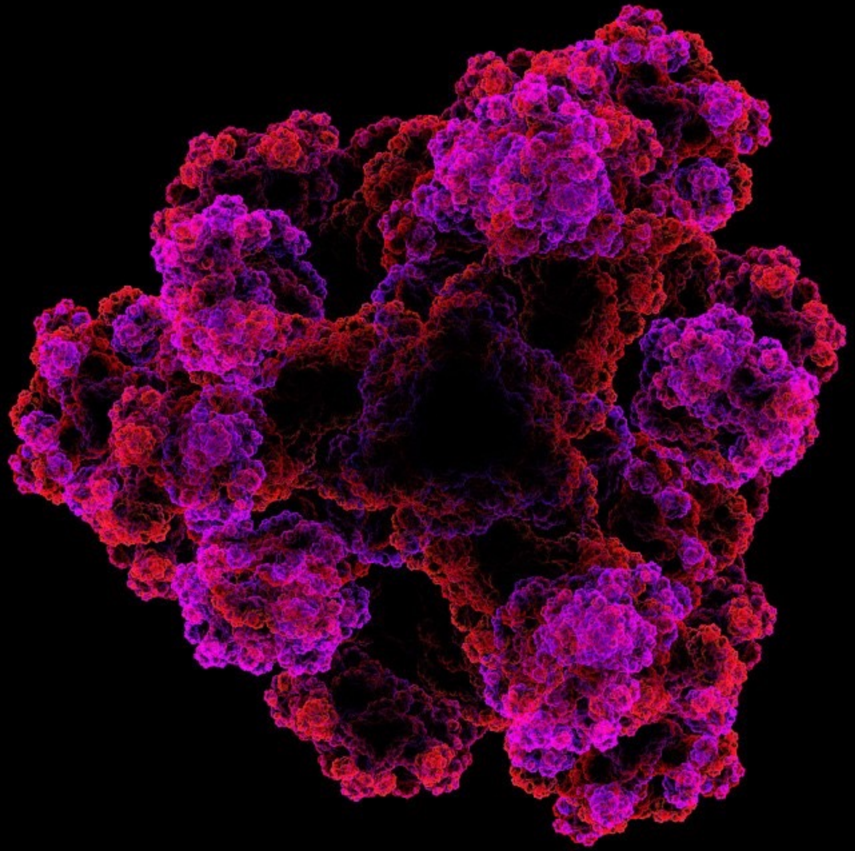
- **The human body contains a subset of 20,000 - 25,000 protein-coding genes with the ability to generate 20,000 to over millions of unique types of proteins.**
  - All processes taking place in an organism have proteins acting somewhere.
- **Each has characteristic shapes, a 3D structure, that allow them to perform a precise function**. As mentioned above, some proteins are structural, others transport other molecules, others are receptors, etc.
  - The specific shape of each protein is tightly related to their function. For example, some proteins form pockets named active sites that perfectly fit to bind a particular target molecule.
- **The distinct "functional native structure" of proteins is important because it exposes several binding sites, channels, receptors and thus impacts how they bind other molecules or how proteins physically interact with others and assemble into complexes for structural or regulatory processes.**



Source: Nature.

Source: "Why Does Predicting a Protein's 3D Structure Matter? The Biological Impact of AI System AlphaFold", https://piip.co.kr/en/blog/alphafold-protein-structure-prediction-importance-1. Accessed Apr. 7, 2022.

Protein Background



Protein Structure

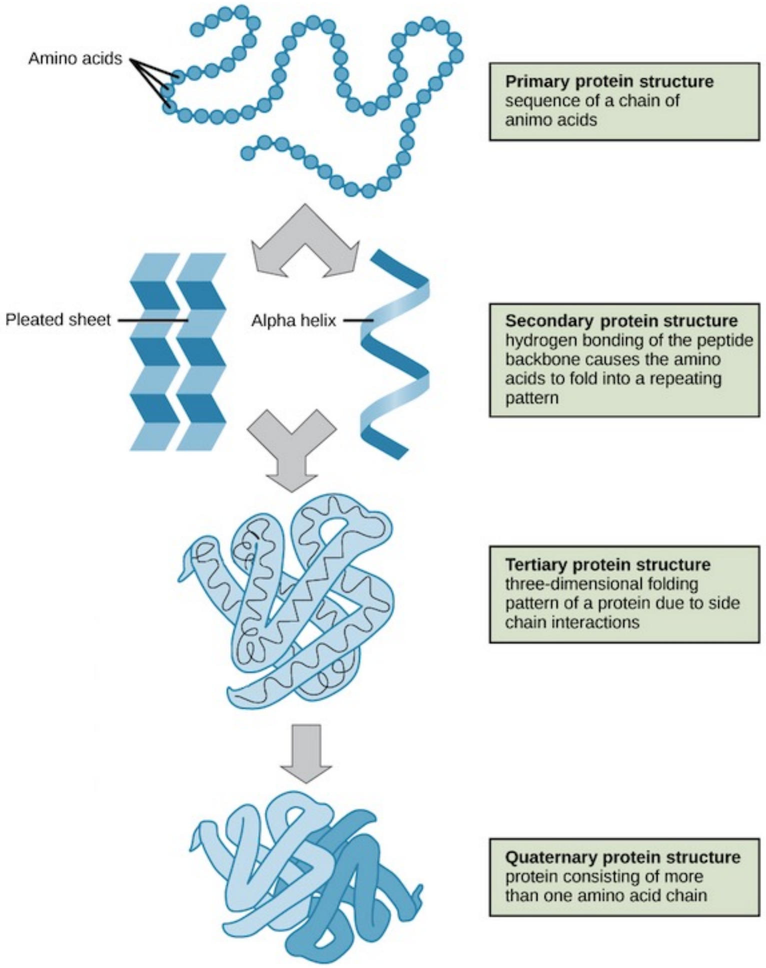
Proteins are chains of amino acids that assemble via amide bonds known as peptide linkages.

The uniqueness of different proteins is then determined by which amino acids it contains, how these amino acids are arranged in a chain, and further complex interactions the chain makes with itself and the environment.

Amino acids are the basic building blocks of proteins, and they serve as the nitrogenous backbones for compounds like neurotransmitters and hormones.

There are **approximately 20,000 unique protein encoding genes responsible for more than 100,000 unique proteins in the human body**.

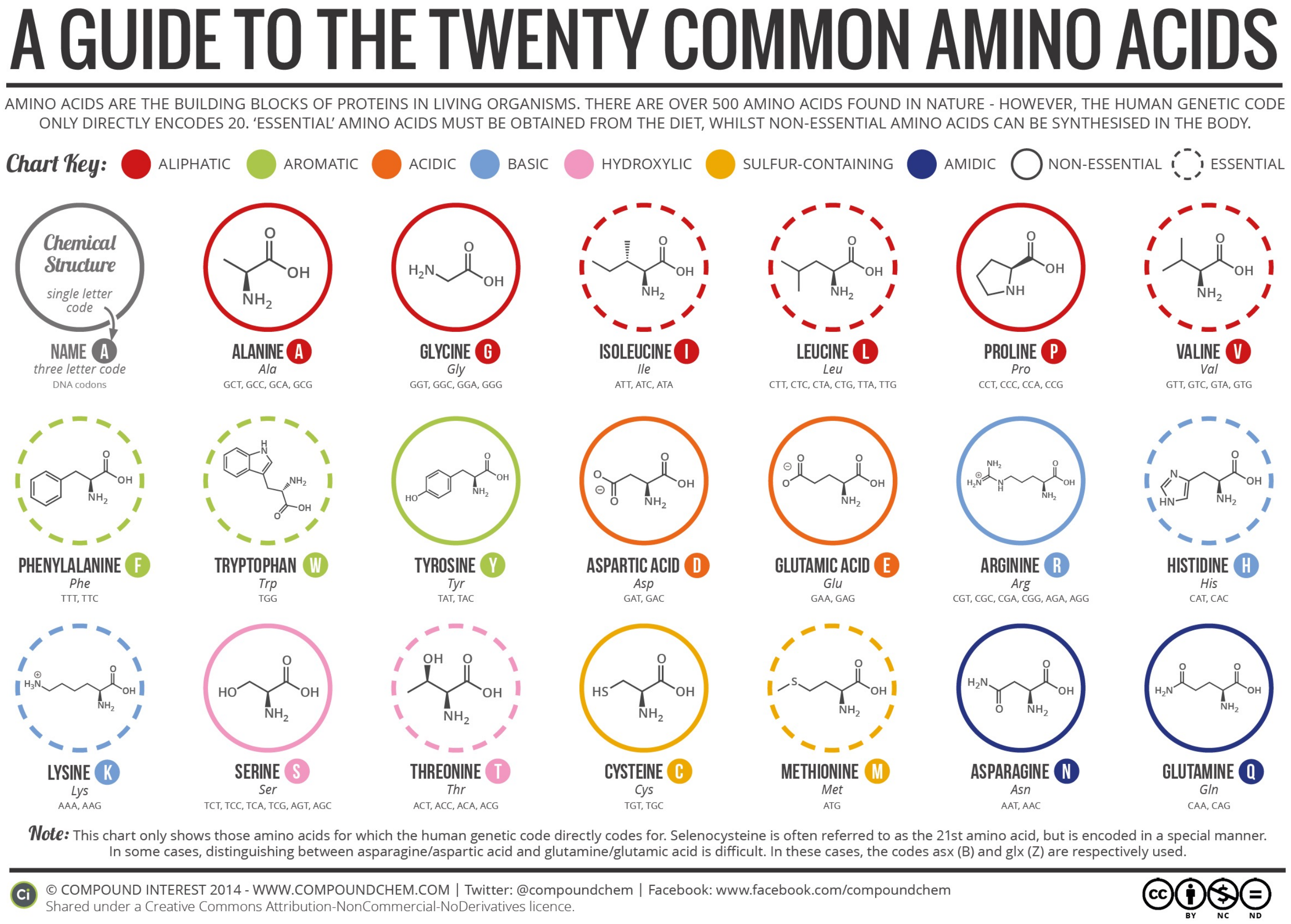
Source: “Biochemistry, Essential Amino Acids”, https://www.ncbi.nlm.nih.gov/books/NBK557845/. Accessed Apr. 5, 2022.


Amino acids
Primary protein structure
sequence of a chain of animo acids
Pleated sheet
Alpha helix
Secondary protein structure
hydrogen bonding of the peptide backbone causes the amino acids to fold into a repeating pattern
Tertiary protein structure
three-dimensional folding pattern of a protein due to side chain interactions
Quaternary protein structure
protein consisting of more than one amino acid chain

Source: Medium.

Amino Acids

- Although there are hundreds of amino acids found in nature, **only about 20 amino acids are needed to make all the proteins found in the human body and most other forms of life.** The amino acids selenocysteine and pyrrolysine are considered the 21st and 22nd amino acids, respectively. They are more recently discovered amino acids that may become incorporated into protein chains during ribosomal protein synthesis.



A GUIDE TO THE TWENTY COMMON AMINO ACIDS

AMINO ACIDS ARE THE BUILDING BLOCKS OF PROTEINS IN LIVING ORGANISMS. THERE ARE OVER 500 AMINO ACIDS FOUND IN NATURE - HOWEVER, THE HUMAN GENETIC CODE ONLY DIRECTLY ENCODES 20. 'ESSENTIAL' AMINO ACIDS MUST BE OBTAINED FROM THE DIET, WHILST NON-ESSENTIAL AMINO ACIDS CAN BE SYNTHESISED IN THE BODY.

Chart Key: ● ALIPHATIC ● AROMATIC ● ACIDIC ● BASIC ● HYDROXYLIC ● SULFUR-CONTAINING ● AMIDIC ○ NON-ESSENTIAL ◌ ESSENTIAL

Chemical Structure
single letter code
NAME A
three letter code
DNA codons

ALANINE A
Ala
GCT, GCC, GCA, GCG

GLYCINE G
Gly
GGT, GGC, GGA, GGG

ISOLEUCINE I
Ile
ATT, ATC, ATA

LEUCINE L
Leu
CTT, CTC, CTA, CTG, TTA, TTG

PROLINE P
Pro
CCT, CCC, CCA, CCG

VALINE V
Val
GTT, GTC, GTA, GTG

PHENYLALANINE F
Phe
TTT, TTC

TRYPTOPHAN W
Trp
TGG

TYROSINE Y
Tyr
TAT, TAC

ASPARTIC ACID D
Asp
GAT, GAC

GLUTAMIC ACID E
Glu
GAA, GAG

ARGININE R
Arg
CGT, CGC, CGA, CGG, AGA, AGG

HISTIDINE H
His
CAT, CAC

LYSINE K
Lys
AAA, AAG

SERINE S
Ser
TCT, TCC, TCA, TCG, AGT, AGC

THREONINE T
Thr
ACT, ACC, ACA, ACG

CYSTEINE C
Cys
TGT, TGC

METHIONINE M
Met
ATG

ASPARAGINE N
Asn
AAT, AAC

GLUTAMINE Q
Gln
CAA, CAG

Note: This chart only shows those amino acids for which the human genetic code directly codes for. Selenocysteine is often referred to as the 21st amino acid, but is encoded in a special manner. In some cases, distinguishing between asparagine/aspartic acid and glutamine/glutamic acid is difficult. In these cases, the codes asx (B) and glx (Z) are respectively used.

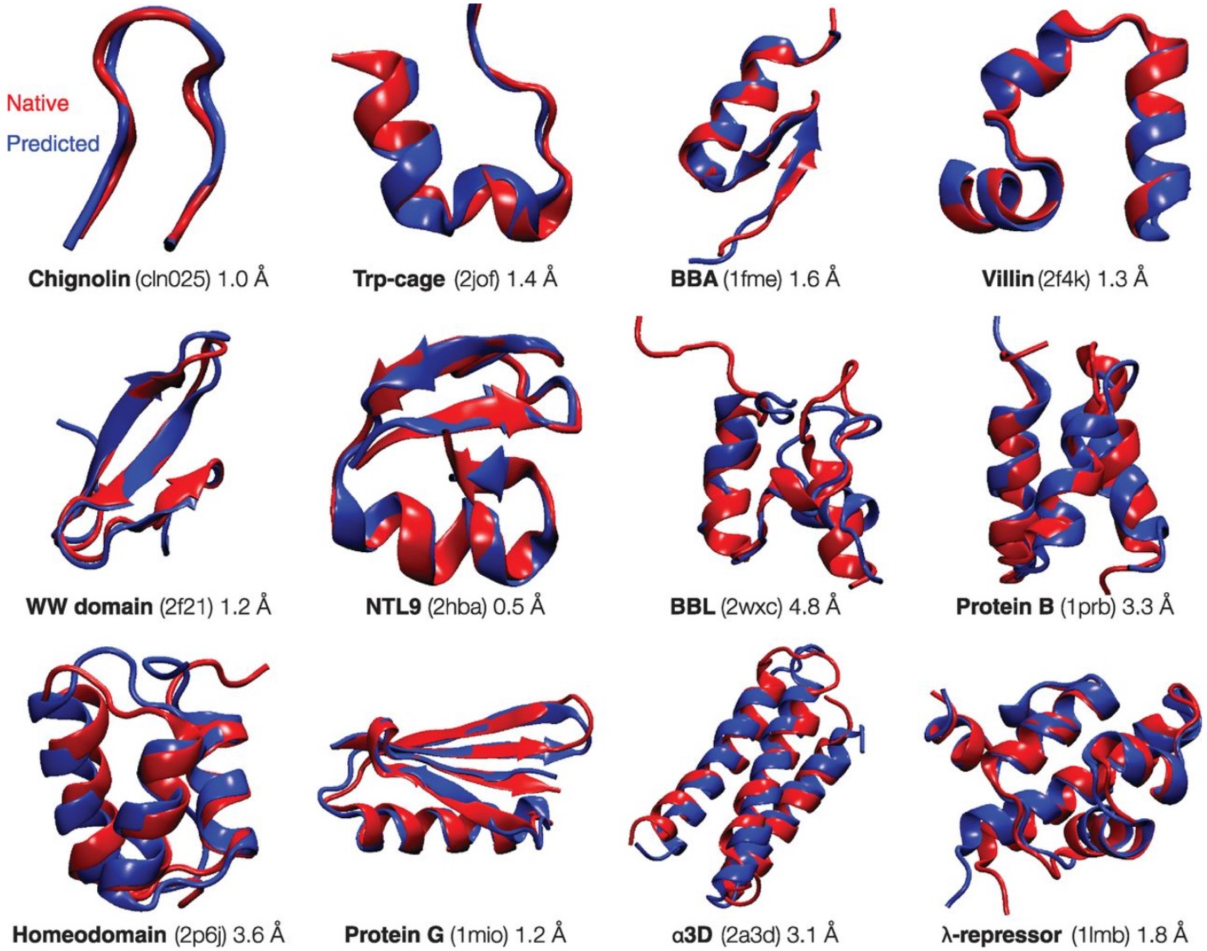
Source: Compound Interest.

Source: "Biochemistry, Essential Amino Acids", https://www.ncbi.nlm.nih.gov/books/NBK557845/. Accessed Apr. 5, 2022.

Protein Structure Prediction Methods

Protein Folding Problem

- **For decades, Scientists have been trying to find a method to reliably determine a protein’s structure just from its sequence of amino acids.** This grand scientific challenge is known as the protein folding problem.

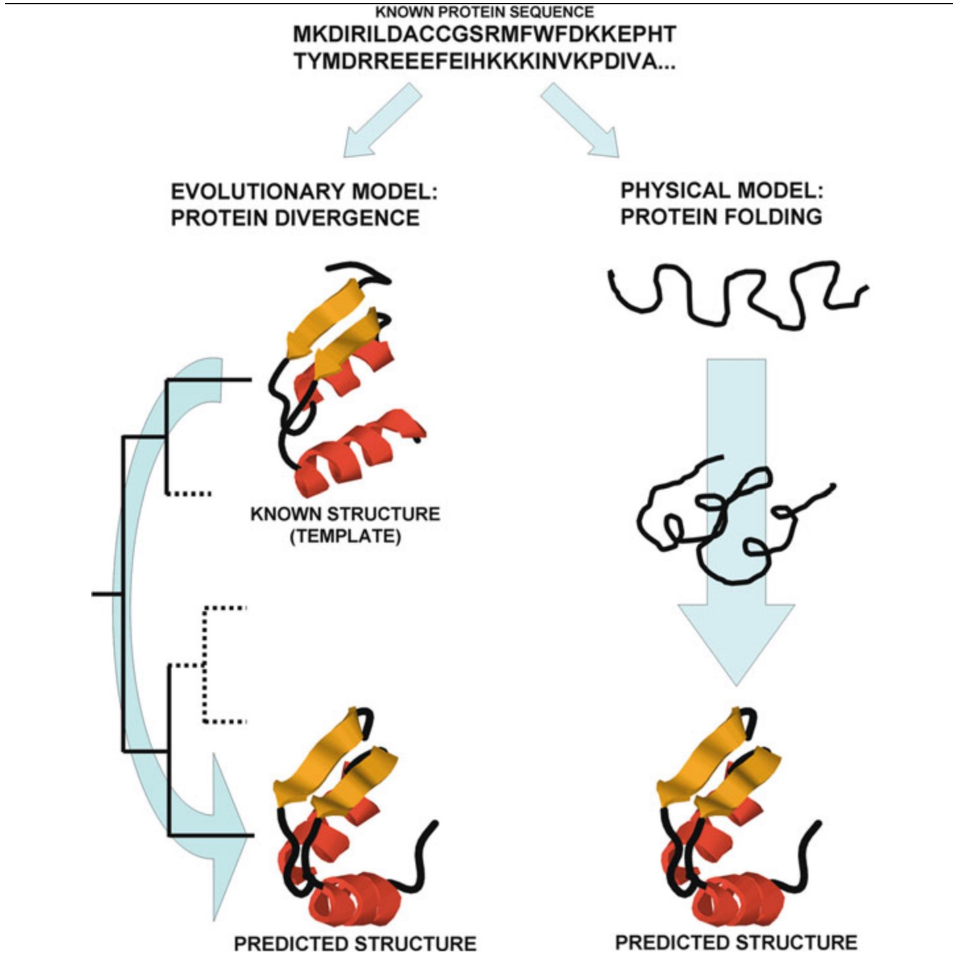


Source: Science.

Source: “AlphaFold”, https://www.deepmind.com/research/highlighted-research/alphafold. Accessed Mar 22, 2022.

Protein Prediction

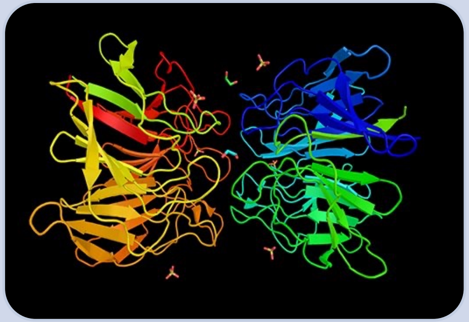
- The development of computational **methods to predict three-dimensional (3D) protein structures from the protein sequence** has proceeded along **two complementary paths** that focus on either the **physical interactions** or the **evolutionary history.**



Source: ResearchGate.

Source: “Highly accurate protein structure prediction with AlphaFold”, https://www.nature.com/articles/s41586-021-03819-2. Accessed Mar. 22, 2022.

Physical Interactions

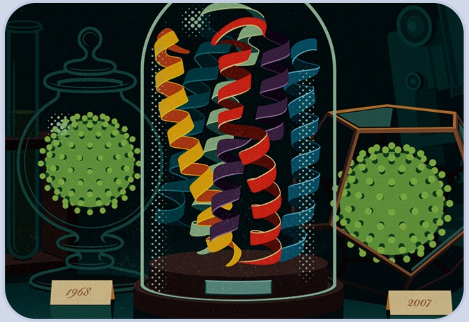


The **physical interaction** program **heavily integrates our understanding of molecular driving forces into either thermodynamic or kinetic simulation of protein physics or statistical approximations** thereof.

Although theoretically very appealing, **this approach has proved highly challenging** for even moderate-sized proteins due to the computational intractability of molecular simulation, the context dependence of protein stability and the difficulty of producing sufficiently accurate models of protein physics.

Source: "Highly accurate protein structure prediction with AlphaFold", https://www.nature.com/articles/s41586-021-03819-2. Accessed Mar. 22, 2022.

Evolutionary History



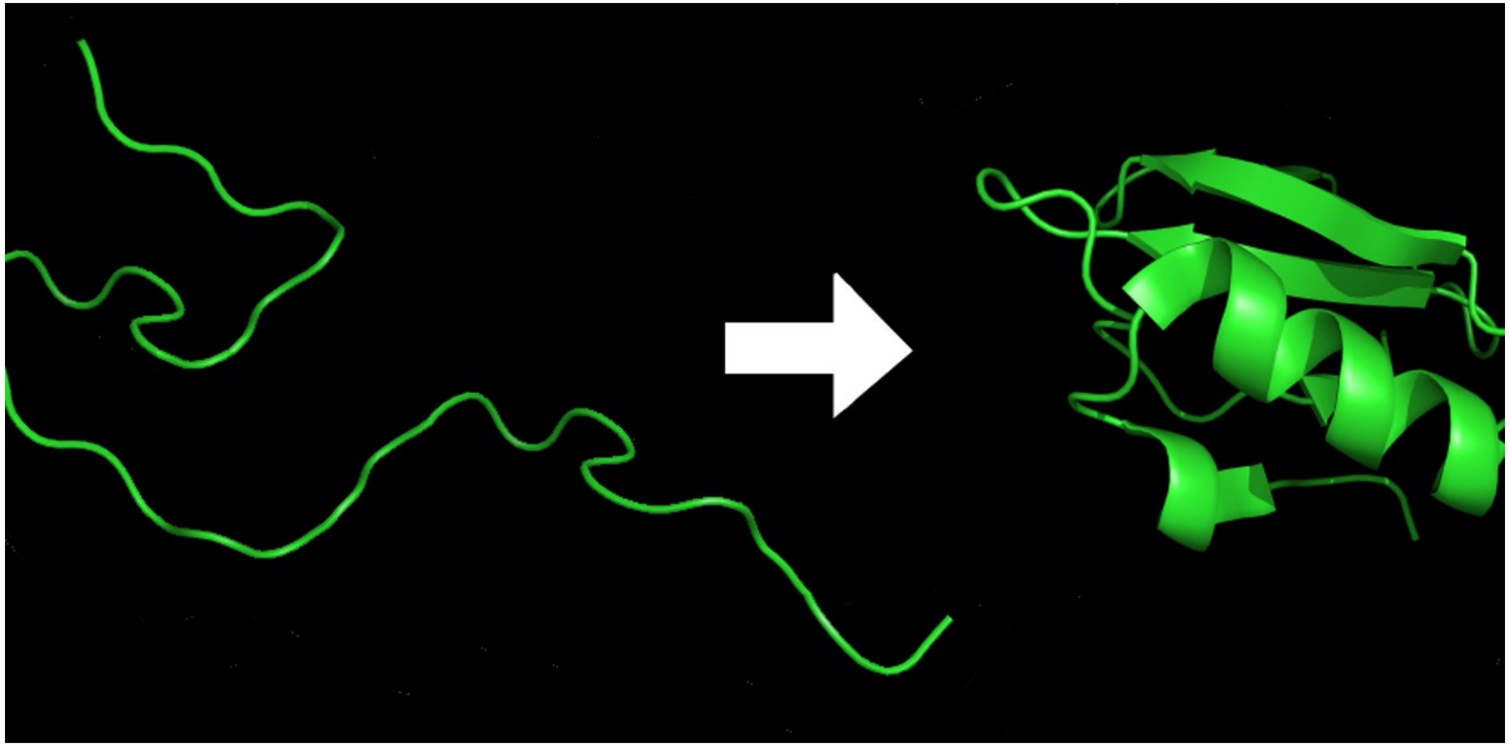
The **evolutionary program** has provided an alternative in recent years, in which the **constraints on protein structure are derived from bioinformatics analysis of the evolutionary history of proteins, homology to solved structures and pairwise evolutionary correlations**.

This **bioinformatics approach has benefited greatly from** the steady growth of experimental protein structures deposited in the **Protein Data Bank (PDB),** the explosion of **genomic sequencing** and the rapid development of **deep learning techniques to interpret these correlations**.

Source: “Highly accurate protein structure prediction with AlphaFold”, https://www.nature.com/articles/s41586-021-03819-2. Accessed Mar. 22, 2022.

Protein Prediction Limitations

- **Despite advances, contemporary physical and evolutionary-history-based approaches produce predictions** that are **far short of experimental accuracy** in the majority of cases in which a close homologue has not been solved experimentally and this has limited their utility for many biological applications.



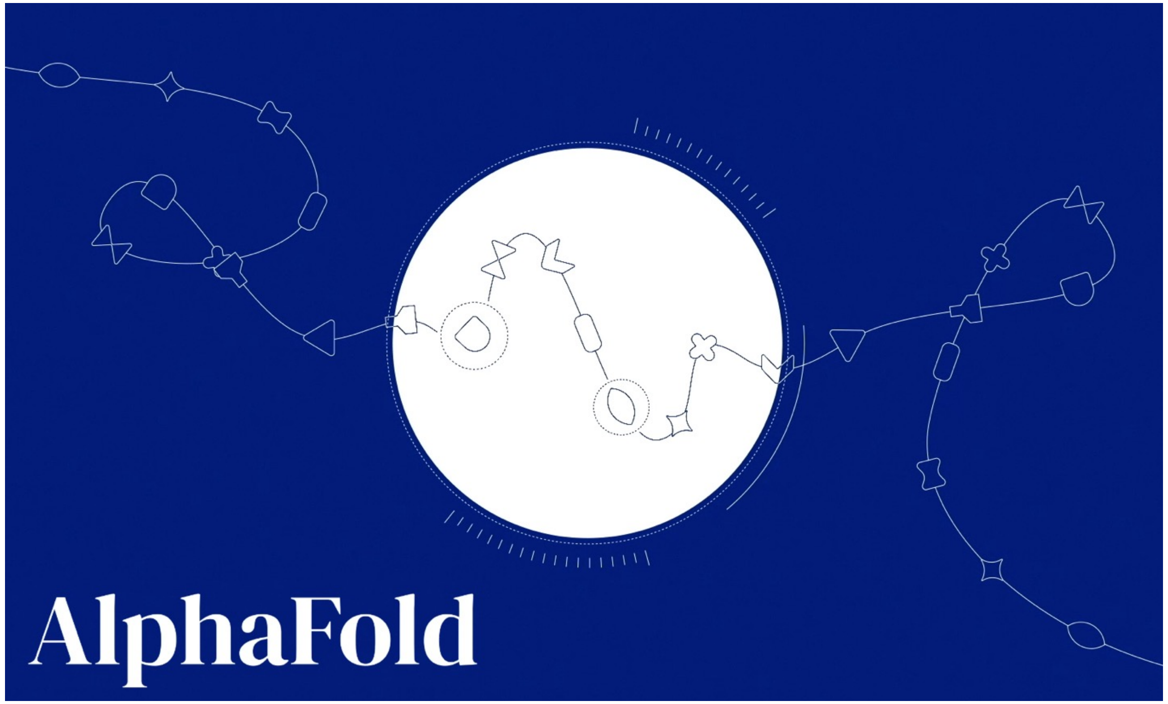
Source: The Conversation.

Source: "Highly accurate protein structure prediction with AlphaFold", https://www.nature.com/articles/s41586-021-03819-2. Accessed Mar. 22, 2022.

AlphaFold Overview

What is AlphaFold?

- AlphaFold is an AI system that can accurately predict 3D models of protein structures.
- **DeepMind** started working on the protein folding problem in 2016 and **have since created an AI system known as AlphaFold**.
- The system is taught by showing sequences and structures of around 100,000 known proteins. The latest version (AlphaFold2) can now predict the shape of the protein, at scale and in minutes, down to atomic accuracy.



Source: TechCrunch.

Source: "AlphaFold", https://www.deepmind.com/research/highlighted-research/alphafold. Accessed Mar. 22, 2022.

What Has AlphaFold Done?

- AlphaFold greatly **improves the accuracy of structure prediction by incorporating novel neural network architectures and training procedures based on the evolutionary, physical, and geometric constraints of protein structures**.
- AlphaFold demonstrates:
  - A new architecture to jointly embed multiple sequence alignments (MSAs) and pairwise features
  - A new output representation and associated loss that enable accurate end-to-end structure prediction
  - A new equivariant attention architecture
  - Use of intermediate losses to achieve iterative refinement of predictions
  - Masked MSA loss to jointly train with the structure
  - Learning from unlabeled protein sequences using self-distillation and self-estimates of accuracy

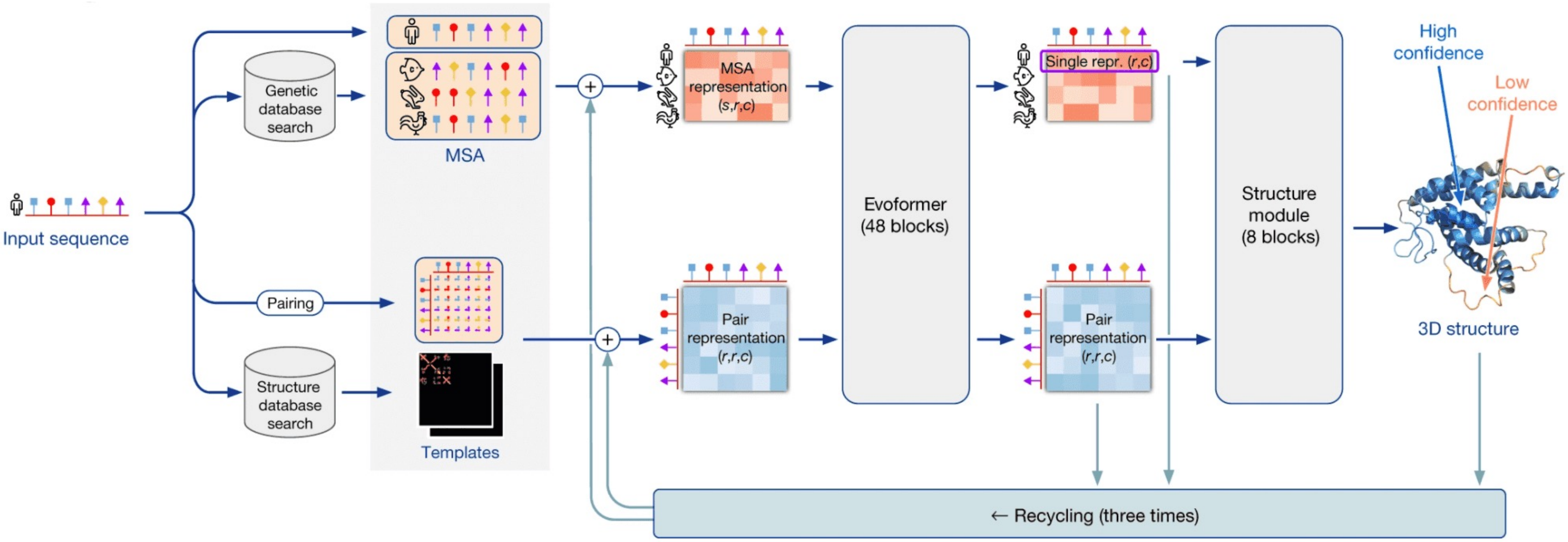
Source: “Highly accurate protein structure prediction with AlphaFold”, https://www.nature.com/articles/s41586-021-03819-2. Accessed Mar. 22, 2022.

How AlphaFold Works

- AlphaFold takes multisequence alignment -which is thought to be evolutionarily related to target.
- AlphaFold then pairs the sequence with an array representing residue pairs in the target protein.
    - The program can make use of templates.
- Evoformer blocks extract information about the relationship between residues.
    - Its Role is to gradually build picture of relationship between protein residues.
    - During this process MSA and pair representations are repeatedly updated until they provide highly informative data about structure.
- Structure module predicts rotation and translation to place each residue to place them in 3D space.
    - Residues are treated as separate objects.
    - Side chains predicted on small simple network.
- Final structure ran through relaxation step to fix any violations in structure.
    - Restrain energy minimization

Source: "How to interpret AlphaFold structures", https://www.youtube.com/watch?v=UqeQfRDA8Yk. Accessed Apr. 12, 2022.

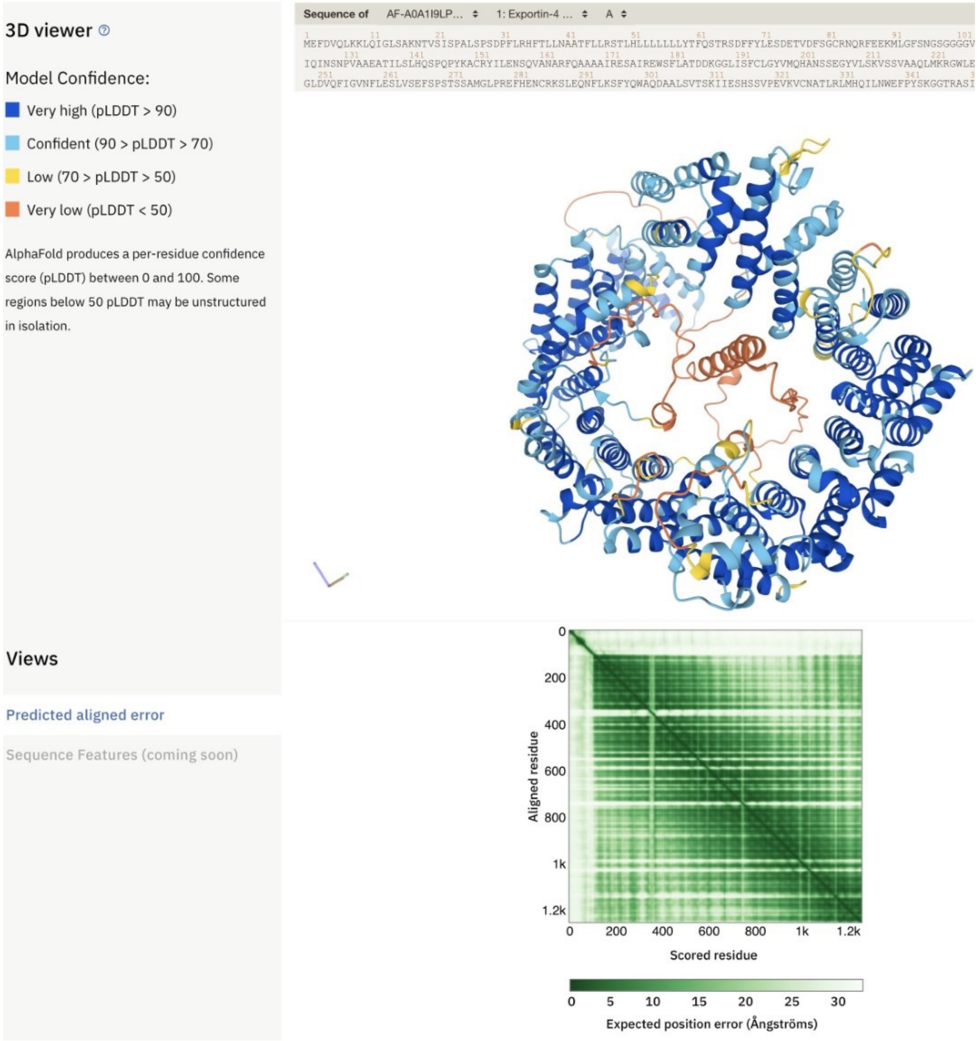
How AlphaFold Works Visual


Input sequence
Genetic database search
Pairing
Structure database search
MSA
Templates
MSA representation (s,r,c)
Pair representation (r,r,c)
Evoformer (48 blocks)
Single repr. (r,c)
Pair representation (r,r,c)
Structure module (8 blocks)
High confidence
Low confidence
3D structure
← Recycling (three times)

Source: borisburkov.net.

How to Interpret AlphaFold Structures

- **In addition to the predicted structure, two confidence metrics are presented.**
- **Predicted LDDT- measure % of correct predicted interatomic distances vs full compared to predicted**
  - Results-high score for locally correct-indicates confidence in local structure (0-100) and shown by color confidence bands raw is in b factors in structure files
  - PLDDT Pitfall- assume because high confidence, on all individual domains must mean AF confident about relative positions- note position can change
- **Predicted Aligned Error (PAE)- expected position error residue x if predicted and true structures aligned on residue y**
  - PAE should be used where pairwise confidence is relevant- interpret domain position in multidomain protein
  - Confident relative position closer to 0 dark green, light green is higher
  - Since confidence higher within domain vs between domain you see squares on plot



3D viewer

Model Confidence:

🟦 Very high (pLDDT > 90)

🟦 Confident (90 > pLDDT > 70)

🟨 Low (70 > pLDDT > 50)

🟧 Very low (pLDDT < 50)

AlphaFold produces a per-residue confidence score (pLDDT) between 0 and 100. Some regions below 50 pLDDT may be unstructured in isolation.

Views

Predicted aligned error

Sequence Features (coming soon)

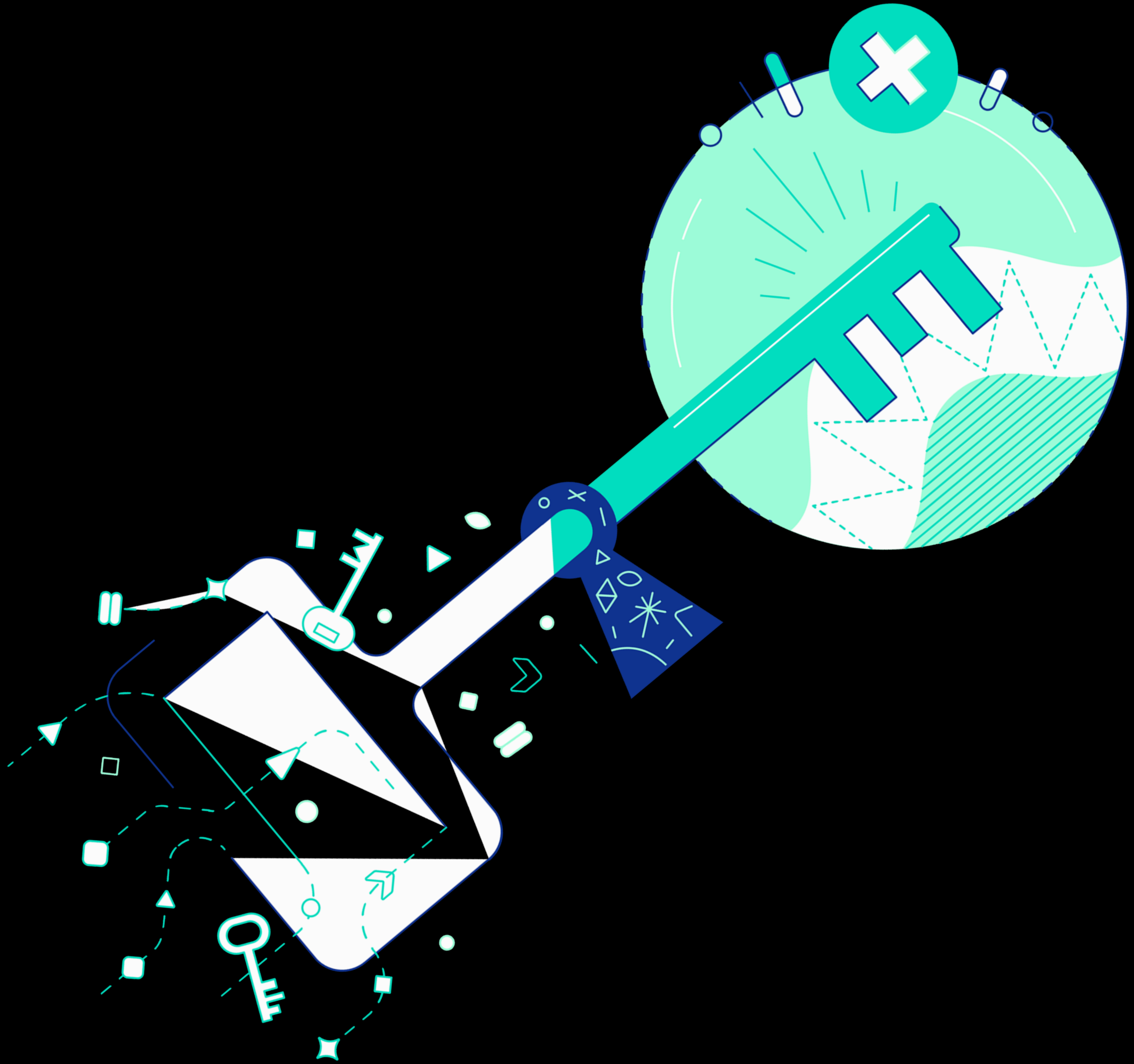
Sequence of AF-A0A1I9LP... 1: Exportin-4 ... A

```
1         11        21        31        41        51        61        71        81        91        101
MEFDVQLKKLQIGLSAKNTVSISPALSPSDPFLRHFTLLNAATFLLRSTLHLLLLLLLYTFQSTRSDFFYLESDETVDFSGCRNQRFEEKMLGFSNGSGGGV
  131       141       151       161       171       181       191       201       211       221
IQINSNPVAAEATILSLHQSPQPYKACRYILENSQVANARFQAAAAIRESAIREWSFLATDDKGGLISFCLGYVMQHANSSEGYVLSKVSSVAAQLMKRGWLE
 251       261       271       281       291       301       311       321       331       341
GLDVQFIGVNFLESLVSEFSPSTSSAMGLPREFHENCRKSLEQNFLKSFYQWAQDAALSVTSKIIESHSSVPEVKVCNATLRLMHQILNWEFPYSKGGTRASI
```

Source: "How to interpret AlphaFold structures", https://www.youtube.com/watch?v=UqeQfRDA8Yk. Accessed Apr. 12, 2022.

Source: FEBS Network.

Where to Access AlphaFold



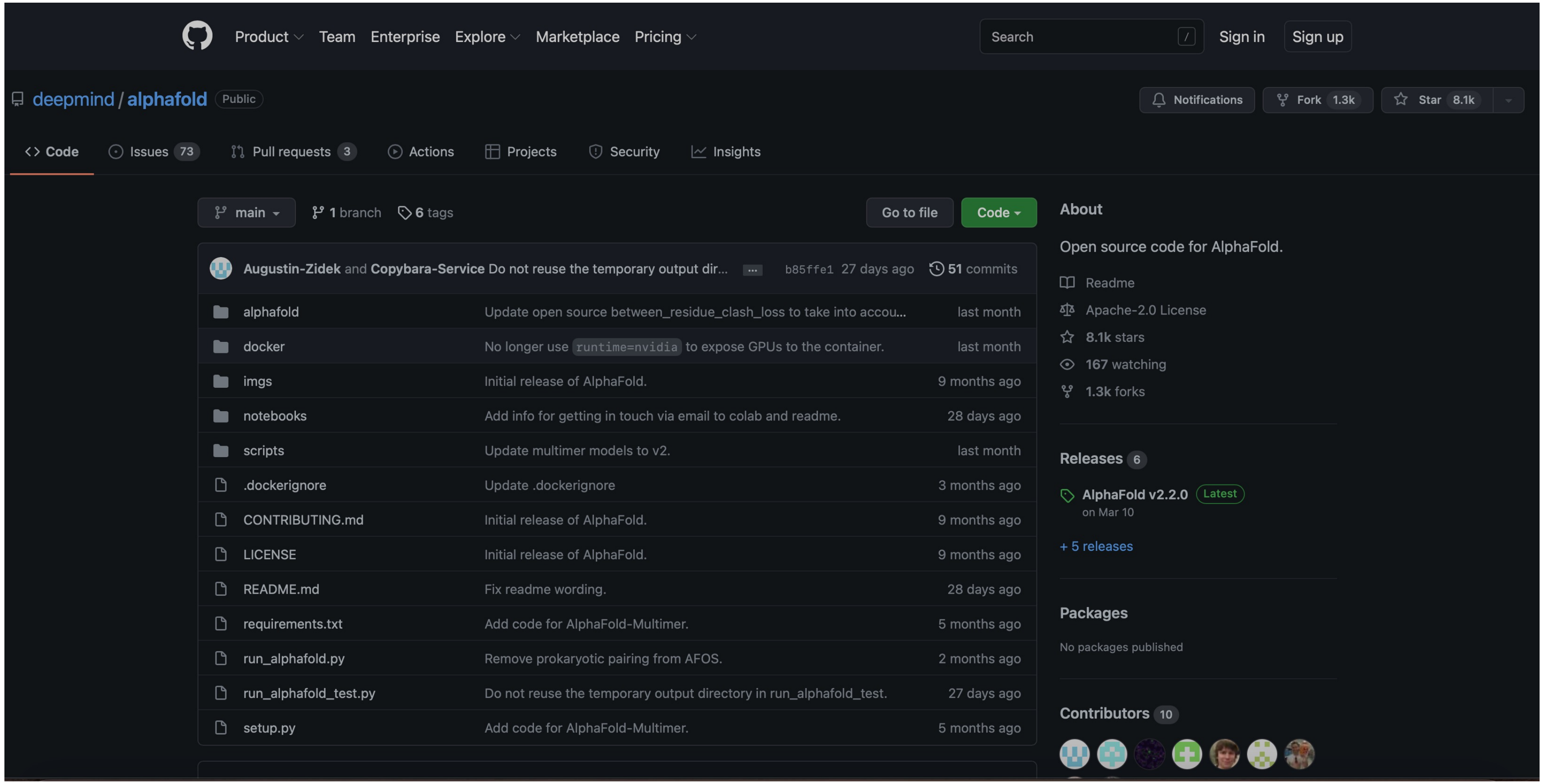
AlphaFold Source Code

- Deep mind made the **source code open source for those wishing to use the AlphaFold2**.
- This open-source code provides an implementation of the AlphaFold v2.0 system. It **allows users to predict the 3-D structure of arbitrary proteins with unprecedented accuracy**.
- AlphaFold v2.0 is a completely new model that was entered in the CASP14 assessment and published in Nature (Jumper et al. 2021). The package contains source code, trained weights, and an inference.
- Note: The download is around 428GB and total size when unzipped is 2.2TB. A large enough hard drive, ample bandwidth, and time is required to download.

Source: “Open source AlphaFold”, https://www.deepmind.com/open-source/alphafold. Accessed Apr. 4, 2022.; “kuixu/alphafold”, https://github.com/kuixu/alphafold. Accessed Apr. 4, 2022

AlphaFold Source Code Link

- **Website**: https://github.com/deepmind/alphafold/



deepmind / alphafold Public

Augustin-Zidek and Copybara-Service Do not reuse the temporary output dir... b85ffe1 27 days ago 51 commits

alphafold	Update open source between_residue_clash_loss to take into accou...	last month
docker	No longer use `runtime=nvidia` to expose GPUs to the container.	last month
imgs	Initial release of AlphaFold.	9 months ago
notebooks	Add info for getting in touch via email to colab and readme.	28 days ago
scripts	Update multimer models to v2.	last month
.dockerignore	Update .dockerignore	3 months ago
CONTRIBUTING.md	Initial release of AlphaFold.	9 months ago
LICENSE	Initial release of AlphaFold.	9 months ago
README.md	Fix readme wording.	28 days ago
requirements.txt	Add code for AlphaFold-Multimer.	5 months ago
run_alphafold.py	Remove prokaryotic pairing from AFOS.	2 months ago
run_alphafold_test.py	Do not reuse the temporary output directory in run_alphafold_test.	27 days ago
setup.py	Add code for AlphaFold-Multimer.	5 months ago

About

Open source code for AlphaFold.

Readme
Apache-2.0 License
8.1k stars
167 watching
1.3k forks

Releases 6

AlphaFold v2.2.0 Latest
on Mar 10

+ 5 releases

Packages

No packages published

Contributors 10

Source: Github.

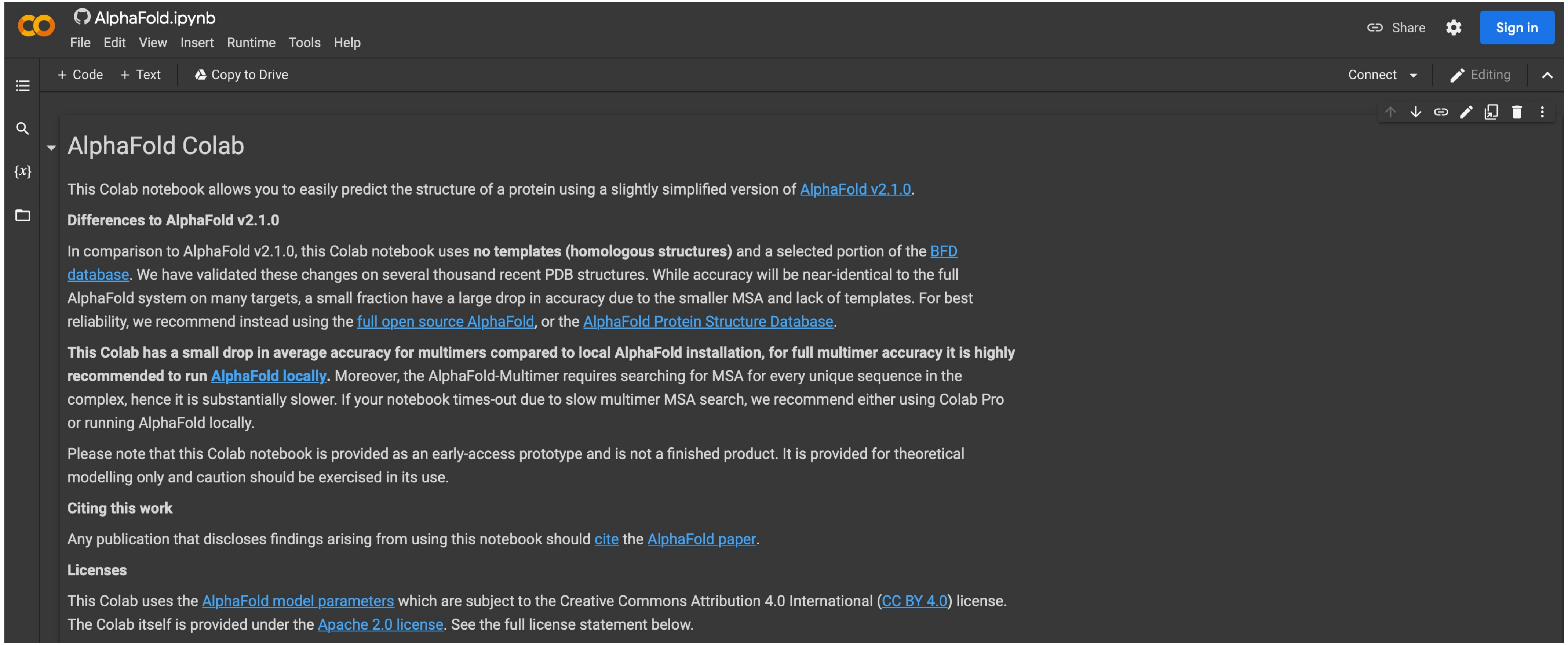
AlphaFold Colab Notebook

- The Collab notebook to run individual sequences. Should be close to accuracy for full code but this is a smaller program with lack of templates. Source code is best for highest accuracy.
- **Advantages**
  - Can target any sequence of interest
  - No coding or install required since the program is on the cloud
- **Disadvantages**
  - Not suitable for large prediction jobs
  - May be unreliable for long sequences
  - Limited ability to influence prediction (more for single sequences rather than batch jobs)
  - Size shouldn’t be greater than 800 residues since the program can't guarantee spec of cloud machine
  - Waits can be lengthy especially when searching large genetics databases as this step could take hours

Source: “AlphaFold Colab”, https://colab.research.google.com/github/deepmind/alphafold/blob/main/notebooks/AlphaFold.ipynb. Accessed Apr. 4, 2022.; “How to interpret AlphaFold structures”, https://www.youtube.com/watch?v=UqeQfRDA8Yk. Accessed Apr. 12, 2022.

AlphaFold Colab Notebook

- **Website:** https://colab.research.google.com/github/deepmind/alphafold/blob/main/notebooks/AlphaFold.ipynb



AlphaFold Colab

This Colab notebook allows you to easily predict the structure of a protein using a slightly simplified version of AlphaFold v2.1.0.

Differences to AlphaFold v2.1.0

In comparison to AlphaFold v2.1.0, this Colab notebook uses **no templates (homologous structures)** and a selected portion of the BFD database. We have validated these changes on several thousand recent PDB structures. While accuracy will be near-identical to the full AlphaFold system on many targets, a small fraction have a large drop in accuracy due to the smaller MSA and lack of templates. For best reliability, we recommend instead using the full open source AlphaFold, or the AlphaFold Protein Structure Database.

This Colab has a small drop in average accuracy for multimers compared to local AlphaFold installation, for full multimer accuracy it is highly recommended to run AlphaFold locally. Moreover, the AlphaFold-Multimer requires searching for MSA for every unique sequence in the complex, hence it is substantially slower. If your notebook times-out due to slow multimer MSA search, we recommend either using Colab Pro or running AlphaFold locally.

Please note that this Colab notebook is provided as an early-access prototype and is not a finished product. It is provided for theoretical modelling only and caution should be exercised in its use.

Citing this work

Any publication that discloses findings arising from using this notebook should cite the AlphaFold paper.

Licenses

This Colab uses the AlphaFold model parameters which are subject to the Creative Commons Attribution 4.0 International (CC BY 4.0) license. The Colab itself is provided under the Apache 2.0 license. See the full license statement below.

Source: AlphaFold.ipynb.

AlphaFold Protein Structure Database

Deep Mind has partnered up with EMBL's European Bioinformatics Institute to create the AlphaFold **Protein Structure Database** to make the protein predictions from their model **freely available to the scientific community**.

The **initial release of the database covers all 20,000 proteins in the human proteome**, **along with the proteomes of several other biologically significant organisms**, from E.coli to yeast, and from the fruit fly to the mouse.

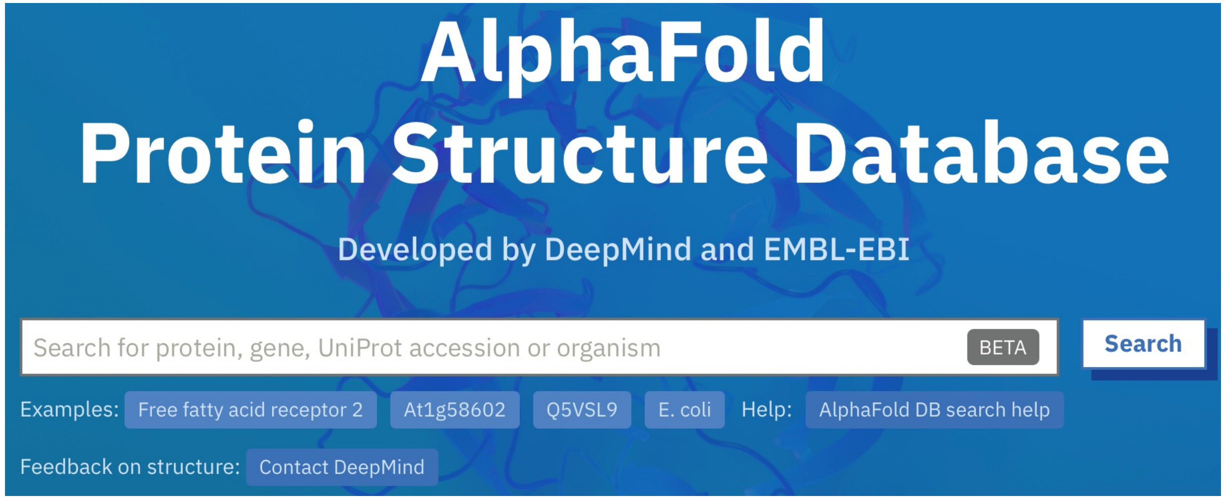
Source: "AlphaFold Protein Structure Database", https://www.deepmind.com/open-source/alphafold-protein-structure-database. Accessed Apr. 4, 2022.

AlphaFold Protein Structure Database Continued

The AphaFold Protein Structure Database will continue to expand over time, with updates on DeepMind and EMBL-EBI's social channels.

In the coming months, Deep Mind plans to expand the database to include all 100 million proteins catalogued in the UniRef90 Database.

Website to access database: https://alphafold.ebi.ac.uk



AlphaFold
Protein Structure Database

Developed by DeepMind and EMBL-EBI

Search for protein, gene, UniProt accession or organism
BETA
Search
Examples: Free fatty acid receptor 2
At1g58602
Q5VSL9
E. coli
Help: AlphaFold DB search help
Feedback on structure: Contact DeepMind

Source: EMBL - EBI.

Source: "AlphaFold Protein Structure Database", https://www.deepmind.com/open-source/alphafold-protein-structure-database. Accessed Apr. 4, 2022.

AlphaFold Limitations

Limitations

- **Interactions with partner proteins or multimers are currently not in the database.**
- **AI also does not predict several other important aspects of protein structures: metal ions, cofactors, and other ligands. Post-translational** modifications, such as glycosylation or phosphorylation, or DNA, RNA, and their complexes, are also absent. In addition, amino acid side chains are not always accurately placed. Each of those features may be crucial for protein function, and many of these are necessary for the integrity of the fold.
- **AlphaFold has not learned from ligands and is actually not aware of the actual energy minima that are essential for folding in real life.**
- **In reality, Alpha Fold has not solved the folding problem as it would occur in solution or in a cell, but it has provided a practical solution**: It has learned the results of folding at the amino acid residue contact level and can therefore accurately predict a single-chain hemoglobin fold that would never exist on its own or in the absence of the heme cofactor in nature.
- Another limitation of AlphaFold predictions is that only a single state is predicted, even if hints for multiple states and dynamic behavior are in the data, like for USP7.

Source: “AI revolutions in biology”, https://www.embopress.org/doi/full/10.15252/embr.202154046. Accessed Apr. 12, 2022.

AlphaFold Prediction Ranking Limitations

- **AlphaFold does not currently predict 3D structures for protein-protein or protein-DNA/RNA/ligand complexes**. In some cases, the single-chain prediction may correspond to the structure adopted in a complex. In other cases (especially if the protein is structured only upon binding partner molecules) the missing context from surrounding molecules may lead to an uninformative prediction.
- **Proteins are dynamic systems and adopt different structures depending on their environment or state within a functional cycle**. Where a protein is known to have multiple conformations AlphaFold will usually only produce one of them.
- **For regions that are intrinsically disordered or unstructured in isolation, AlphaFold is expected to produce a low-confidence prediction and the predicted structure will have an extended, ribbon-like appearance.**
- AlphaFold2 remains an ensemble-based prediction method, predicting structures from families of related proteins instead of individual sequences. This may make it insensitive to sequence-specific structural changes that arise from mutations and suggests that it may not be effective when proteins have few homologues or are human-designed.

Source: "Great expectations – the potential impacts of AlphaFold DB", https://www.embl.org/news/science/alphafold-potential-impacts/. Accessed Apr. 12, 2022.; "Protein structure prediction by AlphaFold2: are attention and symmetries all you need?", https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8329862/. Accessed Apr 12, 2022.

DeepMind’s Perspective on AlphaFold

- **Future of AlphaFold**
  - AlphaFold has limits
  - Does not replace experiment
  - Can't provide all info from experimental structure
  - The model does not predict the position of non-protein components such as DNA, RNA, ligands, and post translational modifications
  - Can't reliably generate particular confirmation for dynamic protein
  - Don’t expect AlphaFold to produce informative prediction for destabilizing point mutations
- **Optimism**
  - The field can tackle broader questions
  - People are building on and experimenting with Alphafold
  - Deep Mind has noticed many positive mentions on Twitter describing plans to build and expand on this program
  - Commercial use permitted on database information

Source: “How to interpret AlphaFold structures”, https://www.youtube.com/watch?v=UqeQfRDA8Yk. Accessed Apr. 12, 2022.

Companies Using AlphaFold For Oncology

Avalon GloboCare Corp. (NasdaqCM: AVCO)

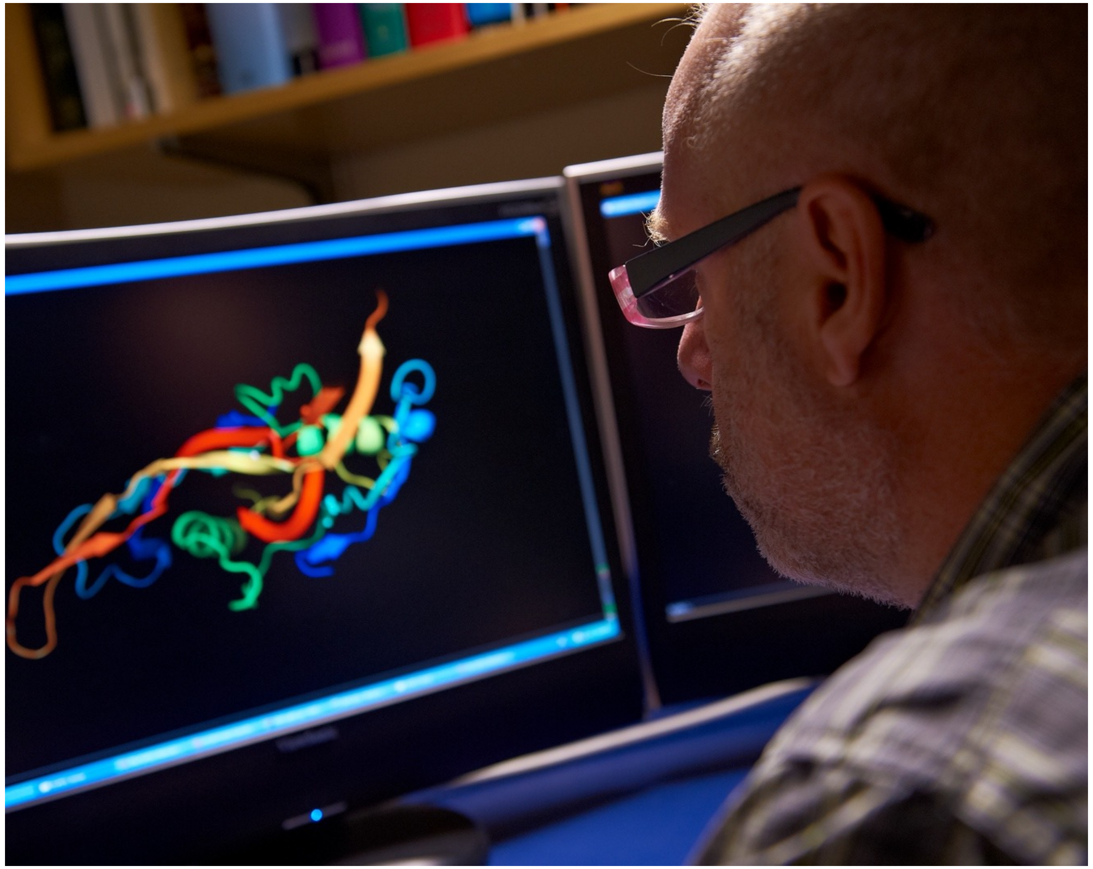
- Avalon and research partner Massachusetts Institute of Technology (MIT) combine their protein design **QTY Code tech with AlphaFold2 to accurately predict 3D structure of protein receptors that have potential as cell therapy targets**. The **goal of this system is to accelerate and advance Avalon's capabilities in developing novel targets for immuno-oncology and cellular medicine.** QTY Code is a protein-design platform that can turn water-insoluble transmembrane receptor proteins into water-soluble proteins, enabling their use in many clinical applications, including drug development.
- This program has shown success through decoy receptors which function to soak up excess chemokines and cytokines produced in the body during a potentially fatal 'cytokine storm.' This is useful for COVID-19 and cancer patient CAR T related cytokine storms.
- The researchers used the QTY code technology to design water-soluble versions of chemokine receptors—water-insoluble proteins involved in cytokine storms, cancer, autoimmune diseases and important drug targets—and then used AlphaFold to accurately predict the structures of these clinically important proteins.

Source: Avalon GloboCare Corp.

Source: "Avalon GloboCare Advances AI-Enhanced Protein Design Technology for Cellular Therapy Development", https://www.globenewswire.com/news-release/2021/12/09/2349316/0/en/Avalon-GloboCare-Advances-AI-Enhanced-Protein-Design-Technology-for-Cellular-Therapy-Development.html. Accessed Apr. 5, 2022.

Fox Chase Cancer Center

- **Statement from Roland L. Dunbrack Jr.,PhD, professor and director of the Molecular Modeling Facility at Fox Chase.**
  - **"What we can do with AlphaFold, specifically in cancer, is predict the structure of the protein in which a mutation occurs**. We can see where that mutation is, see how the mutated protein interacts with other molecules, and **ultimately get an idea about the mechanism, or why that mutation causes a problem for that cell**. **Once you have a mechanism** for how a mutation causes a problem, **you can begin to think about strategies for stopping that problem**," he said.
  - **The lab recently used it to predict how the breast cancer associated protein BRCA1 interacts with one of its partners during DNA repair.** What they found was that that the AlphaFold model was spot on when compared an experimental structure of the complex of these two proteins that came out.
  - As an early adopter, they were unable to set all operational parameters that became available with the new system. Now the new system opens new possibilities.

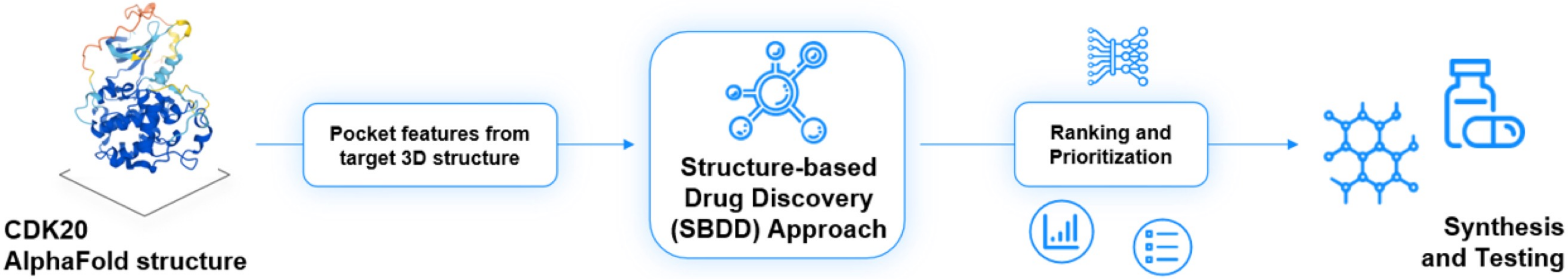


Source: Fox Chase Cancer Center on Twitter.

Source: "AlphaFold System for Protein Structure Prediction Could Enhance Development of New Drugs", https://www.foxchase.org/news/2021-11-22-alphafold-system-for-protein-structure-prediction-could-enhance-development-of-new-drugs. Accessed Apr. 5, 2022.

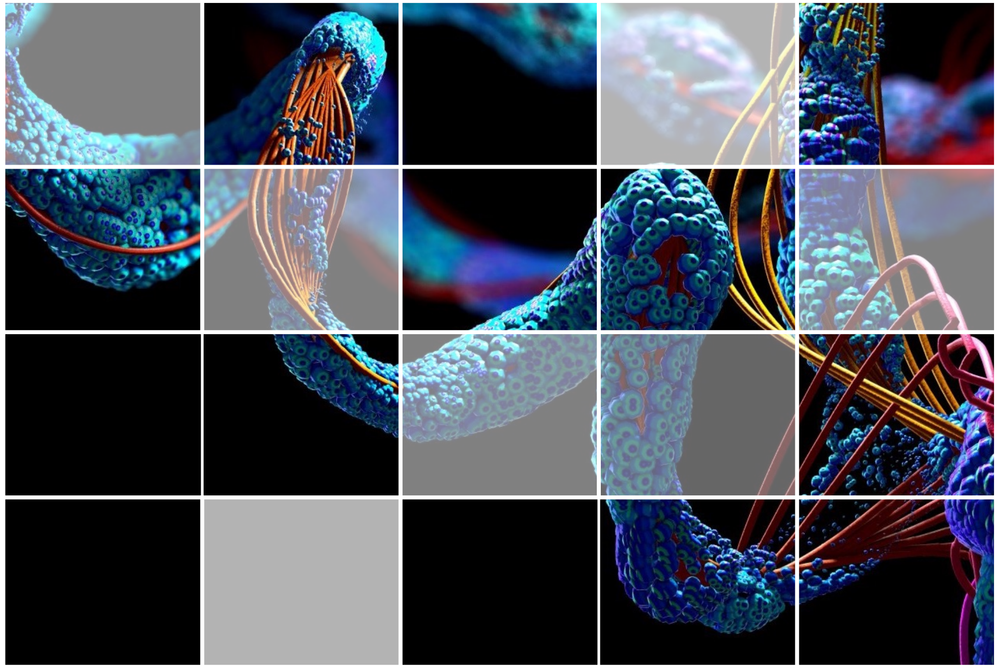
Insilico Medicine

- **Insilico Medicine in Shanghai, China announced the development of a potential drug that combines AlphaFold- predicted protein structure with AI-Powered drug discovery platforms.**
- The **drug is focused on treating a form a liver cancer called hepatocellular carcinoma**, which currently lacks effective treatments. From the list of the top 20 promising proteins, the company selected CDK20, a protein that regulates the way a cell grows and divides.
- The firm then used an Ai powered engine called Chemistry42 to generate molecules that could latch onto CDK20 and disable it. They are studying these molecules with one that seems particularly promising. The entire process to find promising molecules took 30 days.


CDK20
AlphaFold structure
Pocket features from target 3D structure
Structure-based Drug Discovery (SBDD) Approach
Ranking and Prioritization
Synthesis and Testing

Source: "AI Reduces Timescale of Early Drug Development to Weeks", https://www.discovermagazine.com/the-sciences/ai-reduces-timescale-of-early-drug-development-to-weeks, Accessed Apr. 4, 2022.; "AlphaFold Accelerates Artificial Intelligence Powered Drug Discovery: Efficient Discovery of a Novel Cyclin-dependent Kinase 20 (CDK20) Small Molecule Inhibitor", https://arxiv.org/pdf/2201.09647.pdf, Accessed Apr. 4, 2022.

Thanks!



Address：800 3rd Ave 11th floor, New York, NY 10022

Tel：+1(646) 801-2803

Email： info@dgipl.com

Website： www.dgipl.com