BINF 731

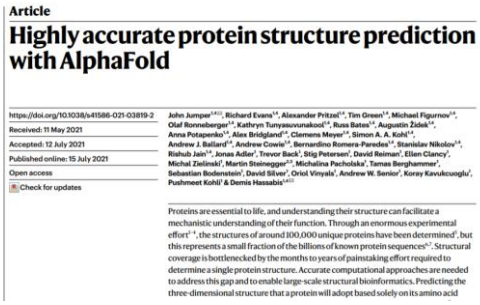
Protein Structure Analysis

Iosif Vaisman

2021

AlphaFold

Nature, volume 596, pages 583–589 (2021)



Article

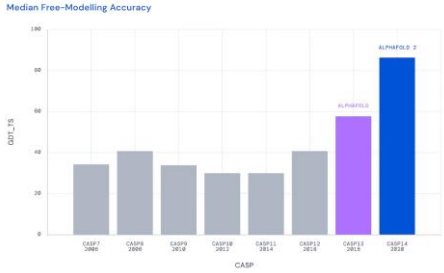
Highly accurate protein structure prediction with AlphaFold

https://doi.org/10.1038/s41586-021-03819-2
Received: 11 May 2021
Accepted: 12 July 2021
Published online: 15 July 2021
Open access
Check for updates

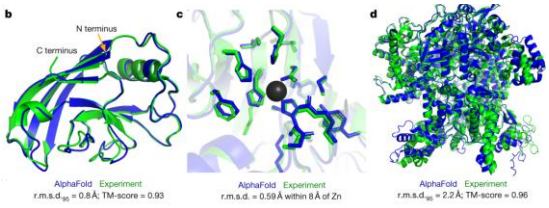
John Jumper[1,2☒], Richard Evans[1,2], Alexander Pritzel[1,2], Tim Green[1,2], Michael Figurnov[1,2], Olaf Ronneberger[1,2], Kathryn Tunyasuvunakool[1,2], Russ Bates[1,2], Augustin Žídek[1,2], Anna Potapenko[1,2], Alex Bridgland[1,2], Clemens Meyer[1,2], Simon A. A. Kohl[1,2], Andrew J. Ballard[1,2], Andrew Cowie[1,2], Bernardino Romera-Paredes[1,2], Stanislav Nikolov[1,2], Rishub Jain[1,2], Jonas Adler[1], Trevor Back[1], Stig Petersen[1], David Reiman[1], Ellen Clancy[1], Michal Zielinski[1], Martin Steinegger[2,3], Michalina Pacholska[1], Tamas Berghammer[1], Sebastian Bodenstein[1], David Silver[1], Oriol Vinyals[1], Andrew W. Senior[1], Koray Kavukcuoglu[1], Pushmeet Kohli[1] & Demis Hassabis[1,2☒]

Proteins are essential to life, and understanding their structure can facilitate a mechanistic understanding of their function. Through an enormous experimental effort[1–4], the structures of around 100,000 unique proteins have been determined[5], but this represents a small fraction of the billions of known protein sequences[6,7]. Structural coverage is bottlenecked by the months to years of painstaking effort required to determine a single protein structure. Accurate computational approaches are needed to address this gap and to enable large-scale structural bioinformatics. Predicting the three-dimensional structure that a protein will adopt based solely on its amino acid

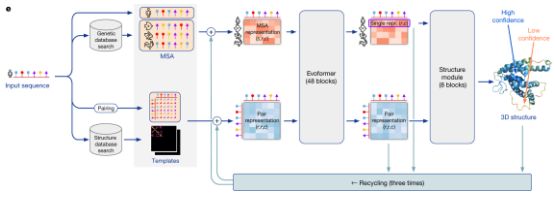
AlphaFold



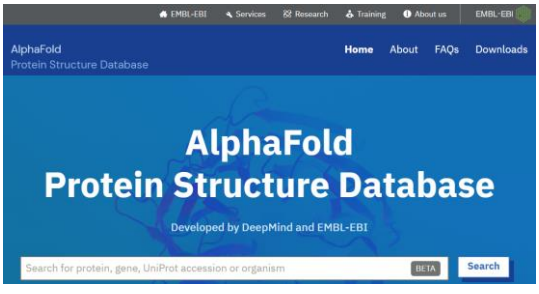
AlphaFold



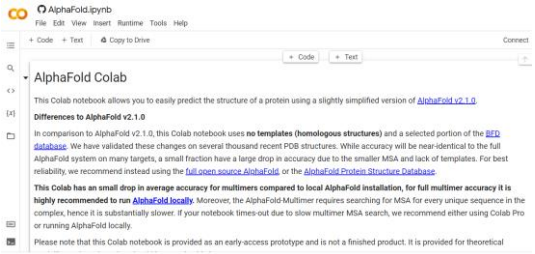
AlphaFold



AlphaFold



AlphaFold


AlphaFold.ipynb
File Edit View Insert Runtime Tools Help
+ Code + Text Copy to Drive
Connect
+ Code + Text
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 an 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