

# 3D Structure Prediction and Disorder Analysis Toolkit

Michał Rembalski, Jagoda Trzeciak, Bruno  
Puczko-Szymański, Marta Korpacz

# Goal

---

Tool with UI to obtain 3d structure based on protein sequence and...

- Multiple methods of input
- Possibility of model selection
- possibility of disorder visualisation with PYMOL

VREVCSEAETGPCR  
AMISRWFYFDVTEGKC  
APFFYGGCGGNRRNF  
DTEEYCMVCGSA



# Input

---

- Uniprot ID
- PDB ID
- FASTA formatted sequence
- Paths to PDB files to run only disorder analysis

## Preprocessing:

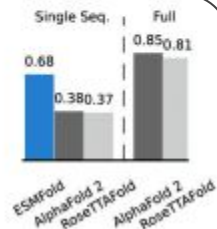
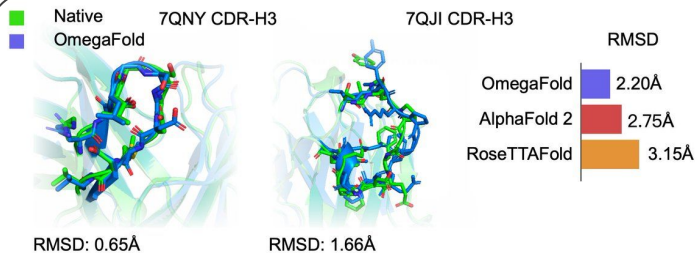
- checking correctness of ids/FASTA
- not accepting non-standard amino acids

# Models

Select Models to Run:

- ☒ SwissModel
- ☒ ESMfold
- ☐ OmegaFold

Run Models



Current CAMEO  
winner

# ESMfold - run via API

---

pros:

- fast (seconds to get structure)
- high accuracy
- does not depend on user computer resources

cons:

- no sequences longer than 400 aa

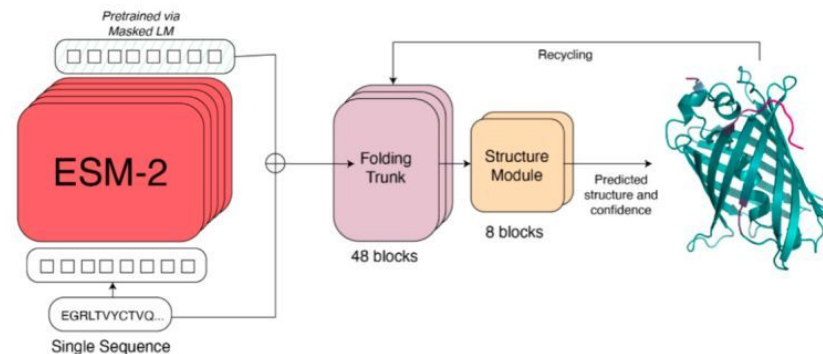


Image source: ESMFold paper

"Evolutionary-scale prediction of atomic level protein structure with a language model,  
<https://www.biorxiv.org/content/10.1101/2022.07.20.500902v3>"

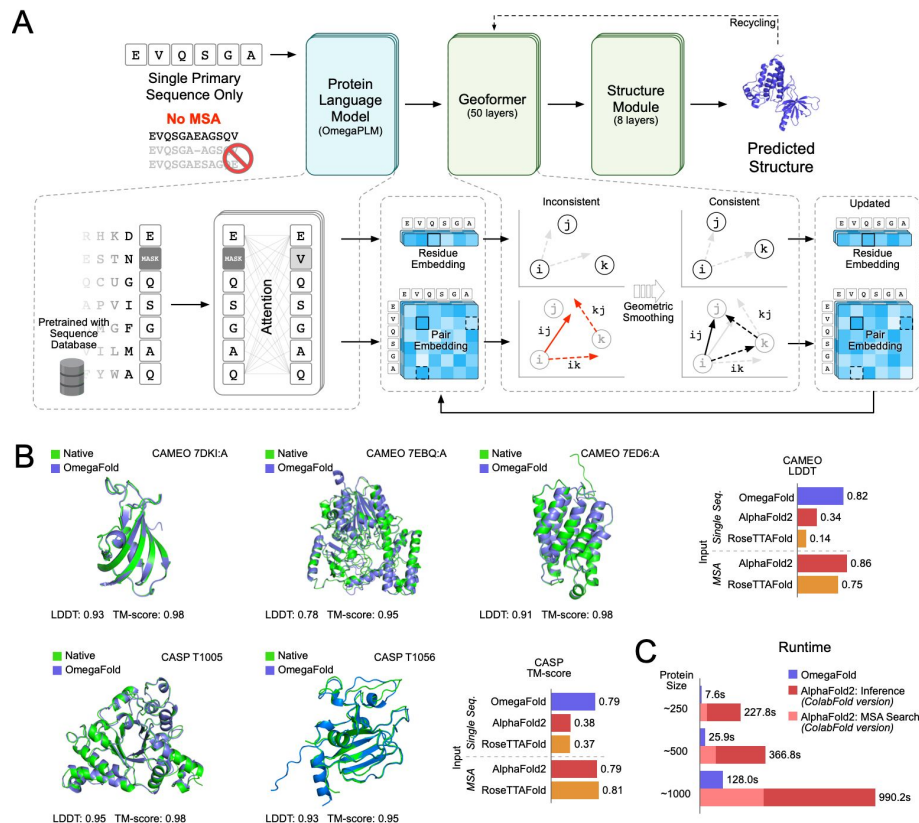
# OmegaFold - run on the CPU

pros:

- no API-related restrictions
- easy to install
- works even on average CPU

cons:

- without GPU long time in comparison to API requests
- depend on user computer resources



# SwissModel - run via API

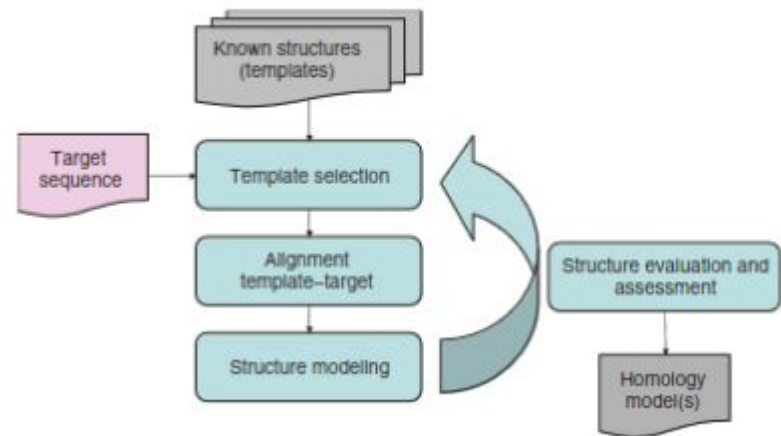
---

pros:

- fast (1-2 minutes to get structure)
- high accuracy
- does not depend on user computer resources

cons:

- requires token for API
- short sequences (<30aa) not allowed

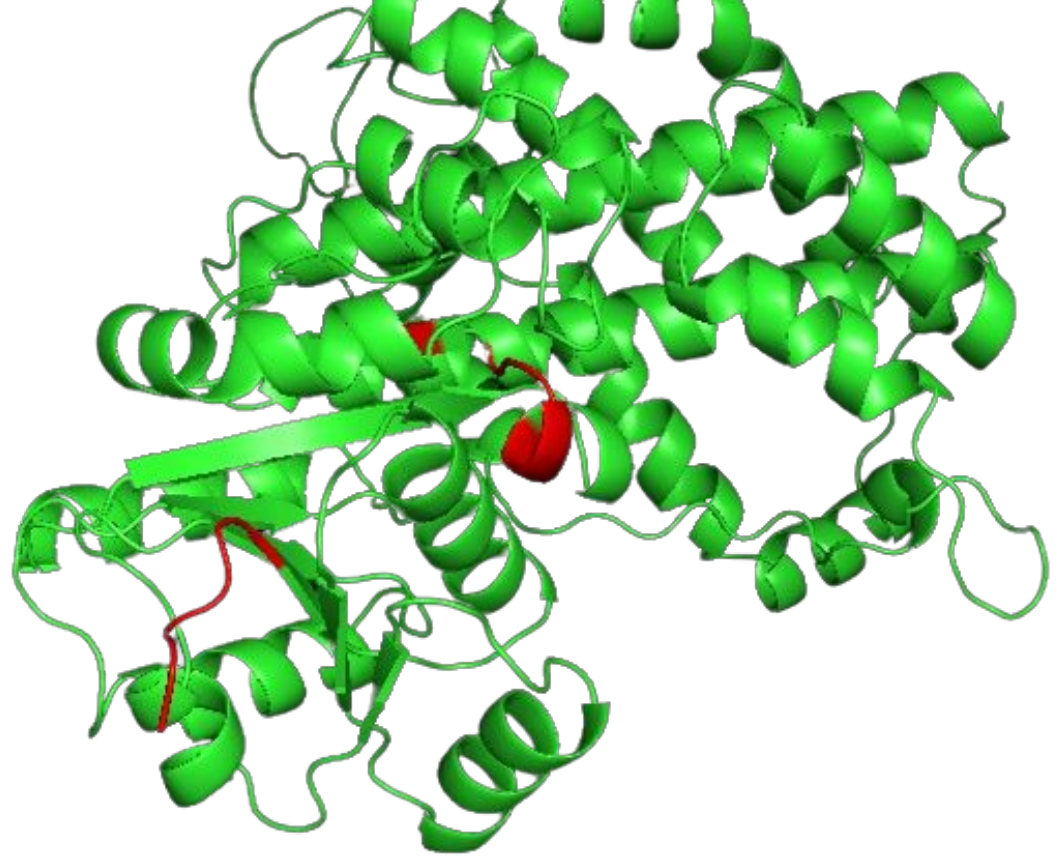


# IUPRED2A runner

---

High accuracy disorder  
prediction

Visualisation with PyMOL





# Bibliography

---

Wu, R., Ding, F., Wang, R., Shen, R., Zhang, X., Luo, S., ... Peng, J. (2022). High-resolution de novo structure prediction from primary sequence. doi:10.1101/2022.07.21.500999

Lin, Z., Akin, H., Rao, R., Hie, B., Zhu, Z., Lu, W., ... Rives, A. (2023). Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science (New York, N.Y.)*, 379(6637), 1123–1130. doi:10.1126/science.ade2574

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., ... Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research*, 46(W1), W296–W303. doi:10.1093/nar/gky427

Mészáros, B., Erdős, G., & Dosztányi, Z. (2018). IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Research*, 46(W1), W329–W337. doi:10.1093/nar/gky384