A Deep Dive into Deep Learning Frameworks for Protein Structure Prediction: Developing and Evaluating Classes of Biomolecular Complexes

Verónica G. Melesse Vergara, Érica Texeira Prates, Manesh Sha, Dan Jacobson

Introduction

- Accurately predicting the structure of a protein has been a long standing and extremely challenging problem in biology.
- In recent years, the rapid evolution and adoption of artificial intelligence (AI) in scientific domains have made the prediction of protein structures leveraging deep learning (DL) frameworks with accuracy rivaling that of experimental crystal structures possible.
- These advances are key to understanding protein function and play a central role in accelerating the drug discovery process.
- This work focuses on comparing the performance of state-of-the-art protein structure prediction models across a predefined set of 7 challenging biomolecule categories.
- *This is a preliminary draft of the poster.*

Background & Motivation

- Accurately predicting the structure of a protein has been a long standing and extremely challenging problem in biology.
- In recent years, the rapid evolution and adoption of artificial intelligence (AI) in scientific domains have made the prediction of protein structures leveraging deep learning (DL) frameworks with accuracy rivaling that of experimental crystal structures possible.

Deep Learning Frameworks

- AlphaFold Family:
  - AlphaFold2 (2021)
  - AlphaFold-multimer (2021)
  - AlphaFold3 (2024)
  - AF2Complex (2022)
- Chai-1 (2024)
- Boltz-1 (2024)

Experimental Methods

- The evaluation was conducted leveraging computational resources at Oak Ridge National Laboratory (ORNL) including Frontier, ORNL's exascale supercomputer.
- In addition, two cloud servers were used in order to study the latest DL framework released by DeepMind and Chai Discovery
  - Experiments conducted explore different features available on each framework:
    - Using MSA or not
    - Including PTMs or not
    - Among others

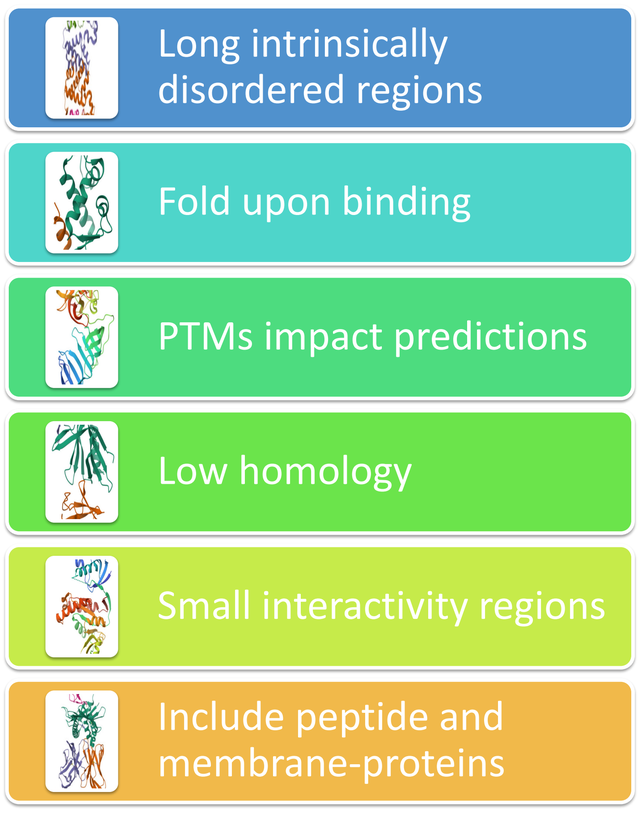
Cloud servers used for Chai-1

Cloud servers used for AF3

Used for AF2, AF-multimer, AF2Complex, Boltz-1

Classification of Biomolecular Complexes

- The study considers 7 categories of biomolecules of interest:



- For each case, 10 protein complexes were selected to be used for evaluation.
- Complexes selected were non-overlapping with the training data sets for each framework.

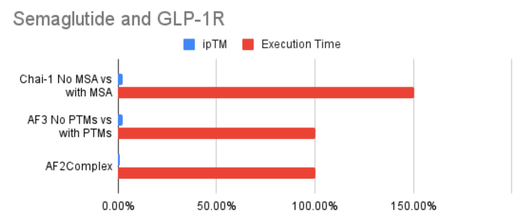
Evaluation & Results

- 10 complexes were collected for each of the classes defined.
- Predictions were conducted with state-of-the-art DL frameworks.
- Accuracy and computational performance across frameworks are compared.
- The derived classification and guidelines are then applied to specific use cases of interest.

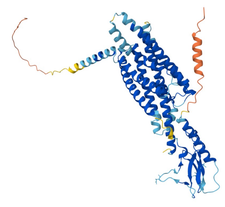
Framework	ipTM	pTM	Execution Time	ipTM	pTM	Execution Time
Chai-1	0.7152	TBD	139	TBD	TBD	TBD
AF2Complex	TBD	TBD	TBD	TBD	TBD	TBD
AF3	0.731	TBD	348	TBD	TBD	TBD
AF-multimer	TBD	TBD	TBD	TBD	TBD	TBD
AF2	TBD	TBD	TBD	TBD	TBD	TBD
Boltz-1	TBD	TBD	TBD	TBD	TBD	TBD

Investigating Off-target effects for GLP-1R

- Use case of interest predicting interactions between semaglutide, a known GLP-1R agonist, with GLP-1R show that post translational modifications (PTMs) play a key role in the resulting predictions.
- The derived classification obtained from the study as well as the guidelines are used to explore off-target effects for GLP-1R are explored.



- An average interface predicted template modeling (ipTM) score of 0.92 is obtained using AlphaFold3 which is ~2.3% higher than without PTMs (i.e., without glycosylation) included.
- Similar experiments executed with Chai-1 result in ipTM scores ~22.4% lower and pTM scores ~2.3% higher than those from AlphaFold3.



Semaglutide and GLP-1R with PTMs as predicted by AF3

- *These are preliminary results for the draft poster.*

Conclusions

- The results compare the accuracy and performance of each method when applied to classes of challenging protein-protein complexes.
- The evaluation was conducted leveraging computational resources at Oak Ridge National Laboratory (ORNL) including Frontier, ORNL's exascale supercomputer.
- The data set constructed, the resulting biomolecular complex type classification, and the comprehensive set of guidelines derived can aid in future experiment design.
- The results from this study can also provide a deeper understanding of the advantages and limitations of each model when applied to specific classes of biomolecular complexes as opposed to individual complexes.

Acknowledgements

- This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725.