MIBO 8110: AlphaFold 3 for Predicting Biomolecular Structures and Interactions

Xiping Gong

Nov 11/13, 2025

Two-Day Workshop

Motivation: This hands-on workshop introduces AlphaFold models and the applications in predicting biomolecular structures and interactions.

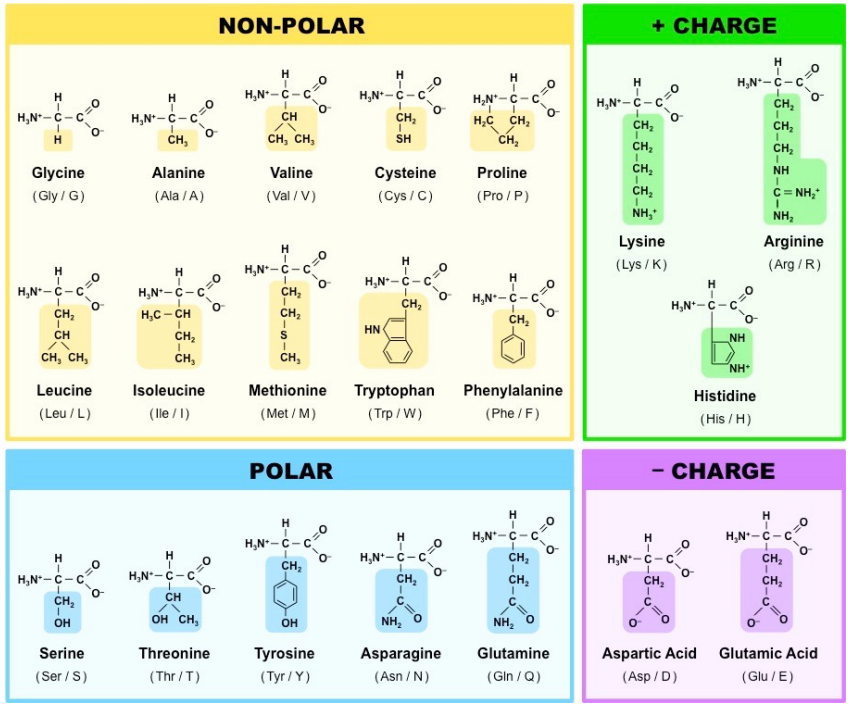
Day 1 (Nov. 11): From 1-D Sequence to 3-D Structure

- AlphaFold Models
- Practical Session — Protein Structure Prediction

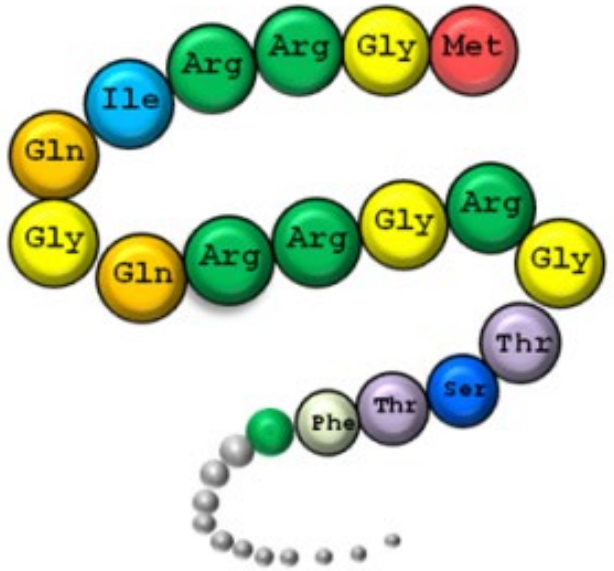
Day 2 (Nov. 13): Protein–Ligand Interactions

- Basics of Molecular Docking and AlphaFold 3 (AF3) Capabilities
- Practical Session — Predicting Protein–Ligand Complexes

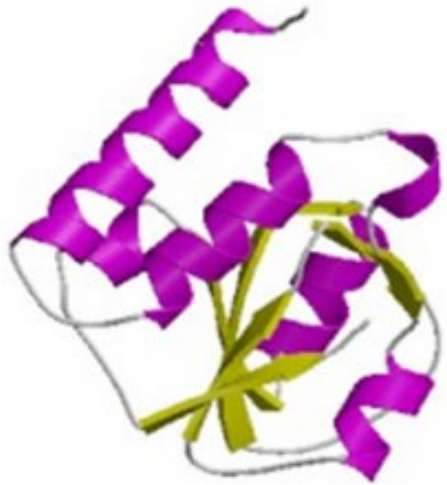
Fundamentals: Protein Structure Prediction



#20 Amino acids



1-D sequence



3-D functional structure

AlphaFold: AI-Based Tools for Protein Structure Prediction

✅ AlphaFold Development Timeline

Year	Version	Key Features & Achievements
2018	AlphaFold 1	Debuted at **CASP13**, first AI system to outperform traditional methods in protein structure prediction.
2020	AlphaFold 2	Presented at **CASP14**, achieved near-experimental accuracy; introduced transformer-based architecture; solved the protein folding problem.
2021	AlphaFold DB	Released **AlphaFold Protein Structure Database** with millions of predicted structures; democratized access for researchers worldwide.
2024	AlphaFold 3	Expanded to **protein–ligand, protein–DNA/RNA, and ion interactions**; improved accuracy for complex biomolecular interactions by 50%+.
2024	Nobel Prize	Awarded to **Demis Hassabis and John Jumper** for AlphaFold's transformative impact on biology and medicine.

https://www.youtube.com/watch?v=cx7l9ZGFZkw

Tutorial: Setup a Computing Environment

Access

Personal computer

Sapelo2@GACRC

- **Service Request**: Service Catalog - GACRC Service Catalog

Note: The Sapelo2 is free to the UGA users, and they provide very useful training resources in computing and software use. **(I assume that you did a GACRC Teaching Cluster Training.)**

- **Installed VMD in your local computer**: VMD - Visual Molecular Dynamics

Tutorial: A Step-by-Step Demo

Prerequisites

- You can log into the teaching node using your MyID account: MyID@teach.gacrc.uga.edu.

- You know how to transfer files between your computer and the teaching node.

- You have VMD installed on your computer.

Tutorial: Human Serum Albumin

Human Serum Albumin (hSA) is the most abundant protein in human blood plasma, playing a critical role in maintaining osmotic pressure and transporting hormones, fatty acids, and drugs throughout the body.

Protein sequence (PDB ID: 7AAI):

HKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLF
GDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKK
YLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQK
FGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSK
LKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYS
VVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVR
YTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCC
TESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAV
MDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAAL

Tutorial: A Step-by-Step Demo

…

Two-Day Workshop

Motivation: This hands-on workshop introduces AlphaFold models and the applications in predicting biomolecular structures and interactions.

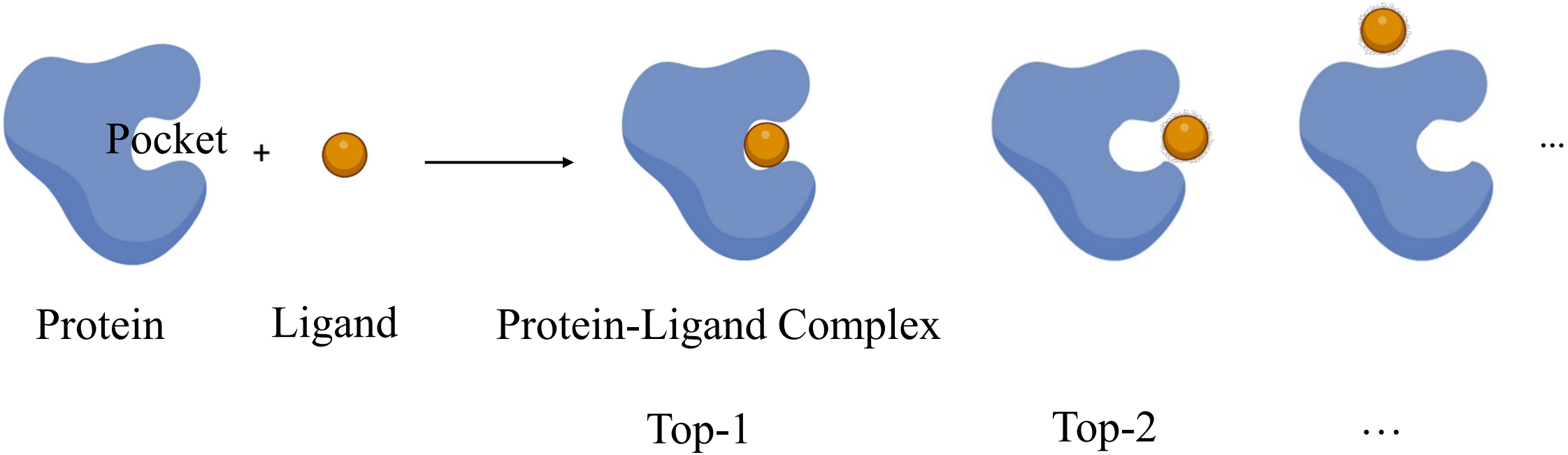
Day 1 (Nov. 11): From 1-D Sequence to 3-D Structure

- AlphaFold Models
- Practical Session — Protein Structure Prediction

Day 2 (Nov. 13): Protein–Ligand Interactions

- Basics of Molecular Docking and AlphaFold 3 (AF3) Capabilities
- Practical Session — Predicting Protein–Ligand Complexes

Molecular Docking



Molecular docking tools: AlphaFold 3, AutoDock Vina, etc.

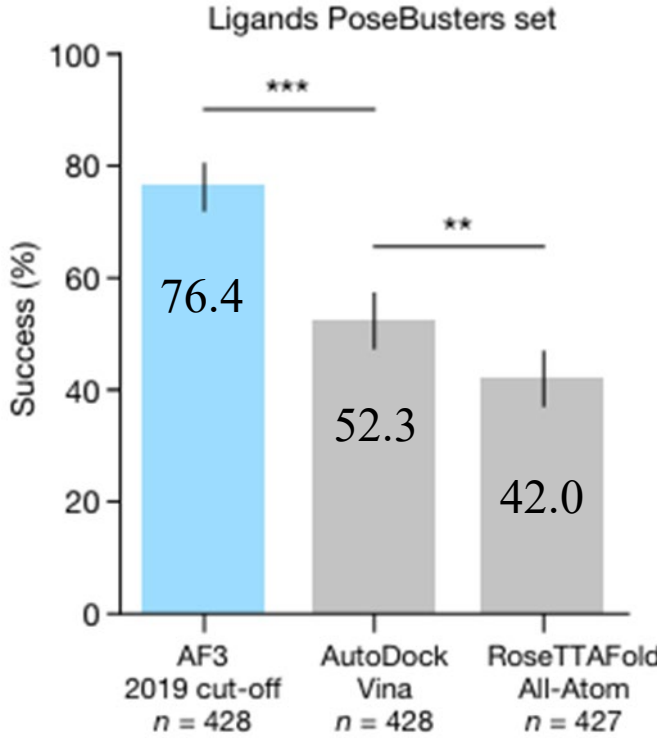
AlphaFold 3 (AF3)

- The AF3 has a capability of predicting the biomolecular interactions, including the protein-ligand docking.

google-deepmind/
alphafold3

AlphaFold 3 inference pipeline.

https://github.com/google-deepmind/alphafold3



AF3 Training Dataset

- Data sources: PDB (proteins and complexes), UniProt (sequences), chemical compound libraries, RNA/DNA databases (e.g., Rfam, RNACentral).

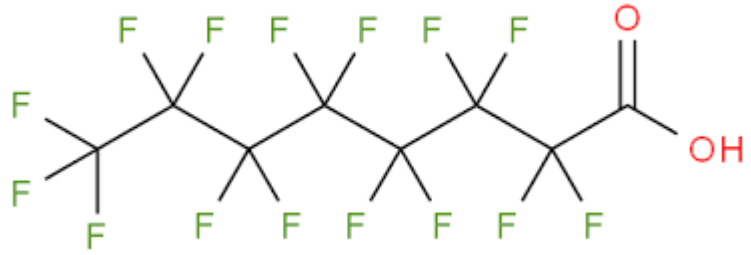
- Training Data Cutoff: Structures released before September 2021.

- Validation: Structures from late 2021–2023 used for testing.

Tutorial: Human Serum Albumin - PFOA

Human Serum Albumin (hSA) is the most abundant protein in human blood plasma, playing a critical role in maintaining osmotic pressure and transporting hormones, fatty acids, and drugs throughout the body.

Perfluorooctanoic acid (PFOA; conjugate base perfluorooctanoate; also known colloquially as C8, from its chemical formula $C_8HF_{15}O_2$) is a perfluorinated carboxylic acid produced and used worldwide as an industrial surfactant in chemical processes and as a chemical precursor.



Protein sequence (PDB ID: 7AAI)
https://onlinelibrary.wiley.com/doi/10.1002/pro.4036

Tutorial: A Step-by-Step Demo

…

Take-Home Exam

1. From the Protein Data Bank (PDB), select another PDB ID that includes a protein structure of interest and at least one ligand. Use AlphaFold 3 (AF3) to predict both the protein structure and the protein–ligand interactions. Then, answer the following questions based on your selected example:

Protein Structure Prediction: How accurate is the predicted protein structure compared to the experimental structure? Please calculate the backbone RMSD values.

Protein–Ligand Interaction Prediction: How accurate are the predicted protein–ligand interactions, please calculate the RMSD values? Which protein residues interact with the ligand, please use VMD to visualize the residues within 0.3 nm of ligand? What types of interactions (e.g., hydrogen bonding, hydrophobic, or electrostatic) can you observe?

In the End

If you have any questions, please contact Dr. Xiping Gong (xipinggong@uga.edu).