MASSIVEFOLD

Practical work

Protein complex

IFB training - AlphaFold and beyond - 10th to 12th December 2025

H1140

The H1140 CASP target is a nanobody/antigen target.
These targets are know to be hard to predict, certainly because of a lack of evolutionary signals.

We are going to model them with the 3 tools and all the neural network versions of AFmassive and ColabFold.

- Load MassiveFold 1.5

```
module load massivefold/1.5.0
```

- Edit the ‘AFmassive_params.json’ and ‘ColabFold_params.json’ files, change the “model_preset” to “multimer” and add “recycles” to the plots (you may want to copy these files to another name which contains ”_multimer” for instance, to keep both file types, for monomers and multimers)

- The H1140.fasta is already present in the input folder because this is the example provided

- Run AFmassive and ColabFold:

```
./run_massivefold.sh -s input/H1140.fasta -r afm_default -p 5 -f AFmassive_params.json
./run_massivefold.sh -s input/H1140.fasta -r cf_default -p 5 -f ColabFold_params.json
```

H1140

For AlphaFold3, you need to specify that the fasta file contains 2 proteins

Edit the ‘AlphaFold3_params.json’ file by setting:

- “fasta_chains” to 2 protein chains
- “model_preset” to “multimer”

```
"AF3_run": {
        "fasta_chains": [“protein”,”protein”],
        "ligand": [
            {"ccdCodes": [""], "smiles": ""}
        ],
        "modifications": [
            [{"type": "", "sequence": "", "positions": []}]
        ],
        "model_preset": "multimer",
        "max_template_date": "2024-11-28",
        "num_diffusion_samples": "5",
        "unpairedMsa": "true",
        "pairedMsa": "true",
        "templates": "true"
    },
```

- Run AlphaFold3

```
./run_massivefold.sh -s input/H1140.fasta -r af3_default -p 5 -f AlphaFold3_params.json
```

H1140 - Results

- Get the results on your local computer to visualize them
- Alternatively, you can get or visualize them here

  https://nextcloud.univ-lille.fr/index.php/s/dnQLmBywbcQdx56

H1140 - Results

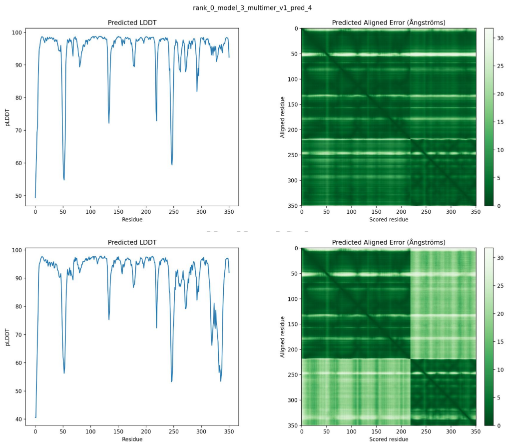
Compare the results for the 3 tools, looking at the ‘ranking_iptm.json’ and the plots.

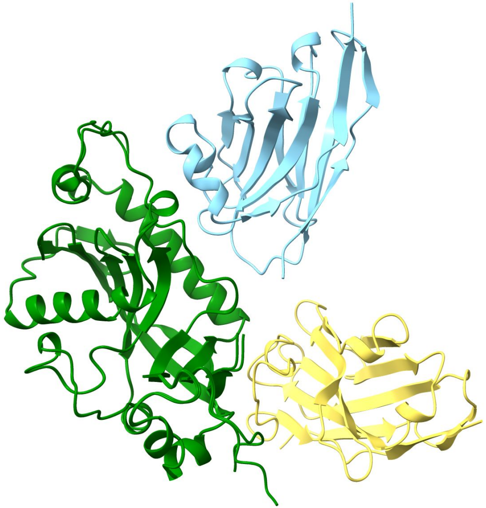
=> Which tool seems to give the best predictions between AFmassive and ColabFold ?

=> Which neural network version provides this prediction ?

=> What about AlphaFold3 predictions ?

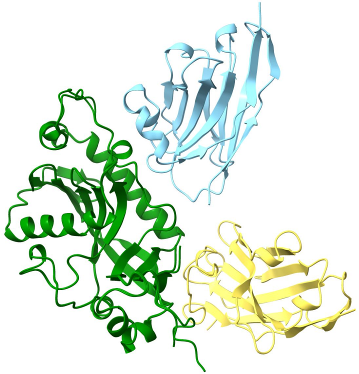
=> Open a high confident prediction in ChimeraX and align it to the released PDB 9ERT to visualize the quality of the prediction. Align the next prediction of AFmassive or ColabFold with a less good PAE matrix. What do you see ?

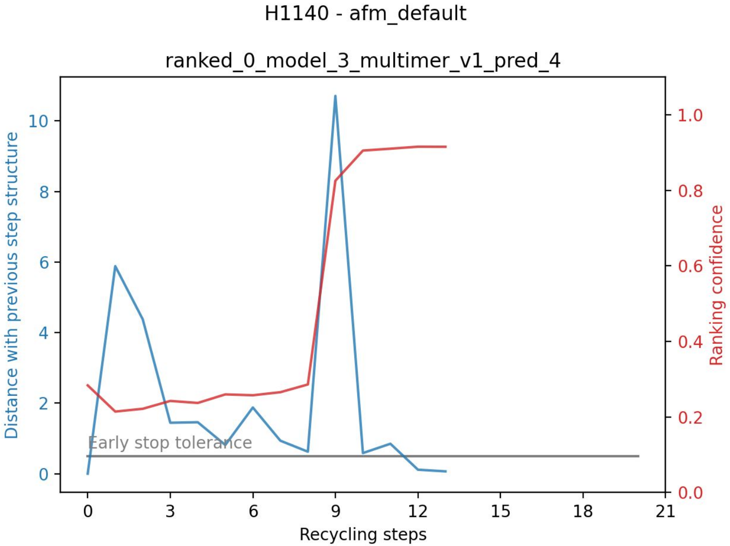


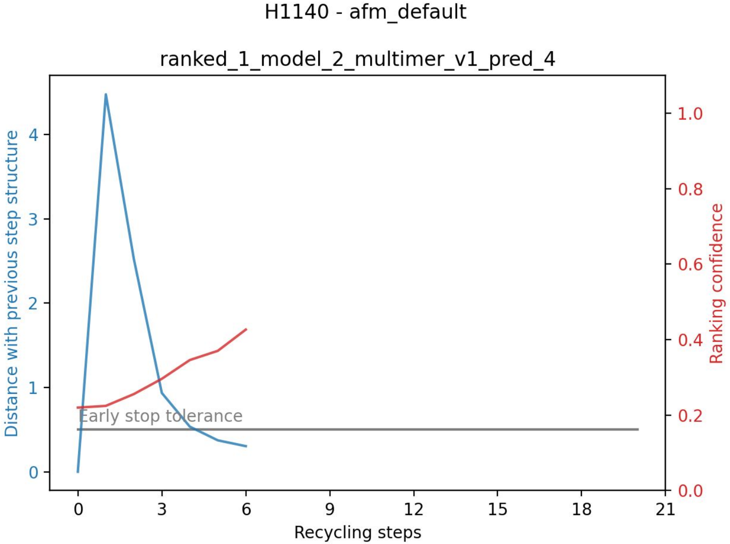


H1140 - Results

Look at the first recycles plots of AFmassive and ColabFold. What do you see ?







T231

T231 is an antibody-antigen target, made of the heavy and light chains bound to a peptide.

- Get the T231.fasta sequence and perform the same runs with MassiveFold as for H1140

```
./run_massivefold.sh -s input/T231.fasta -r afm_default -p 5 -f AFmassive_params.json
./run_massivefold.sh -s input/T231.fasta -r cf_default -p 5 -f ColabFold_params.json
./run_massivefold.sh -s input/T231.fasta -r af3_default -p 5 -f AlphaFold3_params.json
```

N.B.: for AlphaFold3, add 3 times “protein” in the “fasta_chains”

T231 - Results

- Get the results on your local computer to visualize them

- Alternatively, you can get them here

  https://nextcloud.univ-lille.fr/index.php/s/dnQLmBywbcQdx56

T231 - Results

- Compare the results of the 3 tools, looking at the ‘ranking_iptm.json’ and the plots.

=> What can you say about the results for AFmassive and ColabFold ?

=> Why do the ipTM scores are so homogeneous ?