Alphafold on SCC - An Introduction

Alexander Wähling

January 16, 2023

AlphaFold

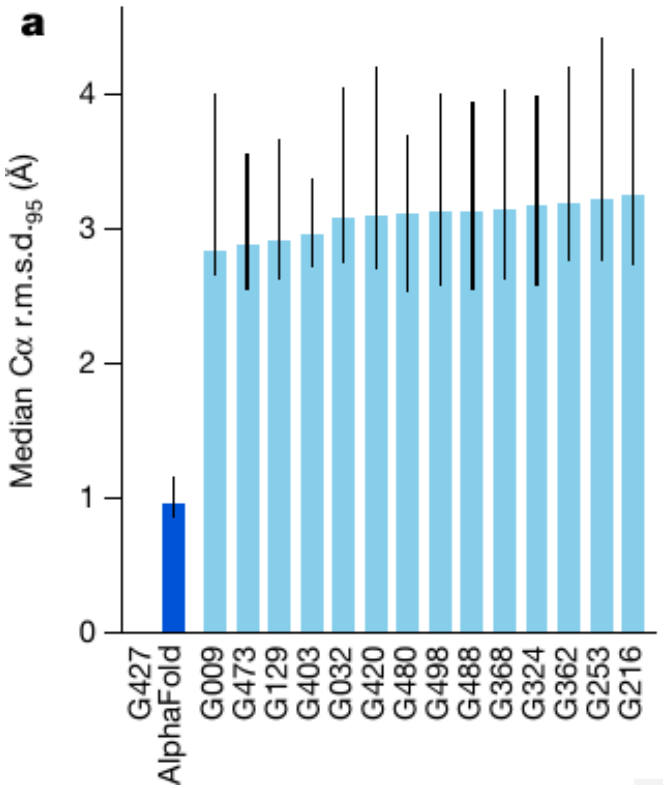

Figure: AlphaFold performance compared to competing teams at CASP14

AlphaFold

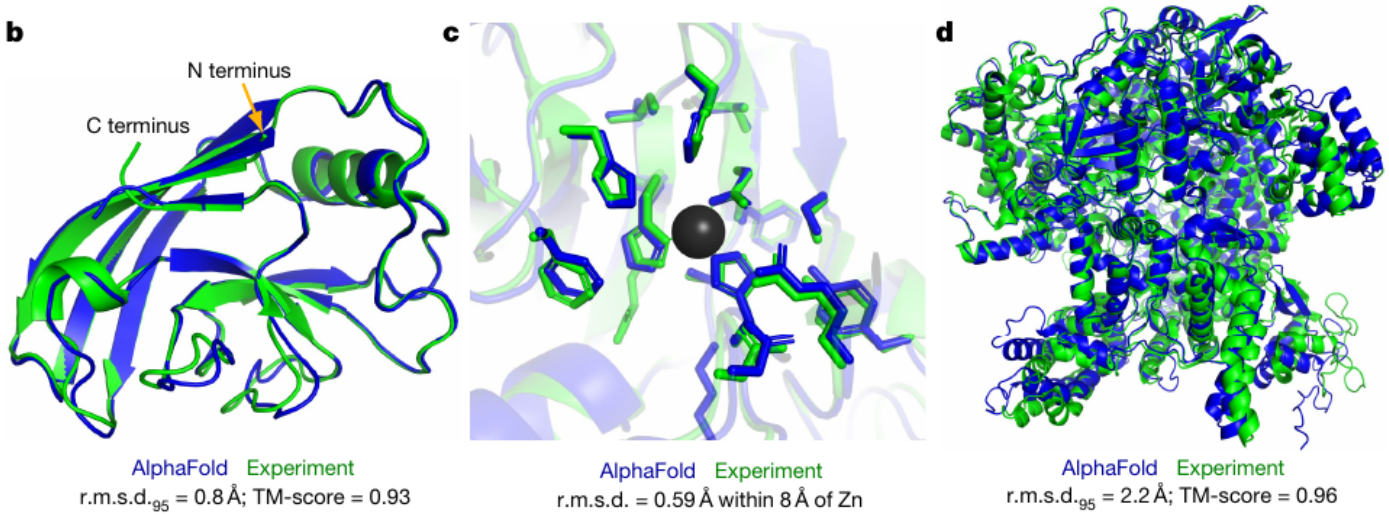

Figure: experimental proteinstructure (green) compared to AlphaFold-predicted structure (blue)

AlphaFold

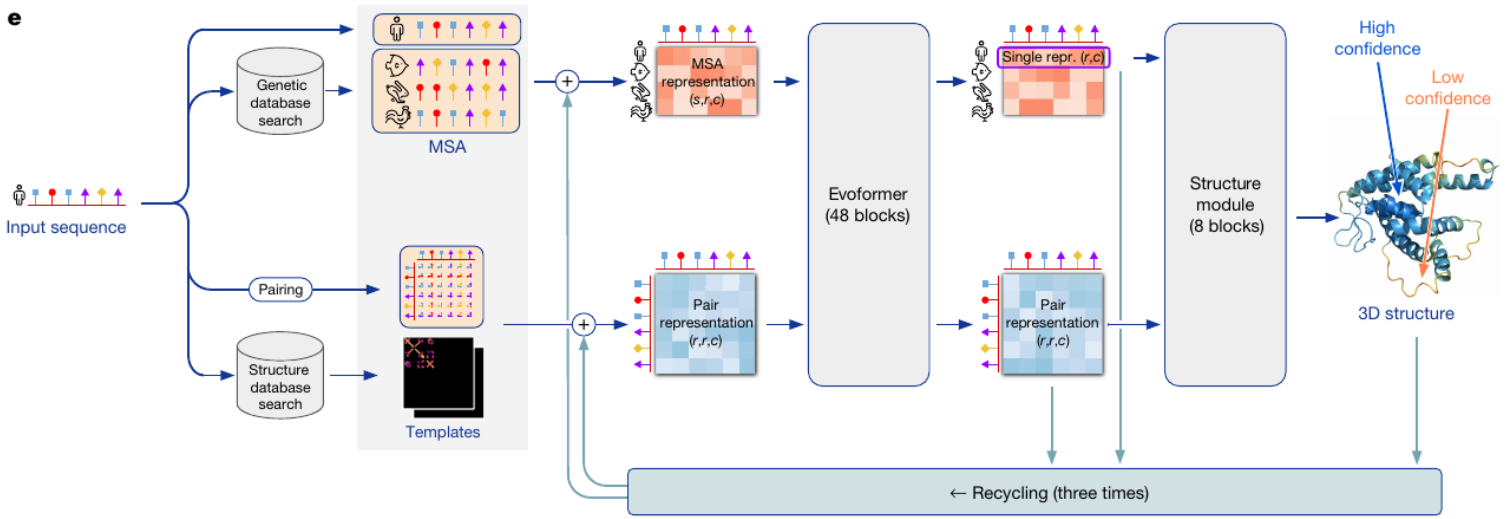

Figure: AlphaFold architecture

Notes

- needed memory not directly linked to sequence length
- MSA size more important
- AlphFold does NOT parallelize the model but shares memory
- limitations at around 3000 AA size
- multiple GPU usage for shared memory can boost performance
- increase number of CPUs with the number of GPUs

Relevant Paths

- Projectdir: /scratch/projects/alphafold
- run script:
  /scratch/projects/alphafold/scripts/run_alphafold.sh
- example batch file:
  /scratch/projects/alphafold/scripts/batch_alphafold_example.sh