AlphaFold

AlphaFold is an AI program developed by DeepMind that predicts protein structures from amino acid sequences.

Databases

The necessary databases are pre-mounted on GPU nodes at `/alphafold_storage/alphafold_db` — no download needed.

Required Parameters

Parameter	Description
`-d <data_dir>`	Path to the supporting data directory
`-o <output_dir>`	Path to store results
`-f <fasta_paths>`	Path to FASTA file(s). Multiple sequences in one file = multimer. Multiple files comma-separated = fold sequentially
`-t <max_template_date>`	Maximum template release date (YYYY-MM-DD)

Optional Parameters

Parameter	Default	Description
`-g`	true	Enable NVIDIA GPU runtime
`-r`	true	Run final relaxation step
`-e`	true	Run relax on GPU
`-n`	all cores	OpenMM threads
`-a`	0	CUDA_VISIBLE_DEVICES — comma-separated GPU list

Parameter	Default	Description
`-m`	monomer	Model preset: `monomer`, `monomer_casp14`, `monomer_ptm`, `multimer`
`-c`	full_dbs	MSA database preset: `reduced_dbs` or `full_dbs`
`-p`	false	Use precomputed MSAs from disk
`-l`	5	Predictions per model (multimer only)
`-b`	false	Benchmark mode — excludes compilation time

Example Job Script

```
#!/bin/bash
#SBATCH --job-name=AlphaFold-Multimer
#SBATCH --partition=gpu2
#SBATCH --nodes=1
#SBATCH --ntasks=1
#SBATCH --cpus-per-task=4
#SBATCH --mem=32G
#SBATCH --gres=gpu:1
#SBATCH --output=alphafold_%j.out
#SBATCH --error=alphafold_%j.err

module load alphafold/alphafold_non_docker_2.3.1

bash $ALPHAFOLD_SCRIPT_PATH/run_alphafold.sh \
  -d $ALPHAFOLD_DB_PATH \
  -o ~/output_dir \
  -f $ALPHAFOLD_SCRIPT_PATH/examples/query.fasta \
  -t $(date +%Y-%m-%d)
```

Memory Guidelines

- **Monomer** — at least 32GB RAM
- **Multimer** — at least 64GB RAM; large/complex structures may need 128GB+

Additional Resources

- Sample FASTA: `/home/alphafold_folder/alphafold_multimer_non_docker/example/query.fasta`
- alphafold_non_docker GitHub repository

Created 2026-06-14 08:23:58 UTC by levk
Updated 2026-06-14 08:24:33 UTC by levk