Making use of averaging methods in MODELLER for protein structure prediction

Serena Rosignoli[1], Elisa Lustrino[1], Iris Di Silverio[1] and Alessandro Paiardini[1, *]

[1] Department of Biochemical Sciences, Sapienza University of Rome, 00185, Rome
* Correspondence: alessandro.paiardini@uniroma1.it

Command-line Arguments

The script can be executed from the command line with the following options:

1. `--alnfile`: Path to the alignment file (e.g., 'tar_tem_alignment.ali').
2. `--knowns`: target structure name (e.g., '1bdm').
3. `--sequence`: label (e.g., 'TvLDH').

Optional Arguments

1. `--n_jobs` (default: 4): Number of parallel jobs to use during model averaging.
2. `--n_runs` (default: 24): Number of MD runs to perform.
3. `--model_name` (default: 'TvLDH_averaged'): Name of the output model file.
4. `--scoring_func` (default: 'dope'): Scoring function for model evaluation.
5. `--select_top` (default: 0.1): Fraction of top decoys to be used for averaging.

Example Usage

```
python altmod_run_averaging.py --alnfile tar_tem_alignment.ali --
knowns 1bdm --sequence TvLDH
```

Notes

- The script initializes an 'automodel' object using the Automodel_averaging class.
- Additional parameters, such as md_level, can be adjusted directly within the script if needed.
- The final averaged model will be written to the specified output file.
- If you want to build multiple averaged models for the same protein, remember to change the random seed of MODELLER within the script, by adding the lines.

  ```
  random_seed = 42
  env.io.random_seed = random_seed
  ```