Phenix *Python-based Hierarchical ENvironment for Integrated Xtallography*

Tutorial: Fit Biomolecules into Cryo-EM Maps using MD Simulation (GUI)

Contents

Overview

This tutorial will show you how to fit biomolecule atomic structures into cryo-EM maps using molecular dynamics simulation within the PHENIX graphical user interface (GUI).

For commandline execution, please see the cryo_fit_commandline_tutorial

Theoretical explanation of cryo_fit is here

For installation of cryo_fit, please see the installation notes for cryo_fit

Input files

<initial_model> and <target_map>

Initial Model

Available format: .cif and .pdb

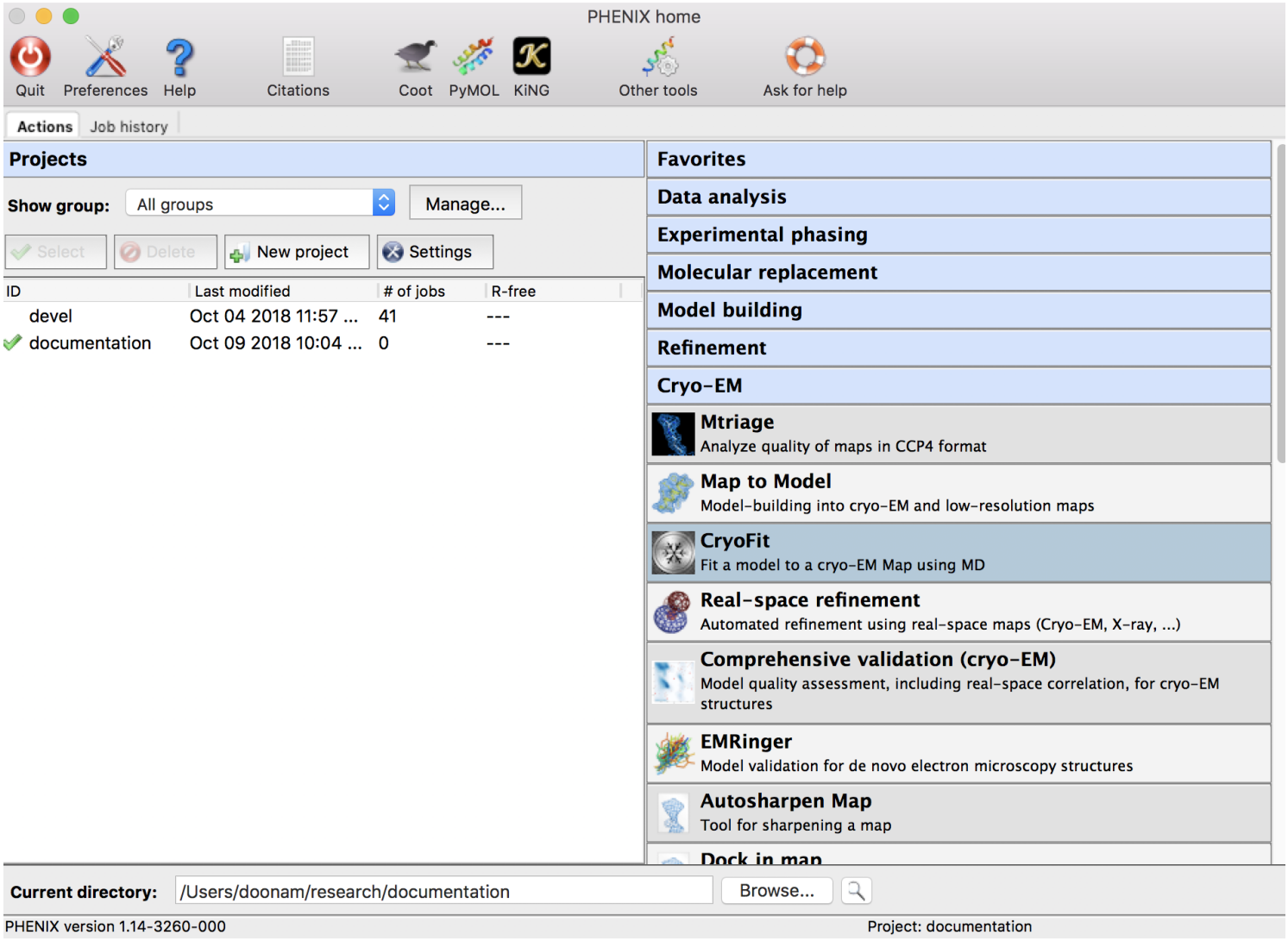
The initial model is a guide or template structure (CIF/mmCIF/pdb) that is close to a target cryo EM map structurally.

You can use either map_to_model or UCSF chimera (Tools -> Volume Data -> Fit in Map)) to prepare the initial model.

Target Map

Available format: .ccp4 and .map (MRC style in binary file) and .sit (Situs style in text file)

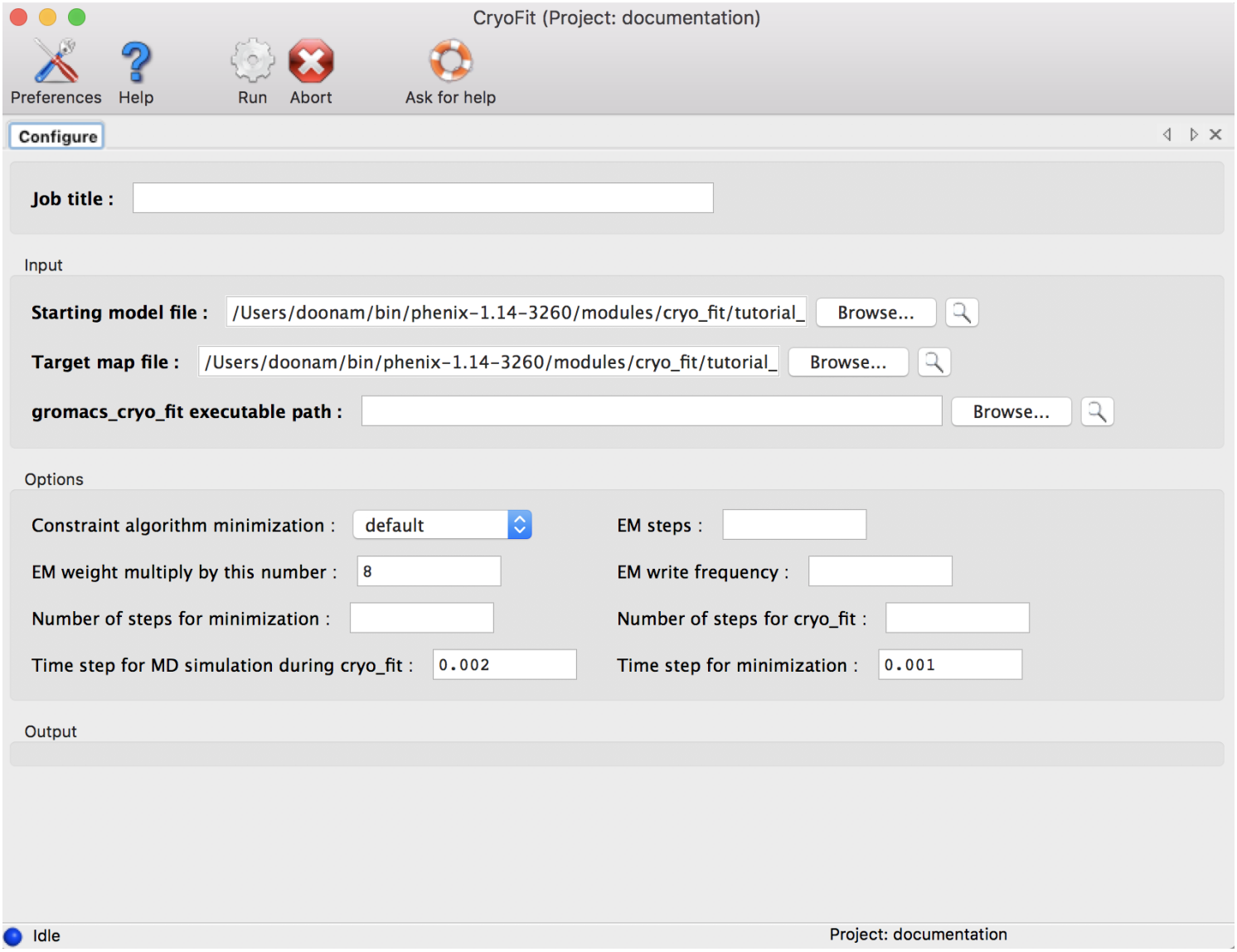
Launch cryo_fit



Enter Input Files

Click browse buttons.

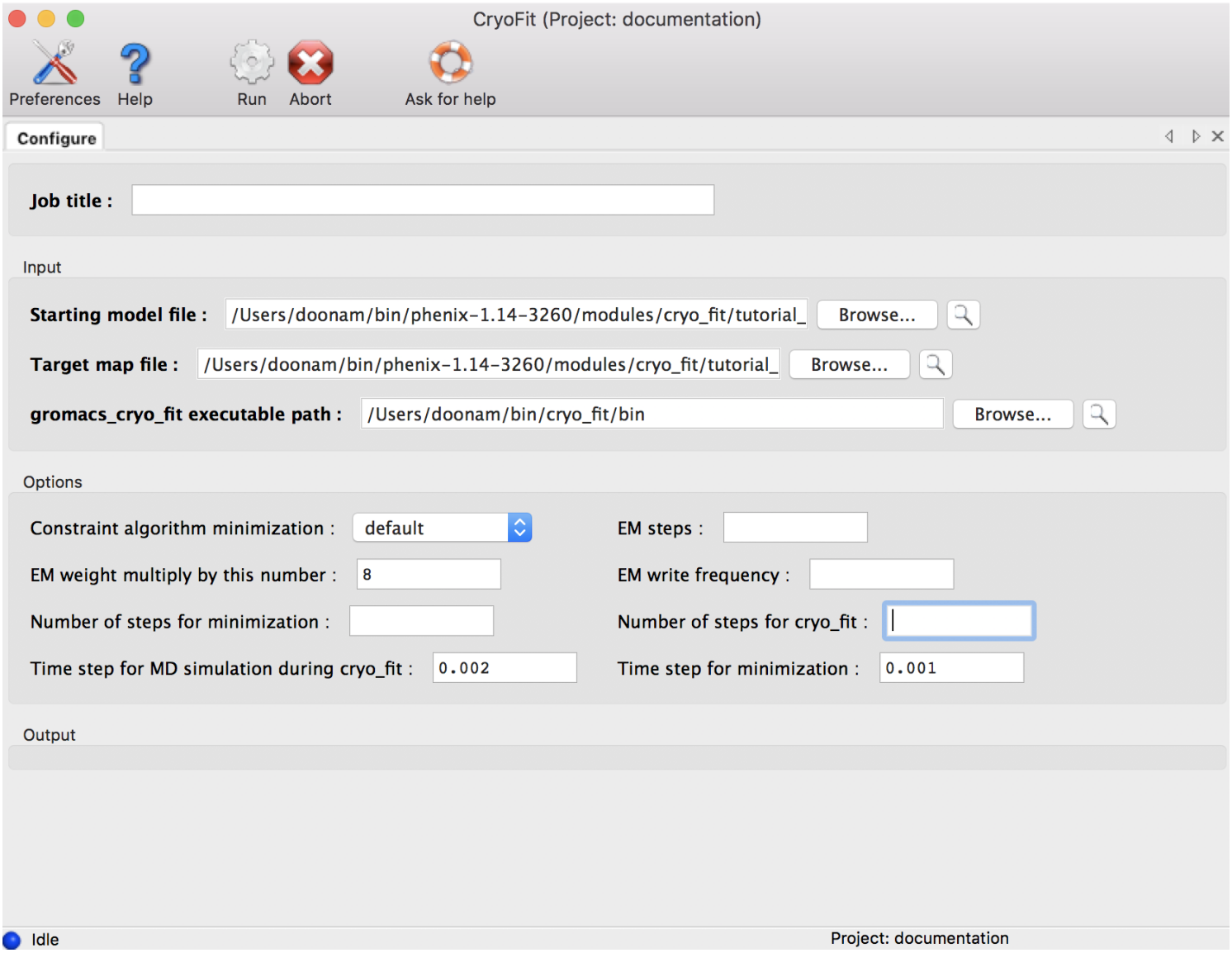
(Tutorial input files live in <User_phenix>/modules/cryo_fit/tutorial_input_files)



Specify cryo_fit executable location

Click browse buttons.

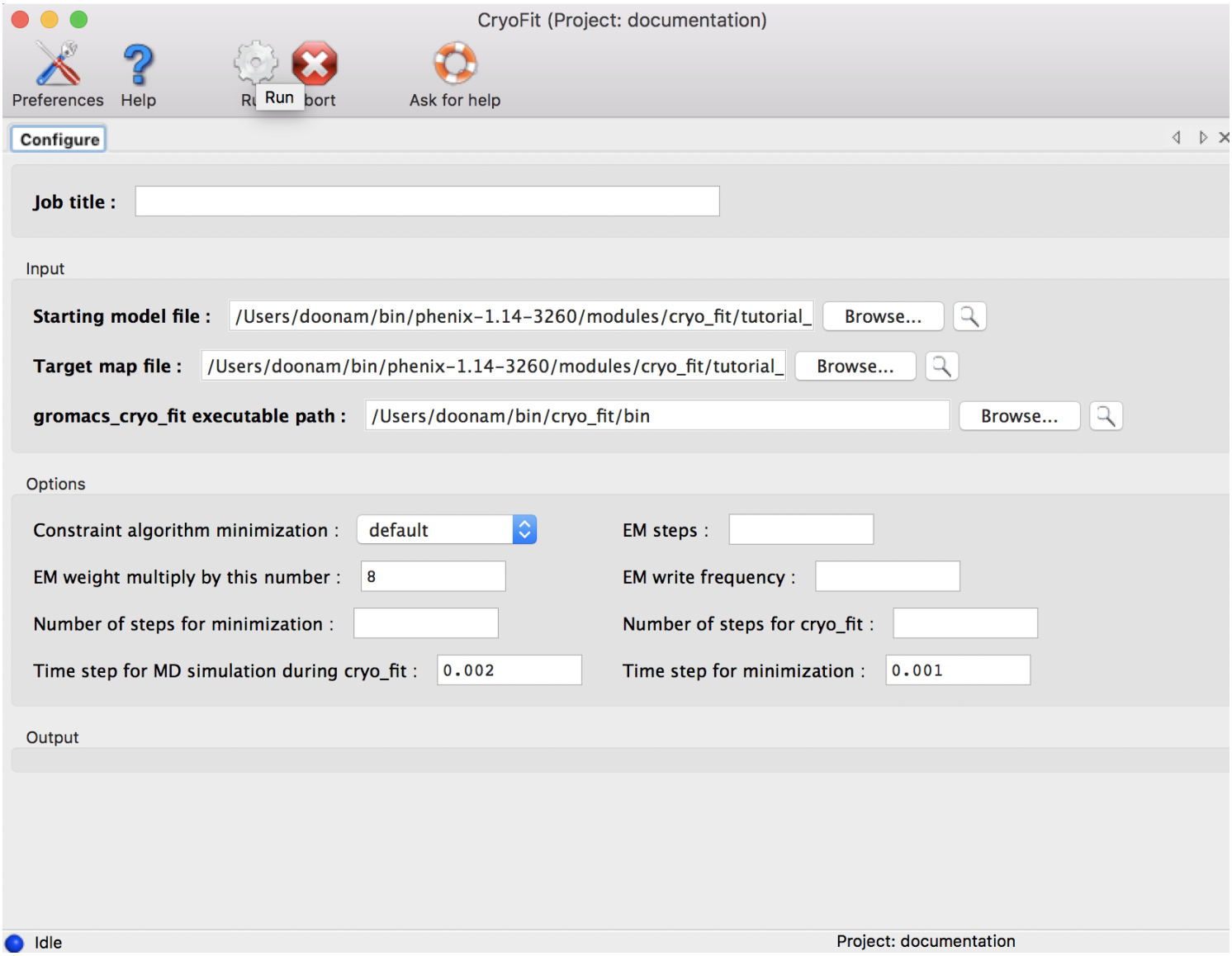
(select for exampple, /Users/doonam/bin/cryo_fit/bin)



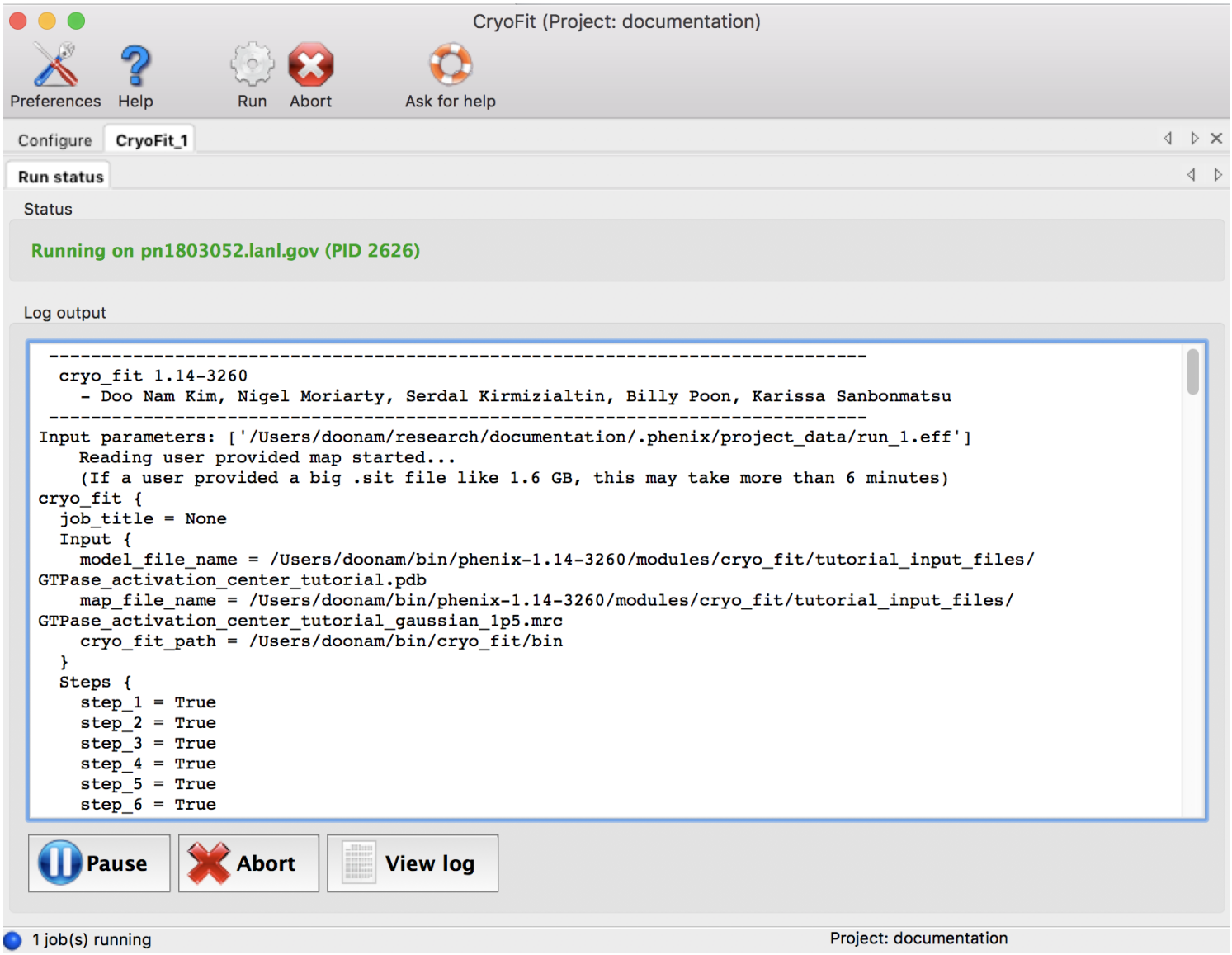
Enter Options

All options can be left blank (cryo_fit will figure out all options automatically).

Run



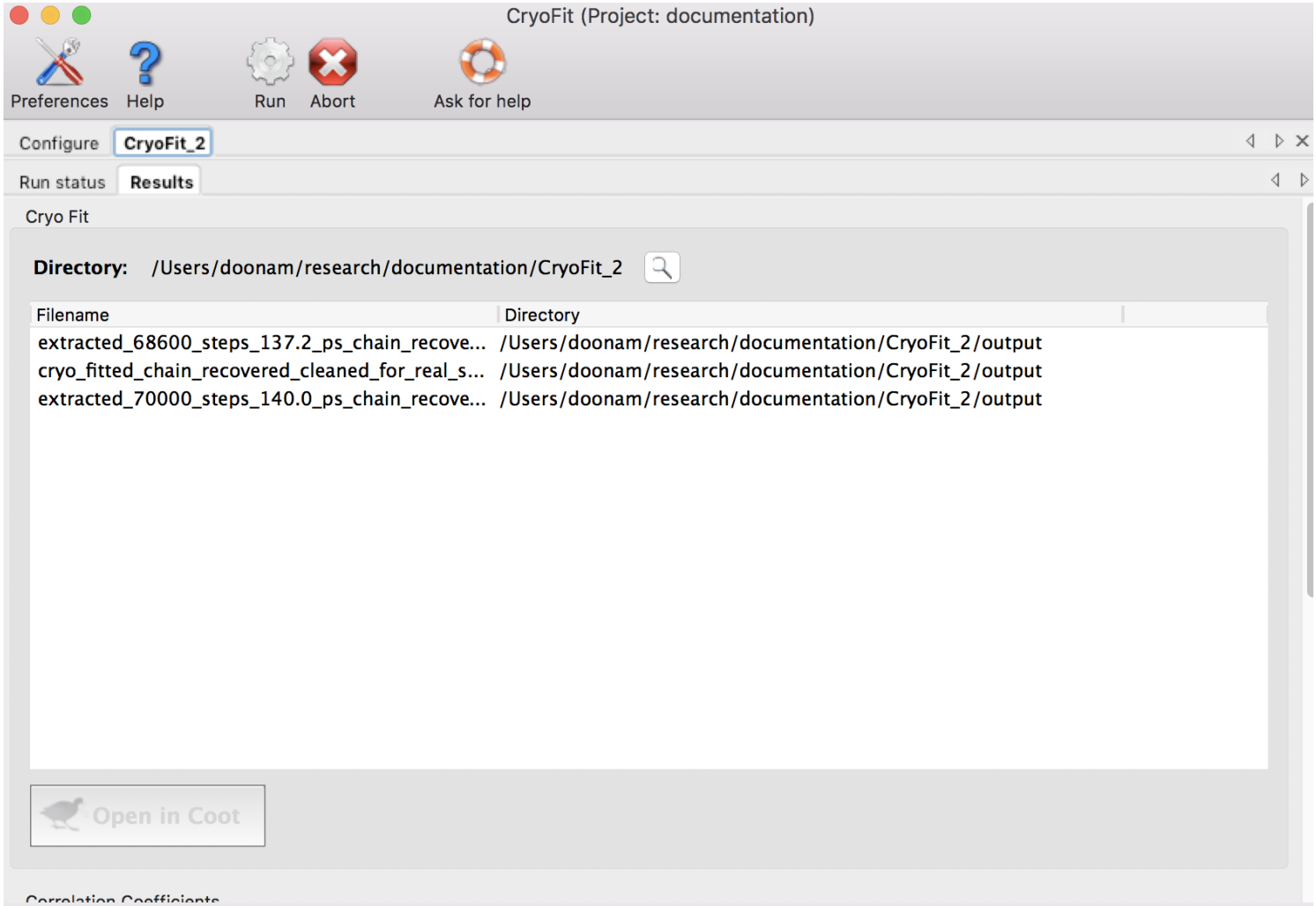
It should start something like this (total steps are 1~8).

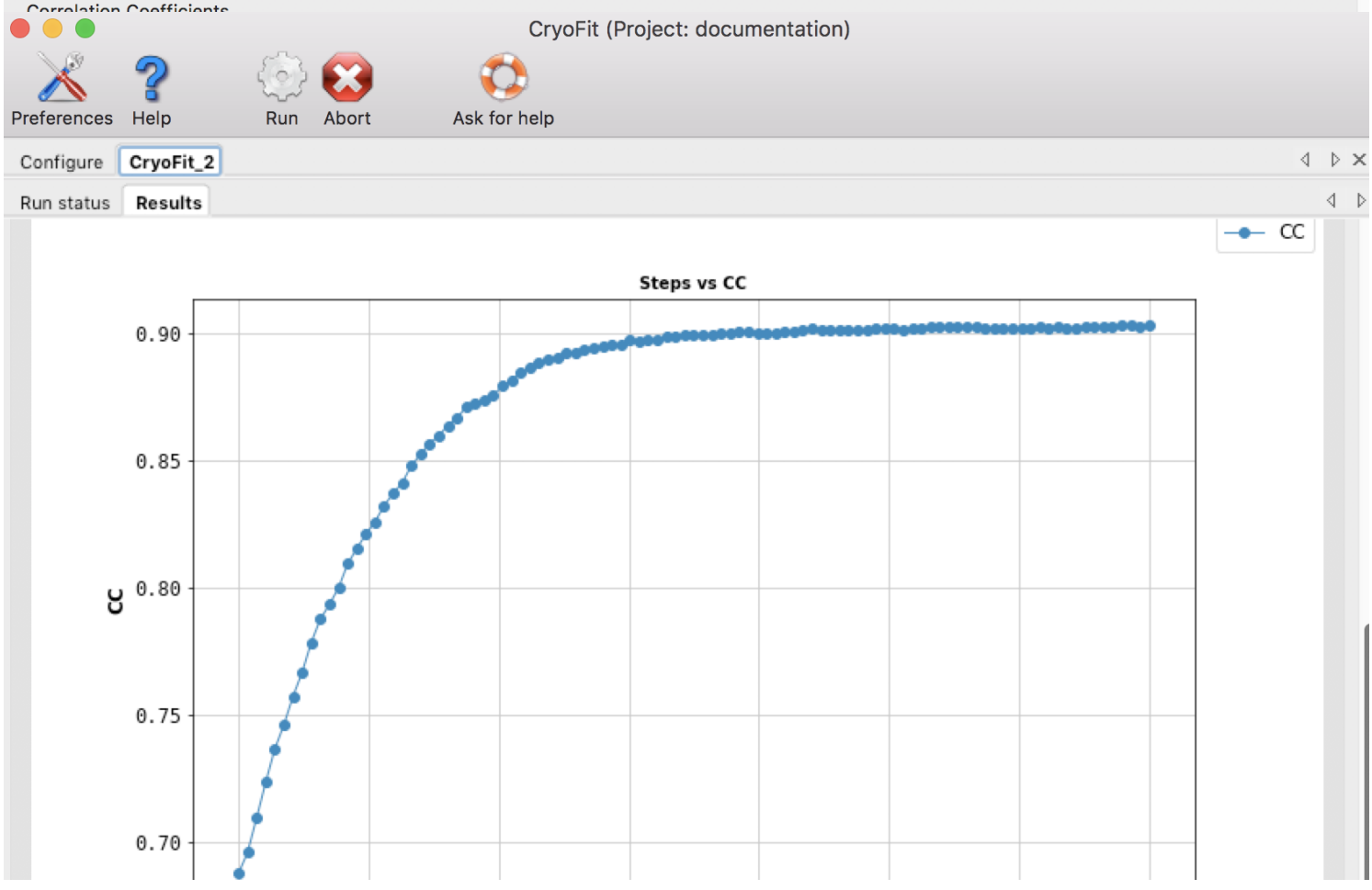


Running time (With 2.7 GHz CPU, macOS)

All default values, it took ~2 minutes With 10k steps, it took ~9 minues

Result


CryoFit (Project: documentation)
Preferences
Help
Run
Abort
Ask for help
Configure
CryoFit_2
Run status
Results
Cryo Fit
Directory: /Users/doonam/research/documentation/CryoFit_2
Filename
Directory
extracted_68600_steps_137.2_ps_chain_recove... /Users/doonam/research/documentation/CryoFit_2/output
cryo_fitted_chain_recovered_cleaned_for_real_s... /Users/doonam/research/documentation/CryoFit_2/output
extracted_70000_steps_140.0_ps_chain_recove... /Users/doonam/research/documentation/CryoFit_2/output
Open in Coot


CryoFit (Project: documentation)
Preferences
Help
Run
Abort
Ask for help
Configure
CryoFit_2
Run status
Results
CC
Steps vs CC
0.90
0.85
0.80
0.75
0.70
CC

Idle Project: documentation

Output files are in steps/8_cryo_fit folder

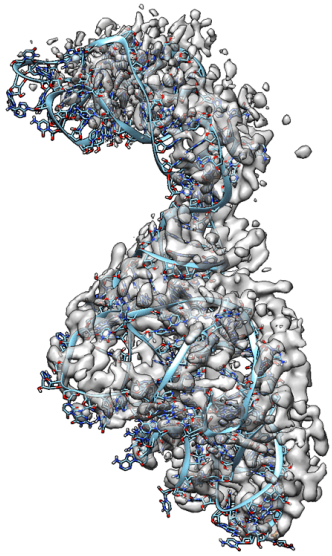
A finally fitted atomic model: cryo_fitted.pdb

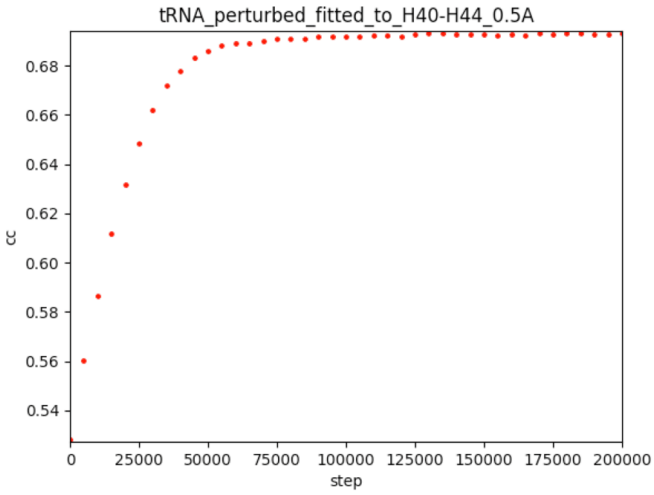
.gro and .pdb files from the highest 3 cc values: extracted_x_steps_x_ps.gro/pdb (.gro files are for vmd)

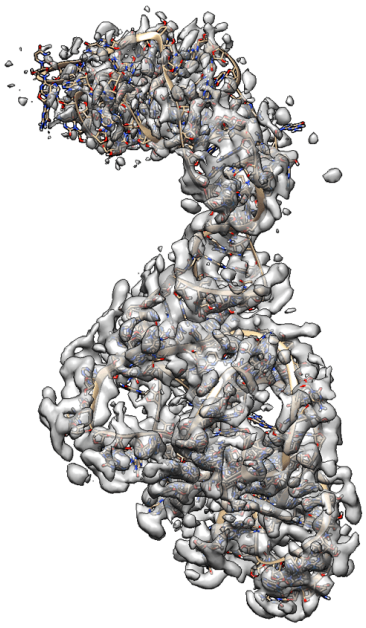
CC means 'correlation coefficient between atomic structure and cryo-EM map': steps/8_cryo_fit/cc_record

Output Examples

With tRNA







Before cryo_fit (cc = 0.53)

After cryo_fit (cc = 0.69)

Limitation

gromacs 4.5.5 seems to not handle H2O (water) heteroatom. cryo_fit will remove water molecules (if any) from the input .cif/.pdb and fit to cryo EM map.

cryo_fit doesn't handle non-canonical "residue"s such as 7C4, BMA, GDP, ILX, NAG, SEP, TRX. The cryo_fit will simply erase those residues.

Reference

S. Kirmizialtin, J. Loerke, E. Behrmann, C. MT. Spahn, K. Y Sanbonmatsu, Using Molecular Simulation to Model High-Resolution Cryo-EM Reconstructions, Methods Enzymol., 558, 2015, 497-514

List of most useful options

Option	Default value	Description of inputs and uses
emweight_multiply_by	8	Multiply by this number to the number of atoms for weight for cryo-EM map bias. For example, emweight = (number of atoms in gro file) x (emweight_multiply_by which is 8) The higher the weight, the stronger bias toward EM map rather than MD force field and stereochemistry preserving constraints. If user's map has a better resolution, higher value of emweight_multiply_by is recommended since map has much information. If user's map has have a worse resolution, lower value of emweight_multiply_by is recommended for more likely geometry. If CC (correlation coefficient) needs to be improved faster, higher number of emweight_multiply_by is recommended.
number_of_cores_to_use	max cores	Specify number of cores for minimization and cryo_fit. If it is not specified, or max is chosen, the cryo_fit will try to use most cores automatically (up to 16)
number_of_steps_for_cryo_fit	None	This is the initial number of steps for cryo_fit. Eventually, cryo_fit will increase it depending on molecule size and cc trend. For tutorial files, this will be 70,000
number_of_steps_for_minimization	None	Specify number of steps for minimization. If this is left blank, cryo_fit will estimate it depending on molecule size.number of steps for cryo_fit. Enough minimization will prevent "blow-up" during MD simulation later.