



*Python-based **H**ierarchical **EN**vironment for **I**ntegrated **X**tallography*

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

Contents

- [Tutorial: Fit Biomolecules into Cryo-EM Maps using MD Simulation \(GUI\)](#)
- [Overview](#)
- [Input files](#)
- [Initial Model](#)
- [Target Map](#)
- [Launch cryo_fit](#)
- [Enter Input Files](#)
- [Specify cryo_fit executable location](#)
- [Enter Options](#)
- [Run](#)
- [Result](#)
- [Output Examples](#)
- [Limitation](#)
- [Reference](#)
- [List of most useful options](#)

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 (Sitius style in text file)

Launch cryo_fit

PHENIX home

Quit Preferences Help Citations Coot PyMOL KiNG Other tools Ask for help

Actions Job history

Projects

Show group: All groups Manage...

Select Delete New project Settings

ID	Last modified	# of jobs	R-free
devel	Oct 04 2018 11:57 ...	41	---
✓ documentation	Oct 09 2018 10:04 ...	0	---

Favorites

- Data analysis
- Experimental phasing
- Molecular replacement
- Model building
- Refinement
- Cryo-EM
 - Mtriage**
Analyze quality of maps in CCP4 format
 - Map to Model**
Model-building into cryo-EM and low-resolution maps
 - CryoFit**
Fit a model to a cryo-EM Map using MD
 - Real-space refinement**
Automated refinement using real-space maps (Cryo-EM, X-ray, ...)
 - Comprehensive validation (cryo-EM)**
Model quality assessment, including real-space correlation, for cryo-EM structures
 - EMRinger**
Model validation for de novo electron microscopy structures
 - Autosharpen Map**
Tool for sharpening a map
 - Dock in map**

Current directory: /Users/doonam/research/documentation Browse...

PHENIX version 1.14-3260-000 Project: documentation

Enter Input Files

Click browse buttons.

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

CryoFit (Project: documentation)

Preferences Help Run Abort Ask for help

Configure

Job title :

Input

Starting model file : Browse...

Target map file : Browse...

gromacs_cryo_fit executable path : Browse...

Options

Constraint algorithm minimization :

EM steps :

EM weight multiply by this number :

EM write frequency :

Number of steps for minimization :

Number of steps for cryo_fit :

Time step for MD simulation during cryo_fit :

Time step for minimization :

Output

Idle Project: documentation

Specify cryo_fit executable location

Click browse buttons.

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

CryoFit (Project: documentation)

Preferences Help Run Abort Ask for help

Configure

Job title :

Input

Starting model file : Browse...

Target map file : Browse...

gromacs_cryo_fit executable path : Browse...

Options

Constraint algorithm minimization :	<input type="text" value="default"/>	EM steps :	<input type="text"/>
EM weight multiply by this number :	<input type="text" value="8"/>	EM write frequency :	<input type="text"/>
Number of steps for minimization :	<input type="text"/>	Number of steps for cryo_fit :	<input type="text"/>
Time step for MD simulation during cryo_fit :	<input type="text" value="0.002"/>	Time step for minimization :	<input type="text" value="0.001"/>

Output

Idle Project: documentation

Enter Options

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

Run

CryoFit (Project: documentation)

Preferences Help Run Abort Ask for help

Configure

Job title :

Input

Starting model file : Browse...

Target map file : Browse...

gromacs_cryo_fit executable path : Browse...

Options

Constraint algorithm minimization : <input type="text" value="default"/>	EM steps : <input type="text"/>
EM weight multiply by this number : <input type="text" value="8"/>	EM write frequency : <input type="text"/>
Number of steps for minimization : <input type="text"/>	Number of steps for cryo_fit : <input type="text"/>
Time step for MD simulation during cryo_fit : <input type="text" value="0.002"/>	Time step for minimization : <input type="text" value="0.001"/>

Output

Idle Project: documentation

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

CryoFit (Project: documentation)

Preferences Help Run Abort Ask for help

Configure **CryoFit_1**

Run status

Status

Running on pn1803052.lanl.gov (PID 2626)

Log output

```

-----
cryo_fit 1.14-3260
- Doo Nam Kim, Nigel Moriarty, Serdal Kirmizialtin, Billy Poon, Karissa Sanbonmatsu
-----
Input parameters: ['/Users/doonam/research/documentation/.phenix/project_data/run_1.eff']
Reading user provided map started...
(If a user provided a big .sit file like 1.6 GB, this may take more than 6 minutes)
cryo_fit {
  job_title = None
  Input {
    model_file_name = /Users/doonam/bin/phenix-1.14-3260/modules/cryo_fit/tutorial_input_files/
GTPase_activation_center_tutorial.pdb
    map_file_name = /Users/doonam/bin/phenix-1.14-3260/modules/cryo_fit/tutorial_input_files/
GTPase_activation_center_tutorial_gaussian_lp5.mrc
    cryo_fit_path = /Users/doonam/bin/cryo_fit/bin
  }
  Steps {
    step_1 = True
    step_2 = True
    step_3 = True
    step_4 = True
    step_5 = True
    step_6 = True
  }
}

```

Pause Abort View log

1 job(s) running Project: documentation

Running time (With 2.7 GHz CPU, macOS)

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

Result

CryoFit (Project: documentation)

Preferences Help Run Abort Ask for help


Configure **CryoFit_2**

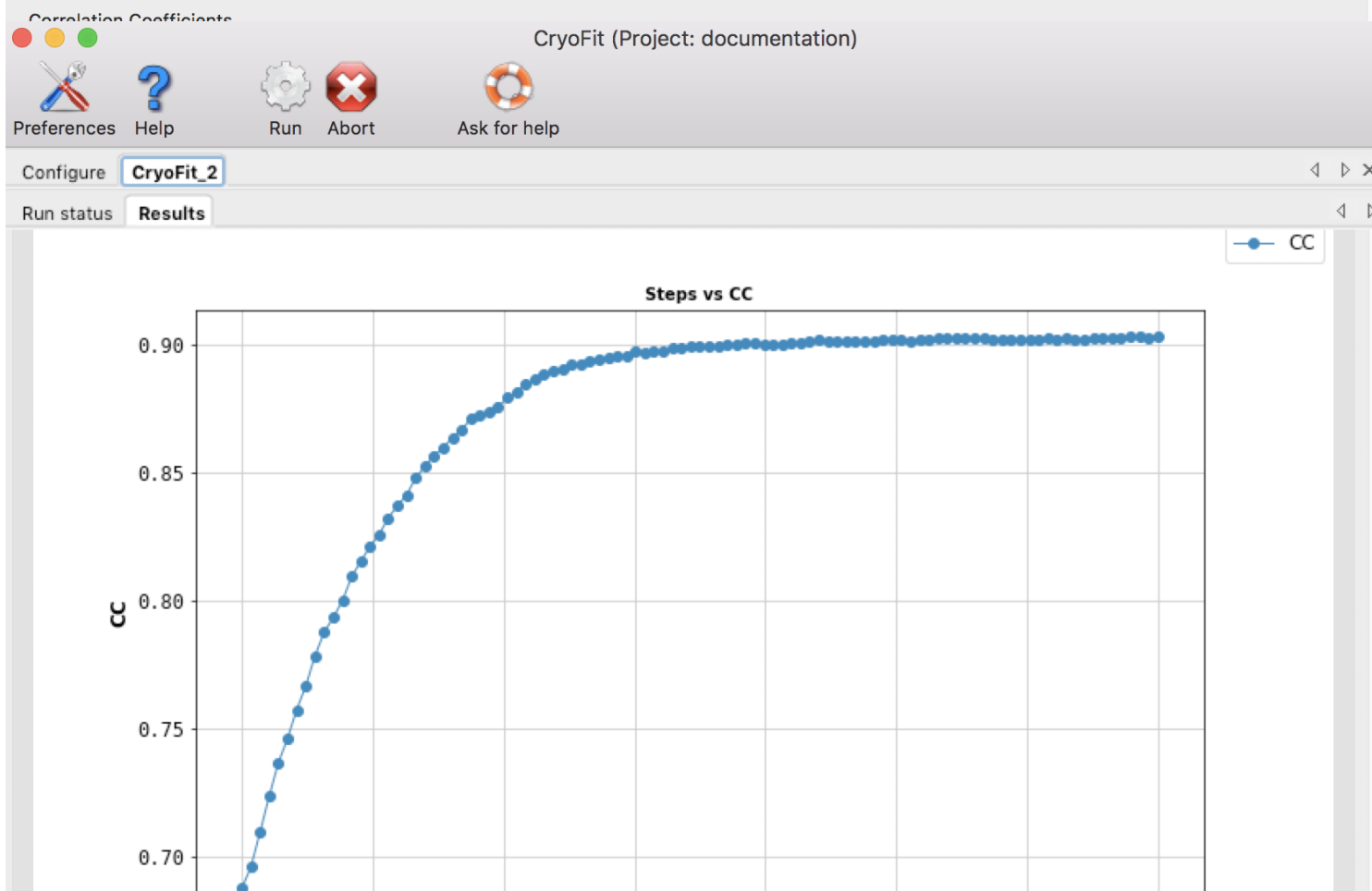
Run status **Results**

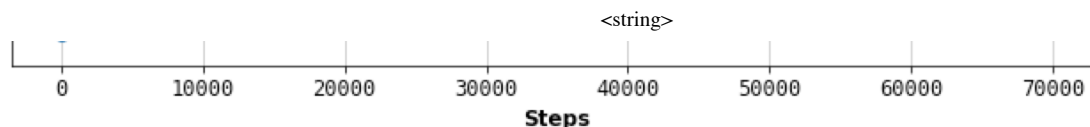
Cryo Fit

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





Idle

Project: documentation

Output files are in steps/8_cryo_fit folder

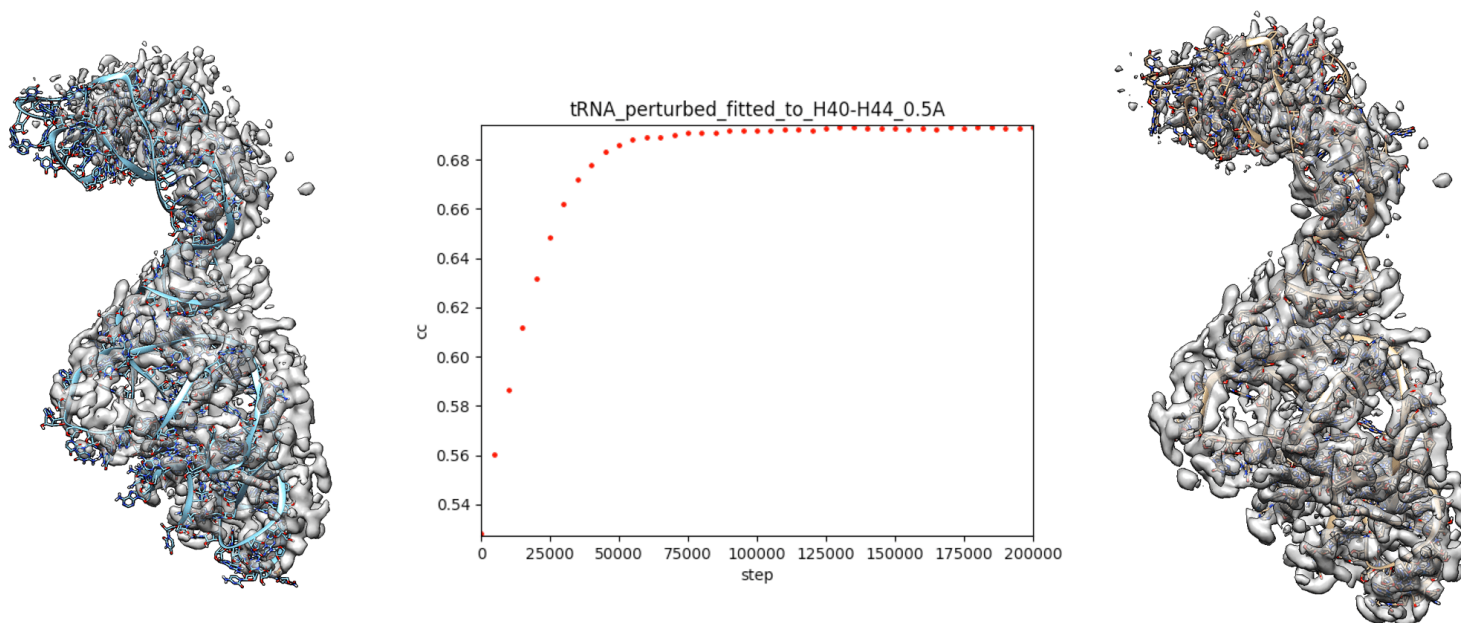
A finally fitted atomic model: cryo_fitted.pdb

.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.