



JOURNAL OF
APPLIED
CRYSTALLOGRAPHY

Volume 56 (2023)

Supporting information for article:

Remote laboratory training for high school students: grocery store based hands-on project in protein crystallography

Ali Lillian Fox, Alyssa R. Teteris and Irimpan I. Mathews

Supplementary Materials:

PyMOL

Download pymol and open the Pymol graphic window.

(PyMOL: <https://pymol.org/2/>)

Type the following commands:

fetch 7knk (this will automatically download the pdb file and display it as “7knk”. PDB code 7knk corresponds to the structure reported in this paper.)

fetch 7knk, type=2fofc (will load the 2Fo-Fc map as “7knk_2fofc”)

fetch 7knk, type=fofc (will load the Fo-Fc map as “7knk_fofc”)

center on 7knk (Select the "center" option by clicking on the "A" next to the "7knk" name on the right panel next to the figure display.

isomesh 2fofcmap,7knk_2fofc, 1.0, poly, carve=2.0 (will display the 2Fo-Fc map for the whole molecule in white color).

color this map blue (select color by clicking on the “C” next to the “7knk_2fofc” name).

isomesh fofcmap,7knk_fofc, 2.5, poly, carve=2.0 (this will display difference maps in white color).

To electron density map for residues 20 to 30:

Type these commands:

select site1, resi 20-30 (this will create an object named “site1” with residues from 20 to 30.

isomesh 2fofcsite1map,7knk_2fofc, 1.0, site1, carve=2.0 (this will display the map for “site1”)

color magenta

hide the previous 2fofcmap (if not, the new map color may not be visible)

(Note: to hide the map “select the option by clicking on the “H” next to the “7knk_2fofc” name).

To Superimpose structures:

Download 7knk.pdb and 135l.pdb from the protein data bank (vide infra)

Load PDB file 1: eg. File > Open > 7knk.pdb

Load PDB file 2: eg. File > Open > 135l.pdb (PDB code “135L” corresponds to the Turkey lysozyme).

type the commad: align 135l, 7knk

This will superpose 135l structure onto 7knk

center on 7knk

Coot

Download Coot crystallographic program.

(Coot: <https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/>)

Search for Protein Data Bank

type 7knk in the "3D Structures" prompt

Download PDB format. Downloaded file name: 7knk.pdb

Download Map Coefficients (MTZ format). Downloaded file name: 7knk_phases.mtz

File > Open Coordinate > 7knk.pdb (this will display the structure).

File > Open MTZ, mmCIF, fcf or phs > 7knk_phases.mtz

Amplitudes FWT

Phase PHWT

Click "OK"

(This will display the 2Fo-Fc map (blue color))

File > Open MTZ, mmCIF, fcf or phs > 7knk_phases.mtz

Amplitudes: FWT

Select Phase as: PHDELWT

Select "Is a Difference Map" option

click "OK"

(This will display Fo-Fc map in green and Fc-Fo map in red).

To superimpose structures:

Load the structures.

Go to Calculate > SSM Superpose >

Select the reference structure: One of the loaded pdb file

Select the Moving structure: Another loaded pdb file

This will superpose the second structure on to the first structure.

Protein Data Bank (PDB)

Search for PDB and locate the PDB site. Eg. <https://www.rcsb.org/>

Go to the PDB site and search for a name of the protein or type the PDB code in the “3D Structures” section. For example: 7knk

This will display details of the crystal structure for PDB code 7knk.

The structural information is stored in various tabs (Structure Summary, 3D View, Annotations, Experiment, Sequence, etc.).

Generally, the “Structure Summary” tab will be displayed.

This shows the display of the structure on the left and text summary of the structure and data on the right. Scrolling down the page shows more details on the structure.

The top right corner of the page has pull down menus with “Display Files” and “download Files”.

By clicking on the “Display Files”, users can display FASTA sequence of the structure and various formats of the deposited data.

The “Download Files” section allows downloading of the FASTA sequence, PDB coordinate, structure factors, maps etc.

The PDB format display shows Title of the PDB entry, name of the protein, related publication, details of the crystal data, refinement statistics, details of the secondary structure, sequence information, unit cell dimensions and space group, and coordinate.

Each line of the coordinate section has

```
ATOM 1 N LYS A 1 -9.800 2.934 -17.926 1.00 20.50 N
```

Numbers “-9.800 2.934 -17.926” gives the X, Y, Z position of the corresponding atom in orthogonal coordinates. Next number “1” gives the occupancy of that atom. This is followed by the b-factor (here, 20.50).

More details about the PDB data is available at <https://pdb101.rcsb.org/learn/guide-to-understanding-pdb-data/introduction>

Crystal mounting video links (provide by Dr. Jim Pflugrath) :

<https://www.youtube.com/watch?v=45Qc3jOPaKY>

<https://www.youtube.com/watch?v=1y89sfsuQBw>

Structure solution and Refinement

Phenix (<https://phenix-online.org/>).

Phenix can be downloaded from this site.

Phenix website has video tutorials and documentation for the various programs.

Start Phenix software, this will display the following window.

PHENIX home

File Projects Utilities Help

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

Actions Job history

Projects

Show group: PHENIX tutorials Manage...

Select Delete New project Settings

ID	Last modified	# of jobs	R-free
✓ tutorial	Feb 07 2023 12:40...	4	0.2902
nsf-d2-ligand_0	May 09 2019 04:2...	0	---
nsf-d2-ligand	May 09 2019 04:2...	2	---
p9-sad_iimath...	May 09 2019 03:3...	1	0.2366
beta-blip	Apr 25 2018 03:22...	2	---

Favorites

AlphaFold (predicted models)

Crystals: Data analysis and manipulation

Validation and map-based

Experimental phasing

Molecular replacement

Maps (create, manipulate)

Enhanced maps (Polder, F)

Model building

Refinement

Ligands

Current directory: /mnt/home1/iimathew Browse...

PHENIX version 1.20.1-4487-000 Project: tutorial

Click on the “New Project”

This gives the following window with an option to input the sequence file.

Create project

Please choose a simple project ID (alphanumeric and underscore characters only) and project directory. The directory should not overlap with a previously defined project directory. PHENIX will store settings and files specific to this project in a subdirectory named '.phenix'.

Project ID :

Project directory :

Sequence file :

Add to group : ▼

Switch to this project

Give a Project ID:

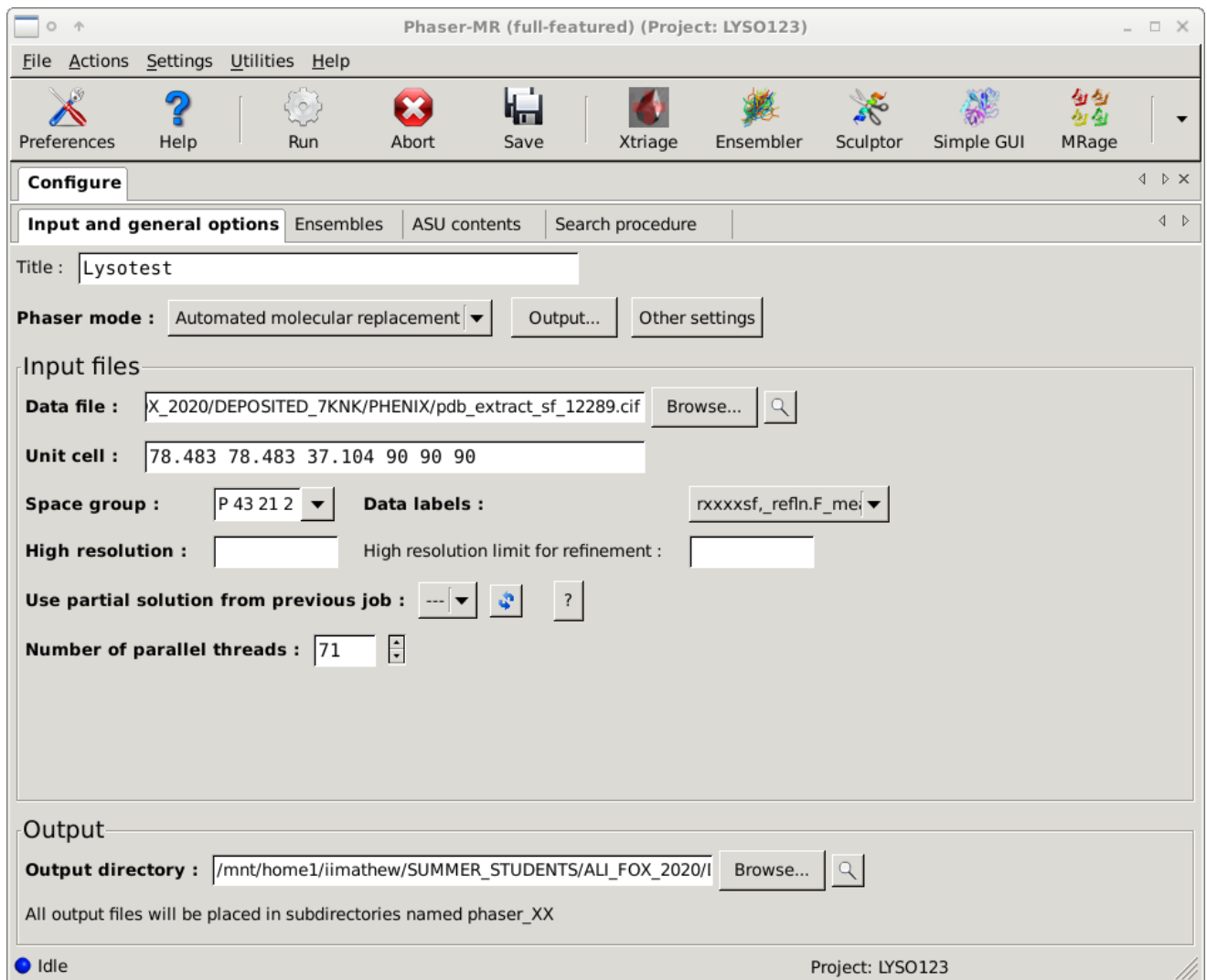
Upload the sequence file for the 193L structure in the Supplementary section or download the sequence file from the PDB ID: 193L (from the PDB site).

Click the “OK” button.

Go to the first Phenix widget and click on the “Molecular Replacement”

This will open the following widget. This widget had different Tabs to input to fill.

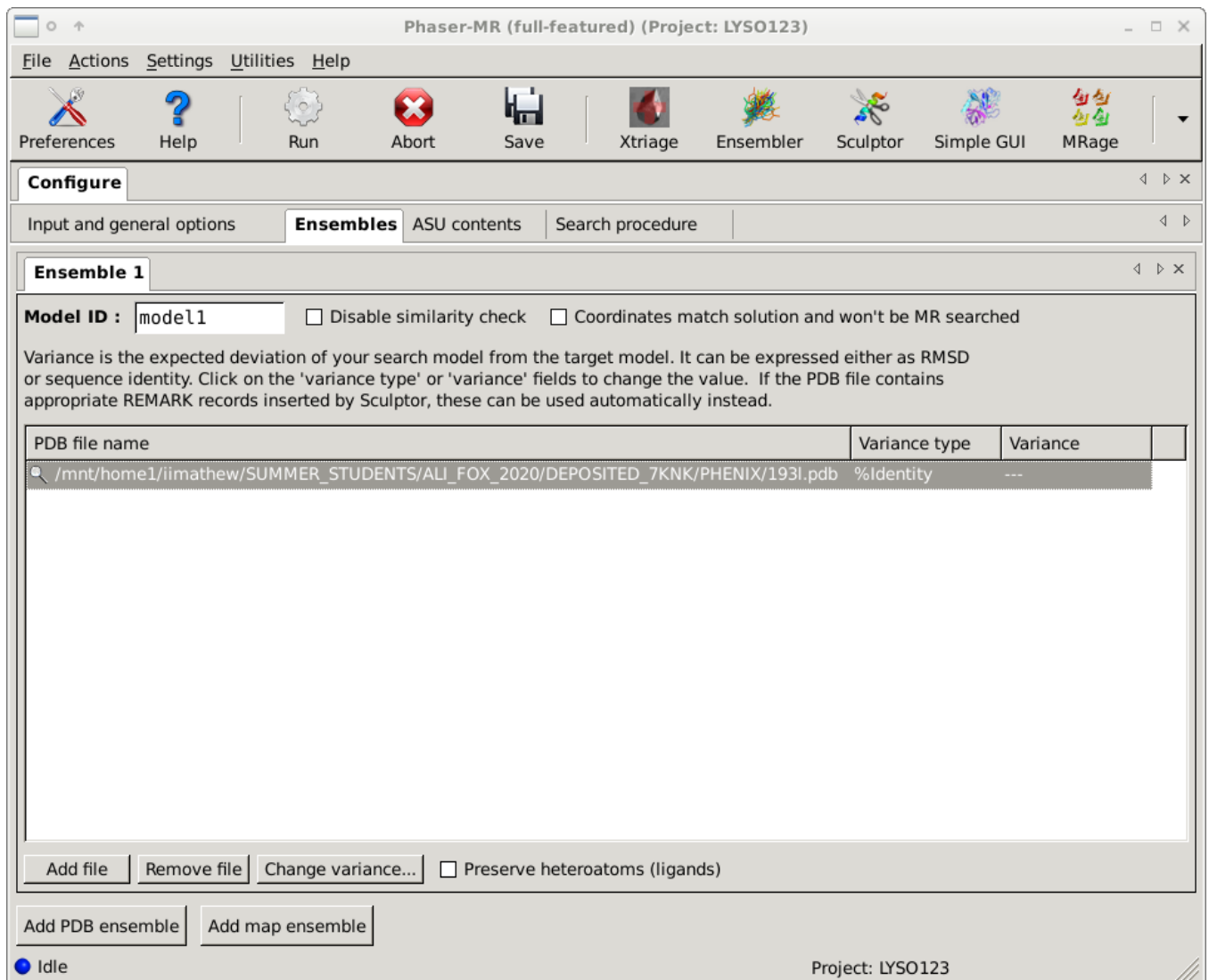
The first Tab is “Input and general options” shown below.



Put a Title and load the reflection file (pdb_extract_sf_12289.cif) in the Supplementary section.

Uploading this file will automatically fill the Unit cell and space group.

Then go to the "Ensembles" Tab (see the picture below).



Put a Model ID: (eg. model1)

Click on the “Add File” and load the pdb file: 193l.pdb (from the Supplementary section or download it from the PDB site).

Click on the “Change variance”. This will open a new window as shown below.

Model variance

Please enter the variance of this search model from the expected structure.
Fractional values for sequence identity will be converted into percent values.

File name:

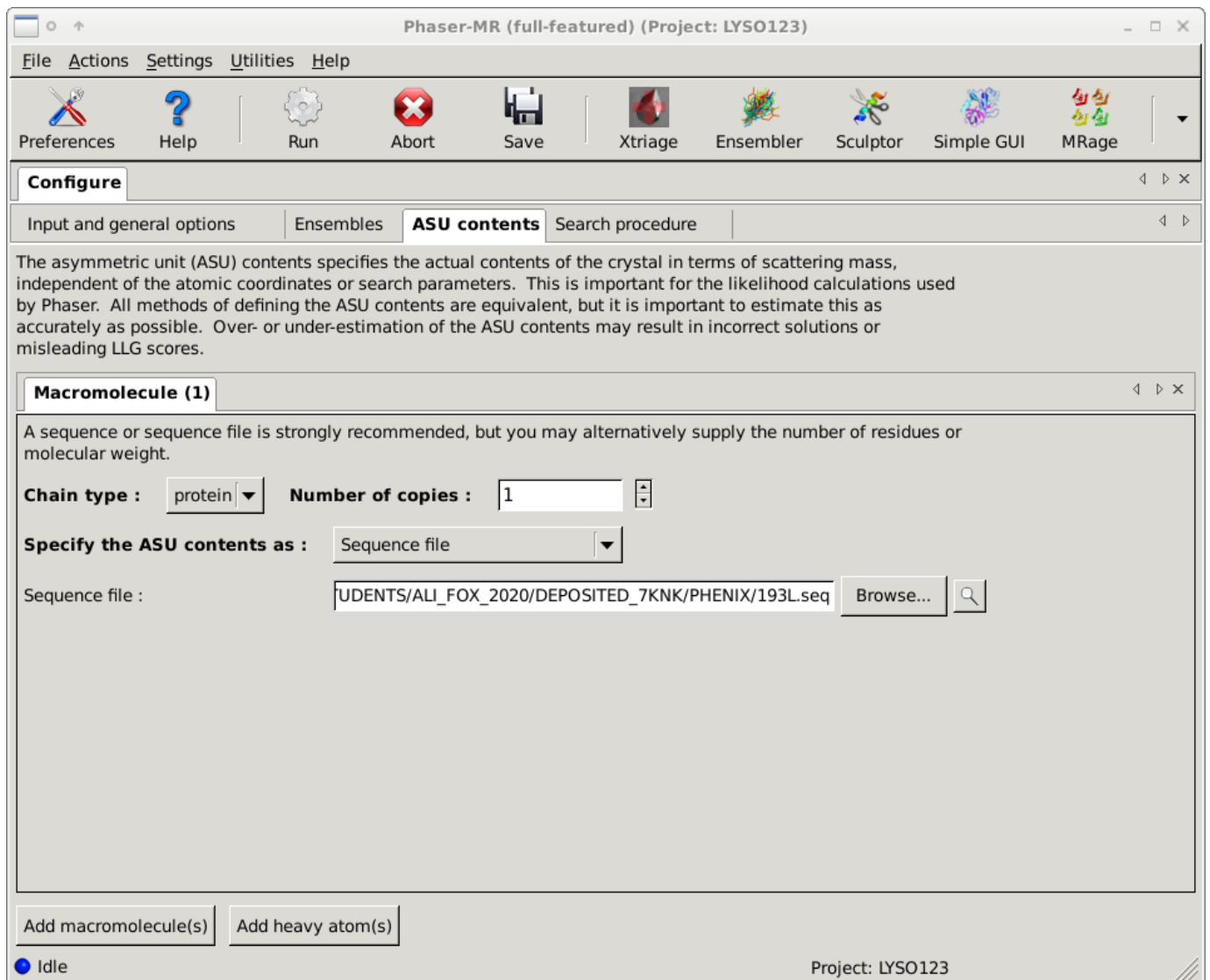
Sequence identity (%)

Automatically extract identity from PDB REMARK section

Put the sequence identity: 95

Click OK to close this widget.

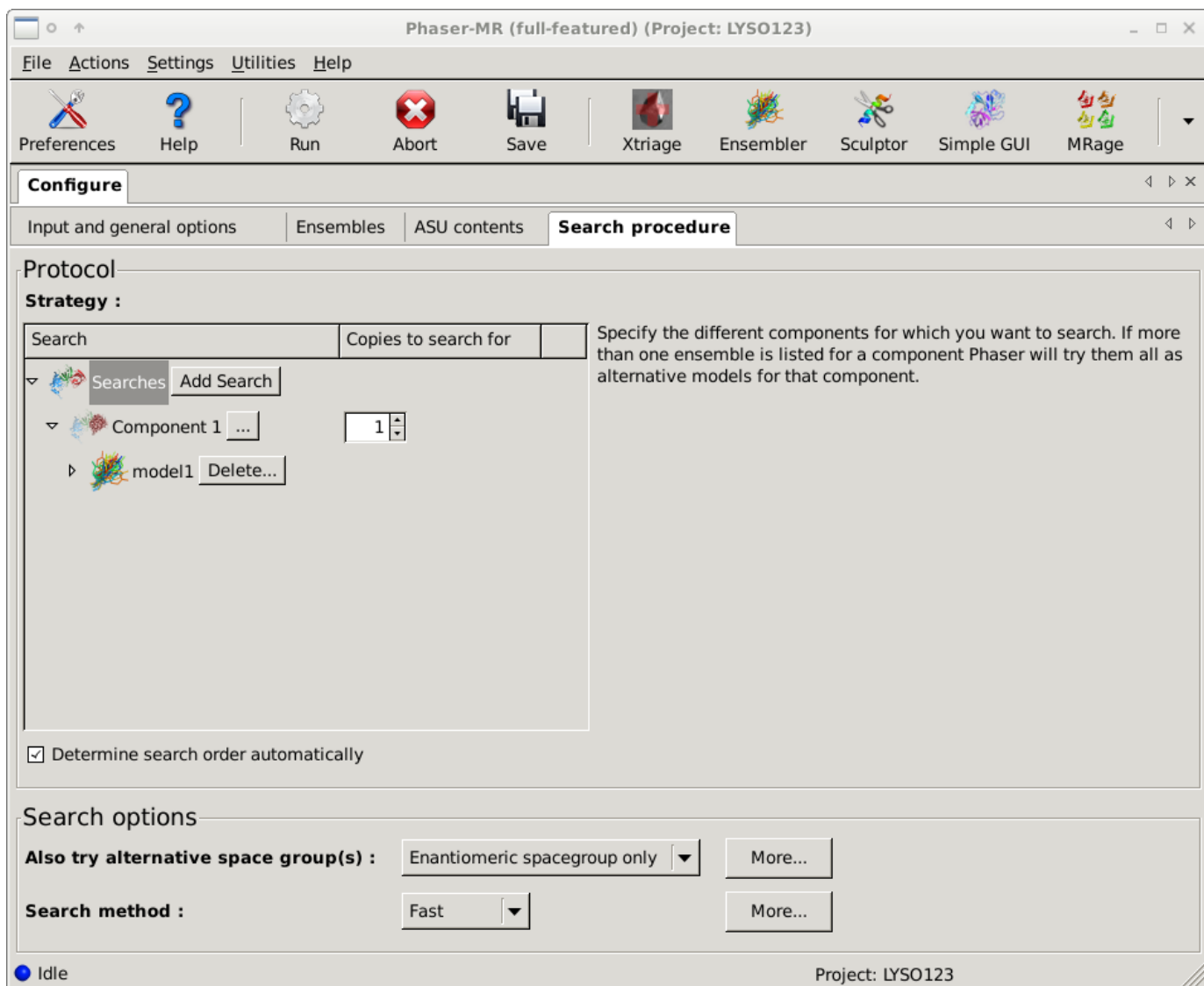
Then, go to the ASU contents Tab (see below)



Number of copies: 1

Upload the sequence file for the 193L structure provided in the Supplementary or download it from the PDB site.

Now go to the "Search Procedure" Tab (see below).



Click on the Component 1, select Add ensemble and select “model1” (this is the Model ID we provided in the Ensembles Tab).

Select “Run” from the top bar and select “Run now”. This will perform the structure solution using the Molecular Replacement method.

Program will run for some time (depending on the CPU power of the machine). There will be constant updates on the window.

When the program completes, it will display the following window.

Phaser-MR (full-featured) (Project: LYSO123)

File Actions Settings Utilities Help

Preferences Help Run Abort Save Xtrriage Ensembler Sculptor Simple GUI MRage

Configure phaser_1

Run status Graphs Results

Summary

Phaser has found 1 MR solution(s).

Top LLG: 5692.300

Top TFZ: 57.9

Spacegroup: P 43 21 2

Run phenix.refine

Run AutoBuild

Output files

Title: Lysotest

Directory: /mnt/home1/iimathew/SUMMER_STUDENTS/ALI_FOX_2020/DEPOSITED_7KNK/PHENIX/phaser_1

File name	Contents
LYSO123_phaser.1.mtz	Phases and maps
LYSO123_phaser.1.pdb	Output model

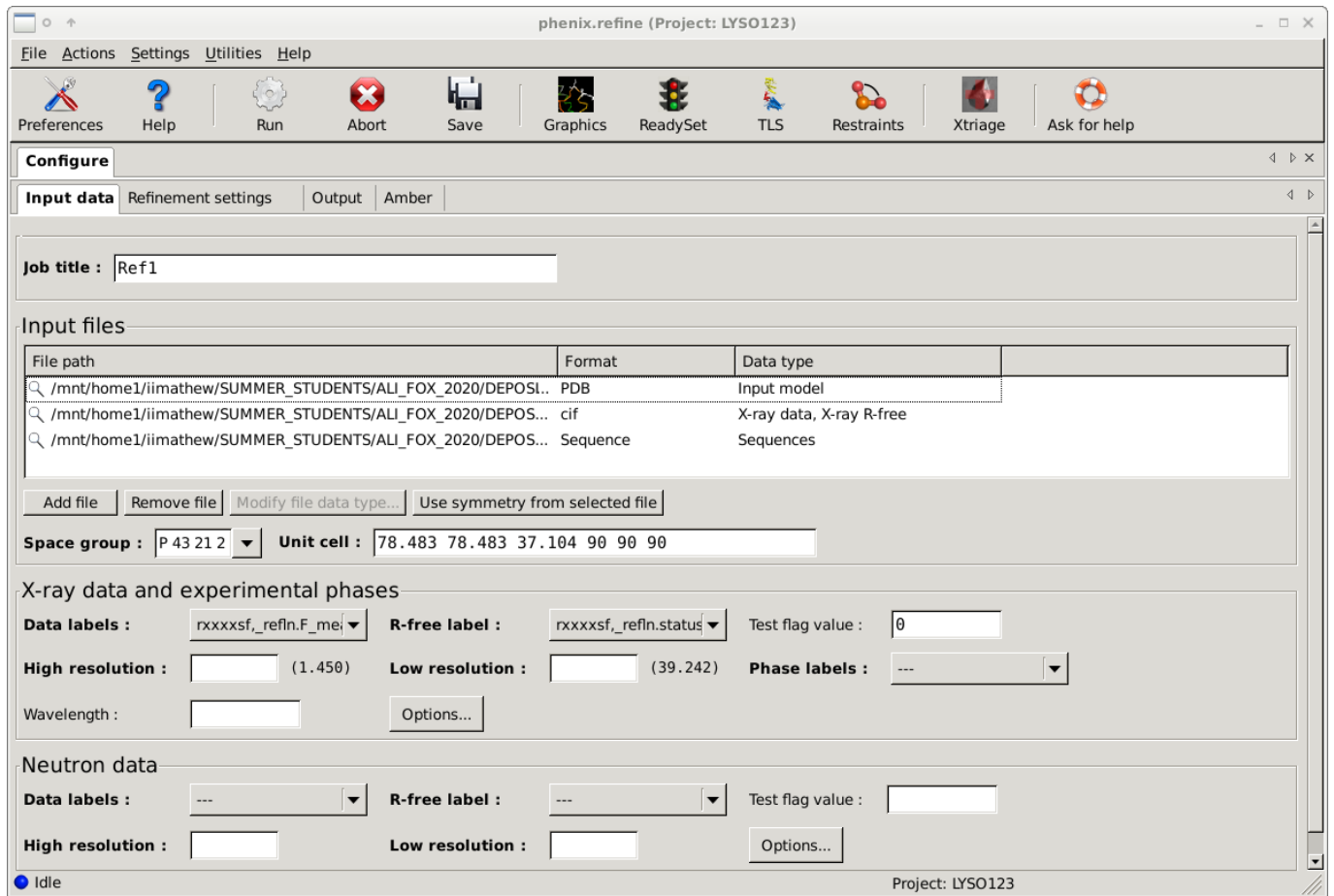
Open in Coot Open in PyMOL

Idle Project: LYSO123

The high values for Top LLG (larger the better) and Top TFZ (higher than 8) show that structure is solved.

Click on the “Run phenix.refine” button.

This will open the following window. Here also information is displayed in a “Tab” format.

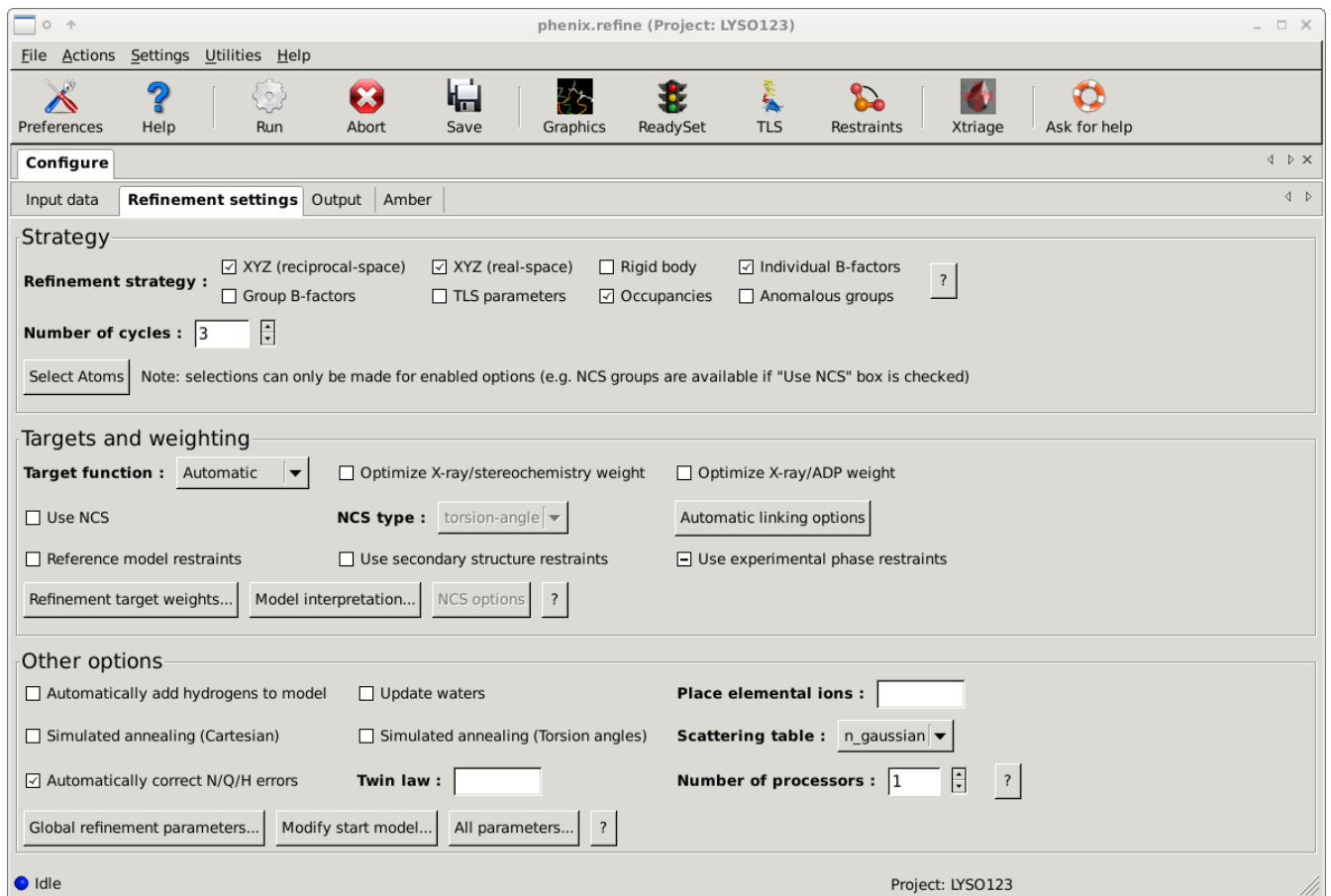


The first Tab is "Input Data".

Put a Job title (eg: Ref1).

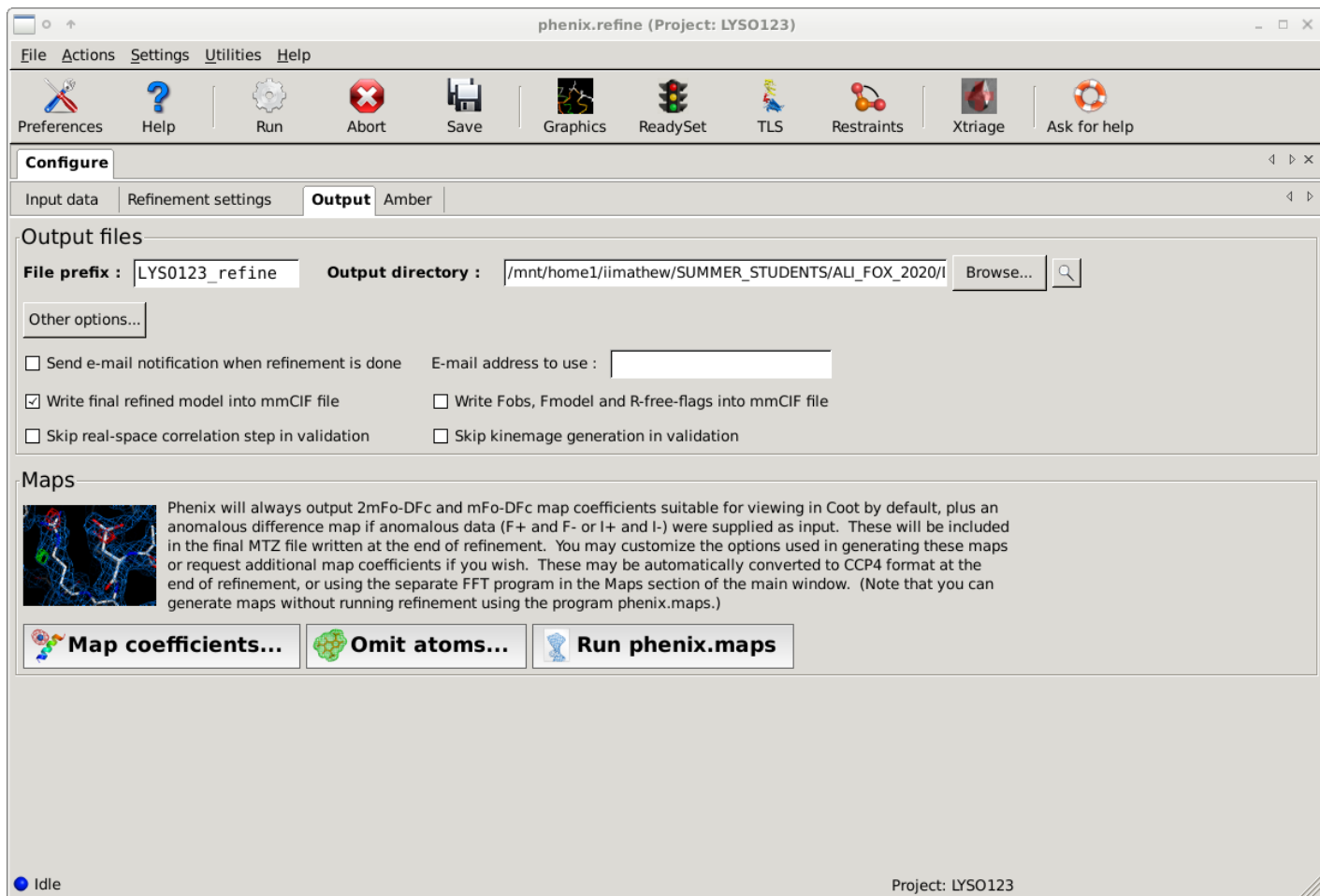
The files needed for the refinement (coordinate, reflection data and sequence) are already uploaded in the window.

Go to the "Refinement settings" Tab.



No need to change any information. However, you could increase “Number of Processors” to higher values, depending on the computing power of the computer.

Select the “Output” Tab



Program automatically fill the “File prefix” based on the two titles we gave at the beginning of Phenix and in the “Input” section of “Refinement”.

Select “Run” from the top bar and select “Run now”. This will run the “Refinement” of the structure obtained in the MR.

When the program completes, it will display the following window.

phenix.refine (Project: LYSO123)

File Actions Settings Utilities Help

Preferences Help Run Abort Save Graphics ReadySet TLS Restraints Xtrriage Ask for help

Configure Refine_2

Log output Run status **Results** MolProbrity Real-space correlation Atomic properties Sequence check

Output files

Title: Ref1

Directory: /mnt/home1/iimathew/SUMMER_STUDENTS/ALI_FOX_2020/DEPOSITED_7KNK/PHENIX/Refine_2

File name	Contents
LYSO123_refine_2.log	phenix.refine log file
LYSO123_refine_data.mtz	Data and R-free set used in refine...
LYSO123_refine_2.geo	Geometry restraints before refine...
LYSO123_refine_2.eff	Effective parameters for this run
LYSO123_refine_2.pdb	Refined model

Convert map coefficients MTZ to CCP4 maps

Open in Coot

Open in PyMOL

Sequence viewer

Refinement statistics

Compare statistics Plot statistics by cycle Plot statistics by resolution

Before and after refinement:

	Starting	Final
R-work	0.3103	0.2625
R-free	0.3220	0.2997
Bonds	0.009	0.006
Angles	1.385	0.955

X-ray statistics by resolution bin:

Idle Project: LYSO123

This widget also has tabs. A short summary of the refinement results is provided in the “Results” tab (shown the above window). Other Tabs provides various information resulted from the Structure solution.

Clicking on the “Open in Coot” will display structure and electron density using the “Coot” program. (NOTE: the proper display require proper setting of the various files).