

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 1351.pdb from the protein data bank (vide infra)

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

Load PDB file 2: eg. File > Open > 1351.pdb (PDB code "135L" corresponds to the Turkey lysozyme).

type the commad: align 1351, 7knk

This will superpose 1351 structure onto 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 loded 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 lint 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 <u>https://pdb101.rcsb.org/learn/guide-to-understanding-pdb-data/introduction</u>

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.

□ • ↑		PHEN	IX home _
<u>F</u> ile Projects <u>U</u> tili	ities <u>H</u> elp		
Quit Prefere	ences Help	Citations Coot	Image: Weight of the sector
Actions Job histo	ry		
Projects			Favorites
Show group: PH	ENIX tutorials	▼ Manage	AlphaFold (predicted models)
Select ⊘ De	elete 🛛 🚑 New proj	ject 😵 Settings	Crystals: Data analysis and manipulation
ID	Last modified	# of jobs R-free	Validation and map-based
< tutorial	Feb 07 2023 12:40 4	0.2902	Experimental phasing
nsf-d2-ligand_0	May 09 2019 04:2 0)	Molecular replacement
nsf-d2-ligand	May 09 2019 04:2 2		Mans (create, manipulate
beta-blip	Apr 25 2018 03:22 2		Maps (create, manipulate
			Enhanced maps (Polder, F
			Model building
			Refinement
			Ligands
Current directory	/mnt/home1/iimath	ew	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 s characters only) a overlap with a pre store settings and named '.phenix'.	imple project ID (alphanumeric and underscore and project directory. The directory should not eviously defined project directory. PHENIX will I files specific to this project in a subdirectory	
Project ID :	LYS0	
Project directory :	/mnt/home1/iimathew/SUMMER_STUDENTS/ALI_FOX_2020/I Browse	٩
Sequence file :	/mnt/home1/iimathew/SUMMER_STUDENTS/ALI_FOX_2020/I Browse	٩
Add to group :	None	
	✓ Switch to this project	
	Set up tutorial data	
	Second Cancel	ĸ

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.

• •	Phaser-MR (full-fea	tured) (Project	t: LYSO123)			×
<u>File</u> <u>Actions</u> <u>Settings</u> <u>U</u> tilities <u>Hel</u>	0					
Preferences Help	Abort Save	Xtriage	🎉 Ensembler	Sculptor	Simple GUI	실실 실실 MRage →
Configure						4 Þ ×
Input and general options Ensen	bles ASU contents Sea	arch procedure				4 Þ
Title: Lysotest						
Phaser mode : Automated molecula	r replacement 🔻 Output	t Other se	ettings			
Input files						
Data file : X_2020/DEPOSITED_7KI	NK/PHENIX/pdb_extract_sf_1	2289.cif Brow	se Q			
Unit cell : 78.483 78.483 37.	104 90 90 90					
Space group : P 43 21 2 ▼	Data labels :	a	xxxxsf,_refln.F	_me 🔻		
High resolution :	High resolution limit for re	finement :				
Use partial solution from previou	s job : ▼ 3					
Number of parallel threads : 71		-				
Outrast						
Output			_			
Output directory : /mnt/home1/iin	nathew/SUMMER_STUDENTS	/ALI_FOX_2020/	Browse	9		
All output files will be placed in subdire	ctories named phaser_XX					
• Idle			P	roject: LYSO	123	

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).

o ↑				Phaser-N	/IR (full-feat	ured) (Projec	t: LYSO123)				- • ×
<u>F</u> ile <u>A</u> ctions	<u>S</u> ettings	<u>U</u> tilitie	s <u>H</u> elp								
Preferences	? Help		() Run	E Abort	Save	Xtriage	j 🎉 Ensembler	X Sculptor	Simple (ෂ්න් මුණු GUI MRage	
Configure											4 Þ ×
Input and ge	neral optio	ns	Ensem	bles ASU con	tents Sear	ch procedure					4 ⊳
Ensemble	1										4 Þ ×
Model ID :	model1		🗌 Dis	able similarity	check 🗌 Co	oordinates ma	atch solution an	d won't be l	MR search	ed	
Variance is th or sequence appropriate F	ne expected identity. Cli REMARK red	d deviat ck on th cords ins	ion of yo e 'varian serted by	ur search mode ice type' or 'var Sculptor, these	l from the targ iance' fields to can be used	get model. It of the state of t	can be expresse value. If the PE r instead.	ed either as DB file conta	RMSD iins		
PDB file nar	ne							Varianc	e type	Variance	
Add file	Remove	file Ch	lange var	riance Pr	eserve hetero	atoms (ligano	ds)				
Add PDB ens	emble A	dd map	ensembl	e							
😑 Idle							P	roject: LYSO	123		

Put a Model ID: (eg. model1)

Click on the "Add File" and load the pdb file: 1931.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									
<pre>values. File name: /mnt/home1/iimathew/SUMMER_STUDENTS/A</pre>									
Sequence i	dentity (%) 💌 = 95								
Automat	Automatically extract identity from PDB REMARK section								
	Sancel	<u>9о</u> к]						

Put the sequence identity: 95

Click OK to close this widget.

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

	○ ↑ Phaser-MR (full-featured) (Project: LYSO123) _ □ ×										
<u>F</u> ile <u>A</u> ctions	<u>S</u> ettings	<u>U</u> tilitie	es <u>H</u> elp								
Preferences	? Help		رکی Run	E Abort	Save	Xtriage	🎉 Ensembler	Sculptor	Simple GUI	성 성 신 십 MRage	•
Configure											4 Þ ×
Input and ger	neral option	ns	Ensembles	ASU con	itents	Search procedure					4 ⊳
The asymmetr independent of by Phaser. All accurately as p misleading LLC	The asymmetric unit (ASU) contents specifies the actual contents of the crystal in terms of scattering mass, independent of the atomic coordinates or search parameters. This is important for the likelihood calculations used by Phaser. All methods of defining the ASU contents are equivalent, but it is important to estimate this as accurately as possible. Over- or under-estimation of the ASU contents may result in incorrect solutions or misleading LLG scores.										
Macromole	cule (1)										4 Þ ×
A sequence of molecular we Chain type Specify the Sequence file	r sequence ight. ASU con :	tents a	Number o Is : Sequ UDEN	mmended, b	ut you ma 1 2020/DEI	ay alternatively s	upply the num	q Browse	Jes or		
Add macromo	lecule(s)	Add h	eavy atom(s	;)							
😑 Idle							F	roject: LYSO	123		

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).

o ↑	Phaser-M	R (full-featu	red) (Project:	: LYSO123)			-	
<u>File Actions Settings Utilities Hel</u>	р							
Preferences Help	E Abort	Save	Xtriage	🎉 Ensembler	X Sculptor	Simple GUI	생성 성실 MRage	•
Configure								⊲ ⊳×
Input and general options Ensen	nbles ASU conte	ents Sear	rch procedur	e				⊲ ⊳
Protocol Strategy :								
Search	Copies to search f	or	Specify the diff than one enser	ferent compo nble is listed	nents for wh for a compo	hich you want to onent Phaser wi	search. If r Il try them a	nore all as
✓ Jearches Add Search		1	alternative mo	dels for that	component.			
▽ 🧼 Component 1 📖	1 .							
model1 <u>Delete</u> Determine search order automatica	lly							
Soarch ontions								
Also try alternative space ground	(s) : Enantion	aric spacegro	up only	More				
Also d y alternative space group(sine spacegro]			
Search method :	Fast	•		More				
• Idle				Р	roject: LYSO	123		

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-N	MR (full-featu	red) (Projec	t: LYSO123)			_	
<u>File</u> <u>Actions</u> <u>Sec</u>	ettings <u>U</u> tiliti	es <u>H</u> elp								
Preferences	? Help	ිටි Run	Abort	Save	Xtriage	Ensembler	X Sculptor	Simple GUI	실실 실실 MRage	•
Configure ph	aser_1									4 Þ ×
Run status Gr	raphs Resu	lts								4 ⊳
Summary Phaser has found 1 MR solution(s). Top LLG: 5692.300 Top TFZ: 57.9 Spacegroup: P 43 21 2 Run AutoBuild										
Title: Ly Directory: /m	/sotest mnt/home1/ii	mathew/SUM	MER_STUDEN	TS/ALI_FOX_2	020/DEPOSI	TED_7KNK/PHI	ENIX/phaser	_1		
File name				Contents						
Q LYSO123_pha	aser.1.mtz			Phases and ma	aps					
Copen i	Q LYS0123_phaser.1.mtz Phases and maps Q LYS0123_phaser.1.pdb Output model									
o Idle							Project: LYSO	123		///

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.

🔲 o 🛧 phenix.refine (Proje	ect: LYS0123) _ 🗆 🗆 🛛
<u>File Actions Settings Utilities Help</u>	
X ? Image: Constraint of the second	Set TLS Restraints Xtriage Ask for help
Configure	4 ▷ ×
Input data Refinement settings Output Amber	4 له
Job title : Ref1	
Input files	
File path Format	Data type
	Input model
Q /mnt/home1/iimathew/SUMMER_STUDENTS/ALI_FOX_2020/DEPOS cif	X-ray data, X-ray R-free
\Q_/mnt/home1/iimathew/SUMMER_STUDENTS/ALI_FOX_2020/DEPOS Sequence	Sequences
Add file Remove file Modify file data type Use symmetry from selected file	
Space group : P 43 21 2 Vnit cell : 78.483 78.483 37.104 90 90 90	
X-ray data and experimental phases	
Data labels : rxxxxsf,_refin.F_me R-free label : rxxxxsf,_refin.stat	us
High resolution : (1.450) Low resolution : (39.2	242) Phase labels : 🔻
Wavelength : Options	
Neutron data	
Data labels : 💌 R-free label :	▼ Test flag value :
High resolution : Low resolution :	Options
• Idle	Project: LYSO123

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.

phenix.refine (Project: LYSO123)	×
<u>F</u> ile <u>A</u> ctions <u>S</u> ettings <u>U</u> tilities <u>H</u> elp	
X Preferences Help Image: Save Image: Save	
Configure	4 Þ ×
Input data Refinement settings Output Amber	4 ⊳
Strategy	
Refinement strategy:	
Number of cycles : 3	
Select Atoms Note: selections can only be made for enabled options (e.g. NCS groups are available if "Use NCS" box is checked)	
Targets and weighting	
Target function : Automatic Optimize X-ray/stereochemistry weight Optimize X-ray/ADP weight 	
Use NCS NCS type: torsion-angle	
Reference model restraints Use secondary structure restraints Use experimental phase restraints	
Refinement target weights Model interpretation NCS options ?	
Other options	
Automatically add hydrogens to model Update waters Place elemental ions :	
□ Simulated annealing (Cartesian) □ Simulated annealing (Torsion angles) Scattering table : n_gaussian ▼	
Automatically correct N/Q/H errors Twin law : Number of processors : 1 7	
Global refinement parameters Modify start model All parameters ?	
Idle Project: LYS0123	

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

• •	phenix.refine (Project: LYSO123) _ D	×
<u>File Actions Settings Utilities Help</u>		
X ? Image: Constraint of the second	Save Graphics ReadySet TLS Restraints Xtriage Ask for help	
Configure	1 Þ	×
Input data Refinement settings Output Amb	ber	⊳
Output files		
File prefix : LYS0123_refine Output dire	rectory : //mnt/home1/iimathew/SUMMER_STUDENTS/ALI_FOX_2020/I Browse	
Other options		
Send e-mail notification when refinement is done	E-mail address to use :	
✓ Write final refined model into mmCIF file	□ Write Fobs, Fmodel and R-free-flags into mmClF file	
Skip real-space correlation step in validation	Skip kinemage generation in validation	
Maps		
Phenix will always output 2mFo-Di anomalous difference map if anom in the final MTZ file written at the or request additional map coefficie end of refinement, or using the seg generate maps without running ref	DFc and mFo-DFc map coefficients suitable for viewing in Coot by default, plus an malous data ($F+$ and $F-$ or $I+$ and $I-$) were supplied as input. These will be included end of refinement. You may customize the options used in generating these maps ients if you wish. These may be automatically converted to CCP4 format at the parate FFT program in the Maps section of the main window. (Note that you can effinement using the program phenix.maps.)	
🌮 Map coefficients 🏼 🎯 Omit a	atoms 🖹 Run phenix.maps	
● Idle	Project: LYSO123	///

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.

○ ↑				pheni	x.refine (Project: I	YSO123)				_ = ×
<u>File</u> <u>Actions</u>	<u>S</u> ettings	<u>U</u> tilities <u>H</u> el	р							
Preferences	? Help	Run	E Abort	Save Graph	ics ReadySet	k TLS	Restraints	Xtriage	Ask for help	
Configure	Refine_2									4 Þ×
Log output	Run statu	IS Results	MolProbity	Real-space correlation	Atomic proper	ties Sec	quence check			4 ⊳
Output fil Title: Directory:	es Ref1 /mnt/hor	nel/iimathew,	/SUMMER_STUD	ENTS/ALI_F0X_2020/DE	POSITED_7KNK/PHE	NIX/Refin	e_2			<u> </u>
File name				Contents				<u>^</u>	🥑 Open i	1 Coot
Q LYSO123	refine_2.lo	g		phenix.refine log file					R open n	reoor
् LYSO123	refine_data	a.mtz		Data and R-free set us	sed in refine				🚧 Open ji	
્ LYSO123_	refine_2.ge	20		Geometry restraints b	efore refine					,
્ LYSO123_	refine_2.ef	f		Effective parameters	for this run					ce viewer
् LYSO123_	refine_2.pd	lb		Refined model				-	Jequer	
Convert map	o coefficien ent statis	ts MTZ to CCP	4 maps							
Com	pare st	atistics	🔒 Plot	statistics by cy	vcle 🔒 Pio	ot stati	stics by re	solution		
Before and R-work R-free Bonds Angles X-ray statis	after refi Starting 0.3103 0.3220 0.009 1.385 tics by re	nement: Final 0.2625 0.2997 0.006 0.955 solution bin	:							
Idle							Proje	ect: 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).