MX Tutorial 3

Refinement and Validation of the crystal structure of Hen Egg White Lysozyme

Pedro Matias, Célia Romão, José Brito

In this tutorial, the crystal structure of Hen Egg White Lysozyme will be refined using REFMAC and PHENIX, and the resulting structure analyzed using the validation tools in PHENIX.

ECS7 - 7th European Crystallographic School – July 13, 2022

Getting started

- Login to your account and start a shell terminal;
- Type the command:

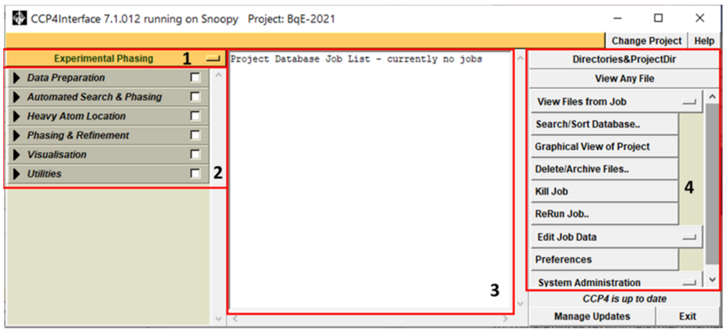
```
%cd ECS7/Lysozyme ↵
```

- Start CCP4i

```
% ccp4i & ↵
```

Overview of the CCP4i interface - recap

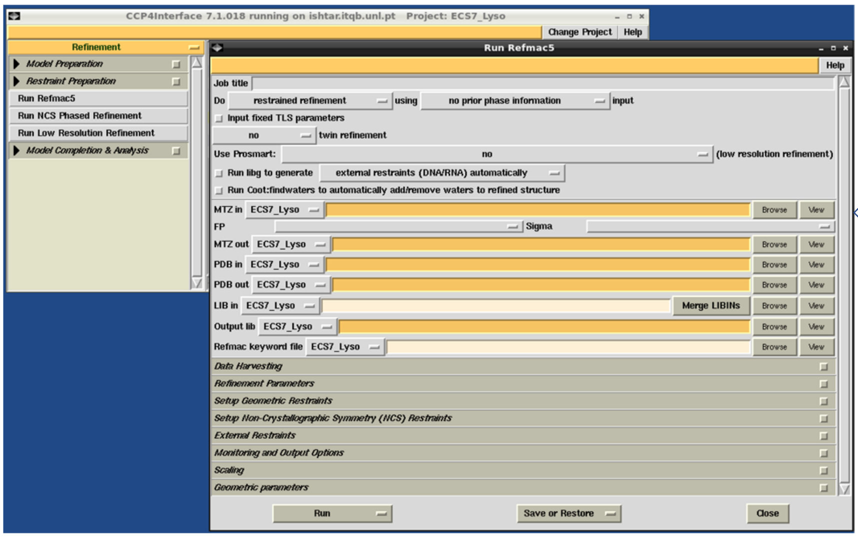
This is what the main window of the CCP4 GUI looks like:



1. *This button selects groups of programs;*
2. *Buttons in this window open a program list or select a program in the group;*
3. *This window lists all the run and running jobs in the current project;*
4. *This window contains buttons for acting on jobs and interface management;*

Start the REFMAC5 job window

- Select the Refinement program group and click on **Run Refmac5**



Job title

Do restrained refinement using no prior phase information input

☐ Input fixed TLS parameters

no twin refinement

Use Prosmart: no (low resolution refinement)

☐ Run libg to generate external restraints (DNA/RNA) automatically

☐ Run Coot:findwaters to automatically add/remove waters to refined structure

MTZ in ECS7_Lyso Browse View

FP Sigma

MTZ out ECS7_Lyso Browse View

PDB in ECS7_Lyso Browse View

PDB out ECS7_Lyso Browse View

LIB in ECS7_Lyso Merge LIBINs Browse View

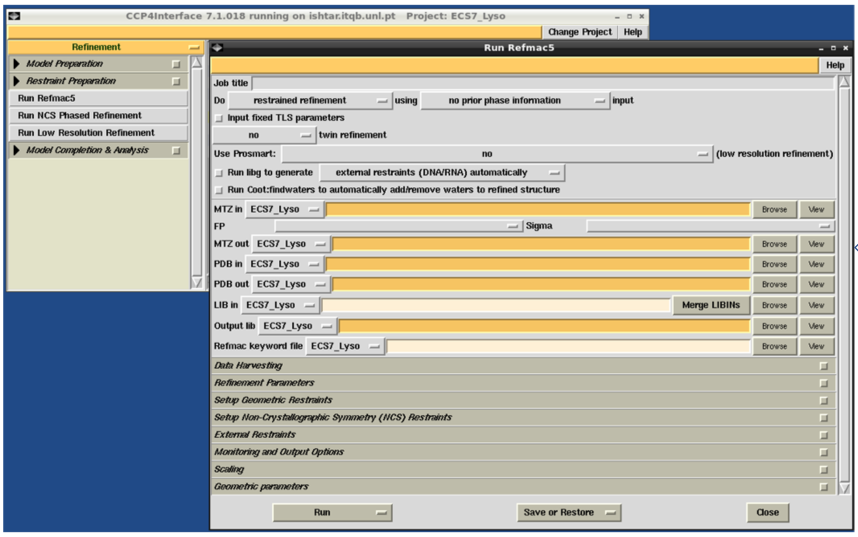
Output lib ECS7_Lyso Browse View

Refmac keyword file ECS7_Lyso Browse View

input reflection file

The REFMAC5 job window – an overview

- Select the Refinement program group and click on **Run Refmac5**

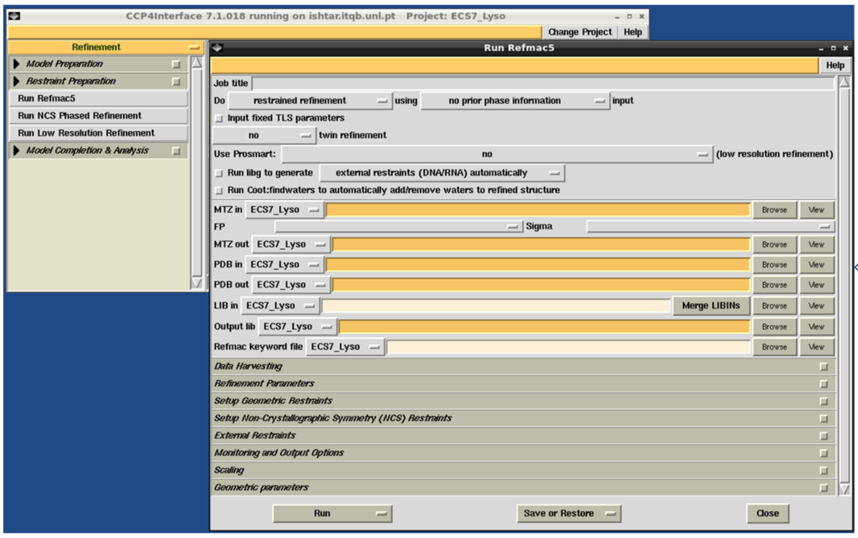


Job title
Do restrained refinement using no prior phase information input
☐ Input fixed TLS parameters
no twin refinement
Use Prosmart: no (low resolution refinement)
☐ Run libg to generate external restraints (DNA/RNA) automatically
☐ Run Coot:findwaters to automatically add/remove waters to refined structure
MTZ in ECS7_Lyso
FP Sigma
MTZ out ECS7_Lyso
PDB in ECS7_Lyso
PDB out ECS7_Lyso
LIB in ECS7_Lyso
Output lib ECS7_Lyso
Refmac keyword file ECS7_Lyso

output reflection file

The REFMAC5 job window – an overview

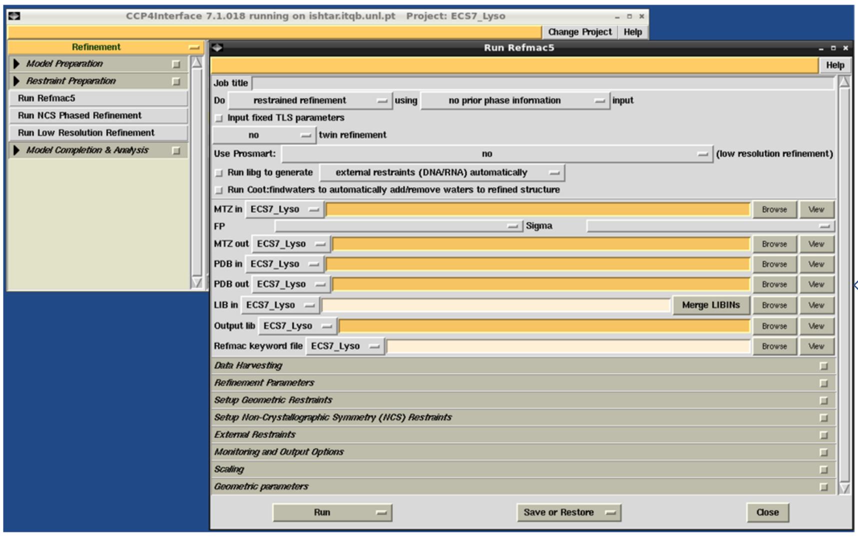
- Select the Refinement program group and click on **Run Refmac5**


Job title
Do restrained refinement using no prior phase information input
☐ Input fixed TLS parameters
no twin refinement
Use Prosmart: no (low resolution refinement)
☐ Run libg to generate external restraints (DNA/RNA) automatically
☐ Run Coot:findwaters to automatically add/remove waters to refined structure
MTZ in ECS7_Lyso
FP Sigma
MTZ out ECS7_Lyso
PDB in ECS7_Lyso
PDB out ECS7_Lyso
LIB in ECS7_Lyso
Output lib ECS7_Lyso
Refmac keyword file ECS7_Lyso

input coordinates file

The REFMAC5 job window – an overview

- Select the Refinement program group and click on **Run Refmac5**



Job title

Do restrained refinement using no prior phase information input

☐ Input fixed TLS parameters

no twin refinement

Use Prosmart: no (low resolution refinement)

☐ Run libg to generate external restraints (DNA/RNA) automatically

☐ Run Coot:findwaters to automatically add/remove waters to refined structure

MTZ in ECS7_Lyso

FP Sigma

MTZ out ECS7_Lyso

PDB in ECS7_Lyso

PDB out ECS7_Lyso

LIB in ECS7_Lyso

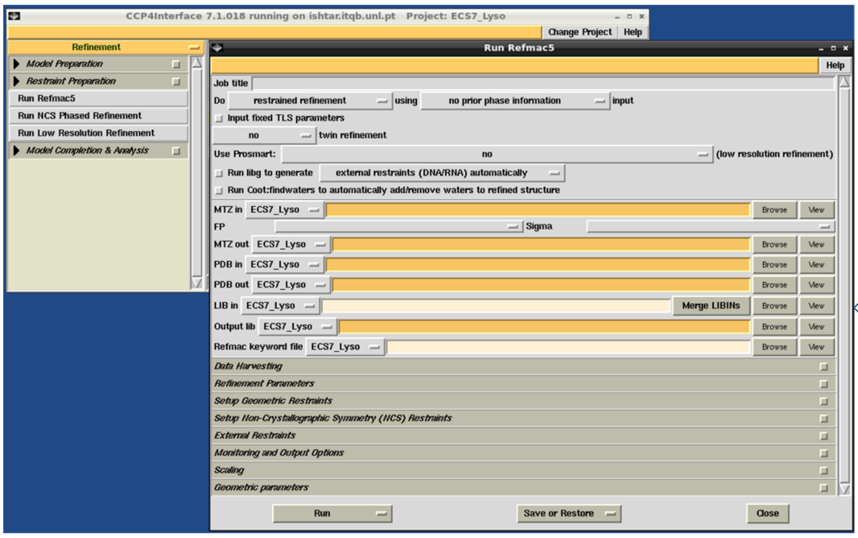
Output lib ECS7_Lyso

Refmac keyword file ECS7_Lyso

output coordinates file

The REFMAC5 job window – an overview

- Select the Refinement program group and click on **Run Refmac5**



Job title

Do restrained refinement using no prior phase information input

☐ Input fixed TLS parameters

no twin refinement

Use Prosmart: no (low resolution refinement)

☐ Run libg to generate external restraints (DNA/RNA) automatically

☐ Run Coot:findwaters to automatically add/remove waters to refined structure

MTZ in ECS7_Lyso

FP Sigma

MTZ out ECS7_Lyso

PDB in ECS7_Lyso

PDB out ECS7_Lyso

LIB in ECS7_Lyso

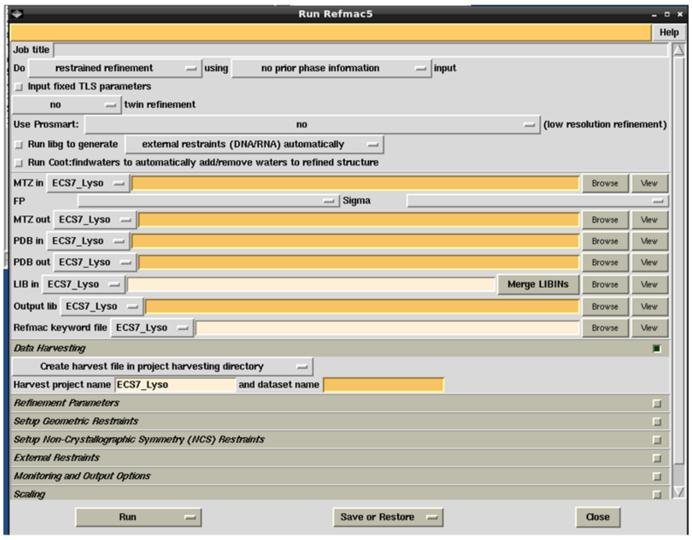
Output lib ECS7_Lyso

Refmac keyword file ECS7_Lyso

Input dictionary file

The REFMAC5 job window – an overview

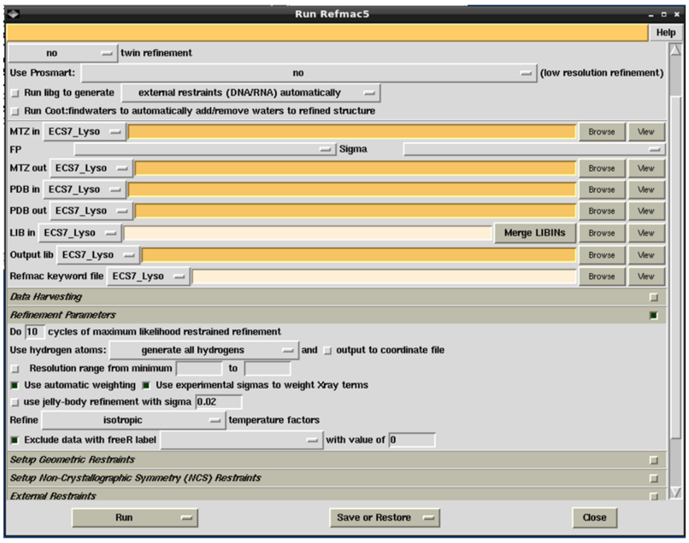
- Data harvesting option (I turn it off)



Job title
Do restrained refinement using no prior phase information input
☐ Input fixed TLS parameters
no twin refinement
Use Prosmart: no (low resolution refinement)
☐ Run libg to generate external restraints (DNA/RNA) automatically
☐ Run Coot:findwaters to automatically add/remove waters to refined structure
MTZ in ECS7_Lyso
FP Sigma
MTZ out ECS7_Lyso
PDB in ECS7_Lyso
PDB out ECS7_Lyso
LIB in ECS7_Lyso
Output lib ECS7_Lyso
Refmac keyword file ECS7_Lyso
Data Harvesting ☑
Create harvest file in project harvesting directory
Harvest project name ECS7_Lyso and dataset name
Refinement Parameters ☐
Setup Geometric Restraints ☐
Setup Non-Crystallographic Symmetry (NCS) Restraints ☐
External Restraints ☐
Monitoring and Output Options ☐
Scaling ☐

The REFMAC5 job window – an overview

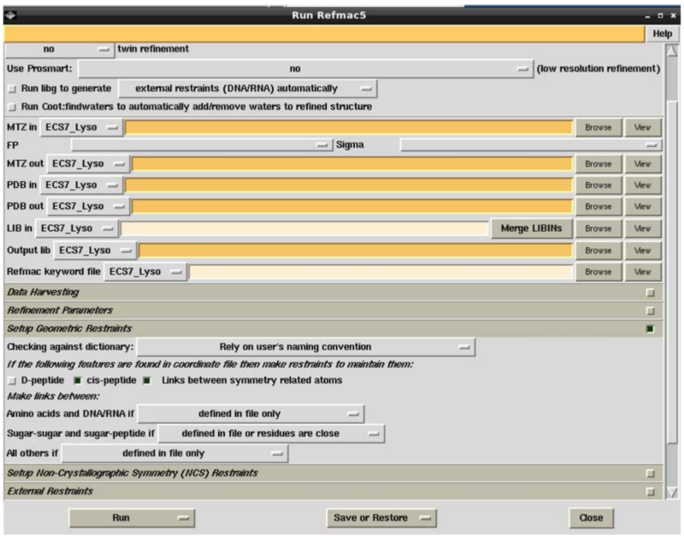
- Refinement parameters (the defaults are usually OK)



no twin refinement
Use Prosmart: no (low resolution refinement)
☐ Run libg to generate external restraints (DNA/RNA) automatically
☐ Run Coot:findwaters to automatically add/remove waters to refined structure
MTZ in ECS7_Lyso
FP Sigma
MTZ out ECS7_Lyso
PDB in ECS7_Lyso
PDB out ECS7_Lyso
LIB in ECS7_Lyso
Output lib ECS7_Lyso
Refmac keyword file ECS7_Lyso
Data Harvesting ☐
Refinement Parameters ☑
Do 10 cycles of maximum likelihood restrained refinement
Use hydrogen atoms: generate all hydrogens and ☐ output to coordinate file
☐ Resolution range from minimum to
☑ Use automatic weighting ☑ Use experimental sigmas to weight Xray terms
☐ use jelly-body refinement with sigma 0.02
Refine isotropic temperature factors
☑ Exclude data with freeR label with value of 0
Setup Geometric Restraints ☐
Setup Non-Crystallographic Symmetry (NCS) Restraints ☐
External Restraints ☐

The REFMAC5 job window – an overview

- Setup geometric restraints (the defaults are usually OK)



no twin refinement

Use Prosmart: no (low resolution refinement)

☐ Run libg to generate external restraints (DNA/RNA) automatically

☐ Run Coot:findwaters to automatically add/remove waters to refined structure

MTZ in ECS7_Lyso

FP Sigma

MTZ out ECS7_Lyso

PDB in ECS7_Lyso

PDB out ECS7_Lyso

LIB in ECS7_Lyso

Output lib ECS7_Lyso

Refmac keyword file ECS7_Lyso

Data Harvesting ☐

Refinement Parameters ☐

Setup Geometric Restraints ☒

Checking against dictionary: Rely on user's naming convention

If the following features are found in coordinate file then make restraints to maintain them:

☐ D-peptide ☒ cis-peptide ☒ Links between symmetry related atoms

Make links between:

Amino acids and DNA/RNA if defined in file only

Sugar-sugar and sugar-peptide if defined in file or residues are close

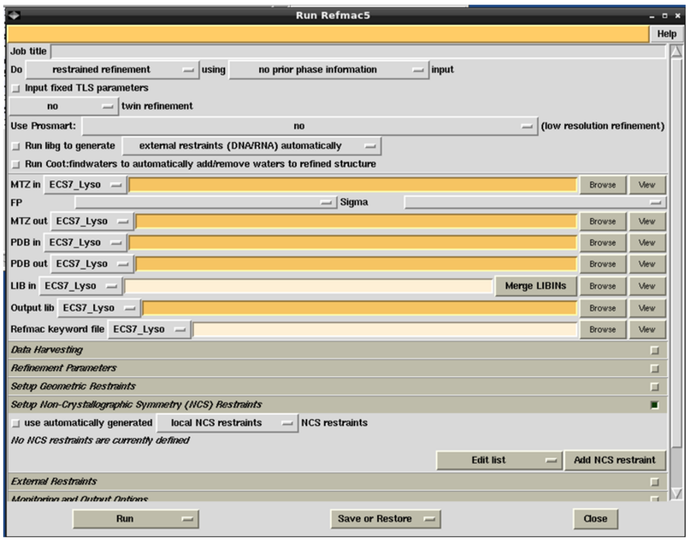
All others if defined in file only

Setup Non-Crystallographic Symmetry (NCS) Restraints ☐

External Restraints ☐

The REFMAC5 job window – an overview

- NCS restraints (**should be used if more than one copy in the asymmetric unit**.)



Job title

Do restrained refinement using no prior phase information input

☐ Input fixed TLS parameters

no twin refinement

Use Prosmart: no (low resolution refinement)

☐ Run libg to generate external restraints (DNA/RNA) automatically

☐ Run Coot:findwaters to automatically add/remove waters to refined structure

MTZ in ECS7_Lyso

FP Sigma

MTZ out ECS7_Lyso

PDB in ECS7_Lyso

PDB out ECS7_Lyso

LIB in ECS7_Lyso

Output lib ECS7_Lyso

Refmac keyword file ECS7_Lyso

Data Harvesting ☐

Refinement Parameters ☐

Setup Geometric Restraints ☐

Setup Non-Crystallographic Symmetry (NCS) Restraints ☒

☐ use automatically generated local NCS restraints NCS restraints

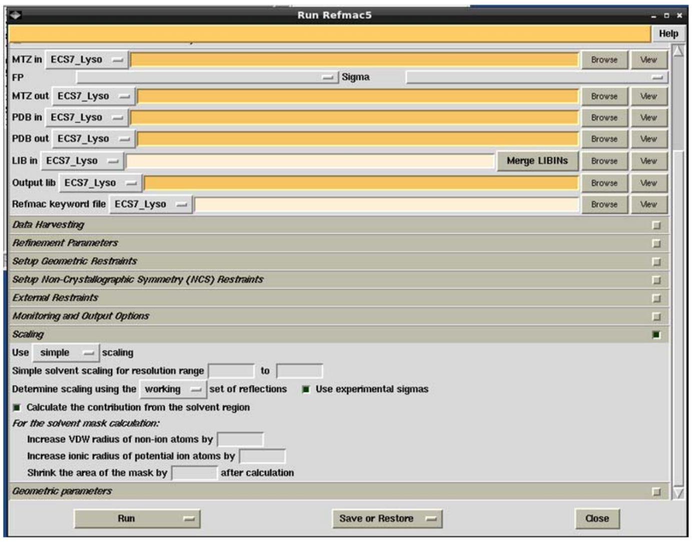
No NCS restraints are currently defined

External Restraints ☐

Monitoring and Output Options ☐

The REFMAC5 job window – an overview

- Scaling between $|F_{obs}|$ and $|F_{calc}|$ (the default settings are usually OK)



Run Refmac5

MTZ in ECS7_Lyso

FP

Sigma

MTZ out ECS7_Lyso

PDB in ECS7_Lyso

PDB out ECS7_Lyso

LIB in ECS7_Lyso

Output lib ECS7_Lyso

Refmac keyword file ECS7_Lyso

Scaling ☒

Use simple scaling

Simple solvent scaling for resolution range ☐ to ☐

Determine scaling using the working set of reflections ☒ Use experimental sigmas

☒ Calculate the contribution from the solvent region

For the solvent mask calculation:

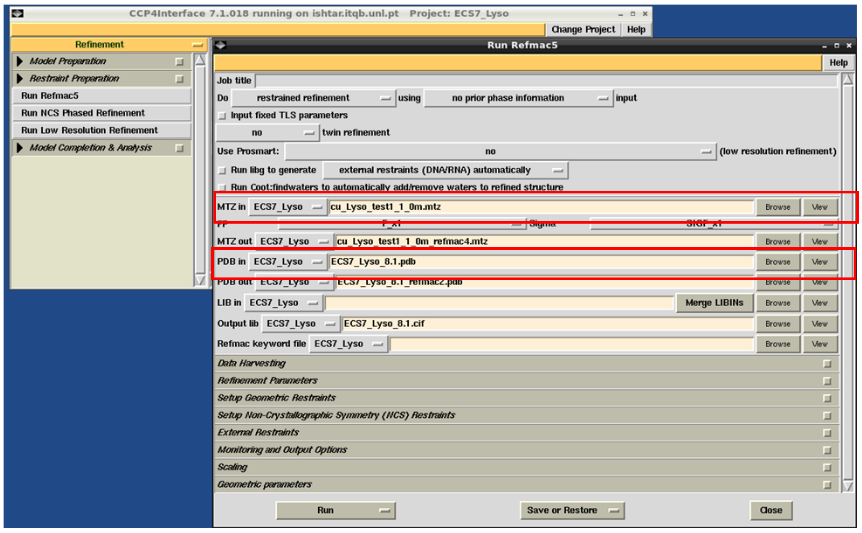
Increase VDW radius of non-ion atoms by ☐

Increase ionic radius of potential ion atoms by ☐

Shrink the area of the mask by ☐ after calculation

Load the refinement files

- Click on the **Browse** button to select the input **reflection** and **coordinate** files to be used in the refinement
- Notice that the output filenames are automatically generated by the program



CCP4Interface 7.1.018 running on ishtar.itqb.unl.pt Project: ECS7_Lyso

Run Refmac5

Job title

Do restrained refinement using no prior phase information input

☐ Input fixed TLS parameters

no twin refinement

Use Prosmart: no (low resolution refinement)

☐ Run libg to generate external restraints (DNA/RNA) automatically

☐ Run Coot:findwaters to automatically add/remove waters to refined structure

MTZ in ECS7_Lyso cu_Lyso_test1_1_0m.mtz

FP F_x1 Sigma SIGF_x1

MTZ out ECS7_Lyso cu_Lyso_test1_1_0m_refmac4.mtz

PDB in ECS7_Lyso ECS7_Lyso_8.1.pdb

PDB out ECS7_Lyso ECS7_Lyso_8.1_refmac2.pdb

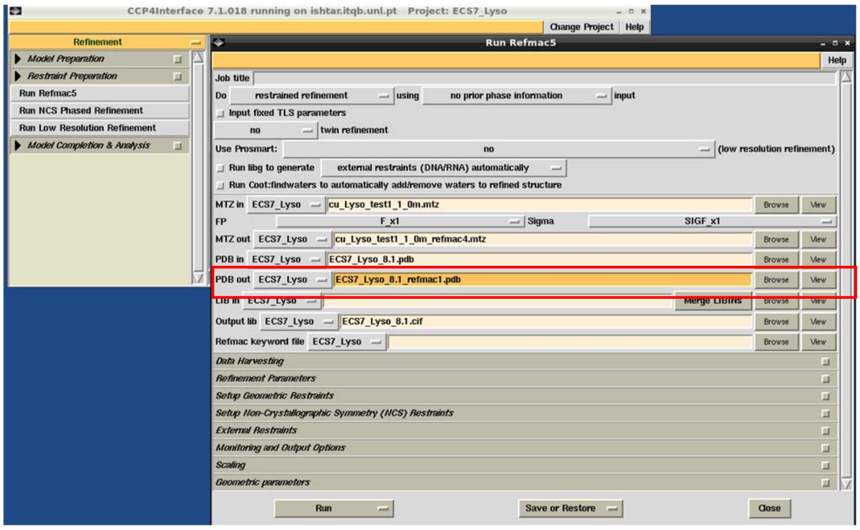
LIB in ECS7_Lyso

Output lib ECS7_Lyso ECS7_Lyso_8.1.cif

Refmac keyword file ECS7_Lyso

Load the refinement files

- A **dark orange background** indicates a **missing mandatory file** or a **file that already exists** and will be overwritten if the program is executed.



Job title

Do restrained refinement using no prior phase information input

☐ Input fixed TLS parameters

no twin refinement

Use Prosmart: no (low resolution refinement)

☐ Run libg to generate external restraints (DNA/RNA) automatically

☐ Run Coot:findwaters to automatically add/remove waters to refined structure

MTZ in ECS7_Lyso cu_Lyso_test1_1_0m.mtz

FP F_x1 Sigma SIGF_x1

MTZ out ECS7_Lyso cu_Lyso_test1_1_0m_refmac4.mtz

PDB in ECS7_Lyso ECS7_Lyso_8.1.pdb

PDB out ECS7_Lyso ECS7_Lyso_8.1_refmac1.pdb

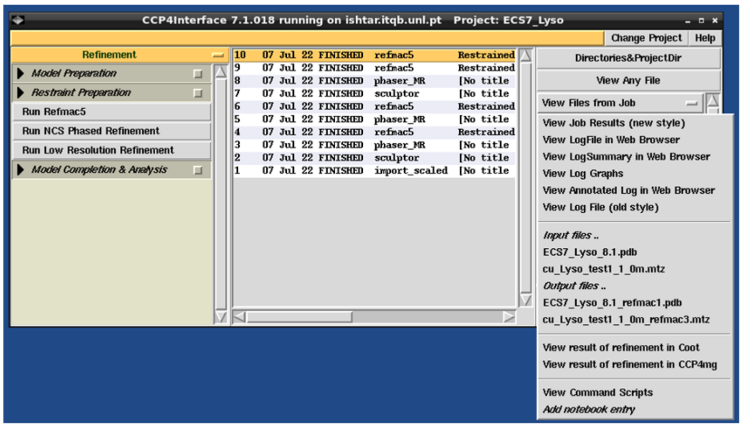
LIB in ECS7_Lyso

Output lib ECS7_Lyso ECS7_Lyso_8.1.cif

Refmac keyword file ECS7_Lyso

Running the REFMAC5 job and looking at the results

- Click on the **Run** button, select **Run Now** from the drop-down menu and then click the **Close** button.
- When the program is **FINISHED** in the central GUI window, click on the **View Files from Job** button and inspect the results.



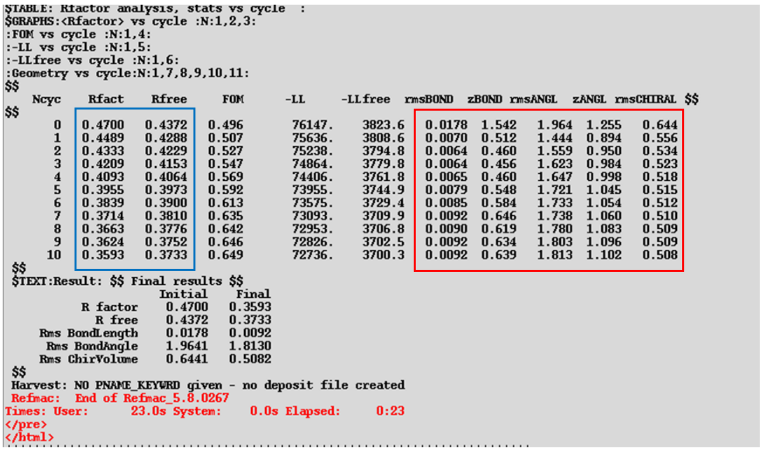
```
10   07 Jul 22 FINISHED  refmac5        Restrained
9    07 Jul 22 FINISHED  refmac5        Restrained
8    07 Jul 22 FINISHED  phaser_MR      [No title
7    07 Jul 22 FINISHED  sculptor       [No title
6    07 Jul 22 FINISHED  refmac5        Restrained
5    07 Jul 22 FINISHED  phaser_MR      [No title
4    07 Jul 22 FINISHED  refmac5        Restrained
3    07 Jul 22 FINISHED  phaser_MR      [No title
2    07 Jul 22 FINISHED  sculptor       [No title
1    07 Jul 22 FINISHED  import_scaled  [No title
```

Log files and graphs

Open a Coot window with the coordinate and maps from the refinement

Running the REFMAC5 job and looking at the results

- Click on the **View Log file (old style)** button and scroll down to the bottom.



```
$TABLE: Rfactor analysis, stats vs cycle  :
$GRAPHS:<Rfactor> vs cycle :N:1,2,3:
:FOM vs cycle :N:1,4:
:-LL vs cycle :N:1,5:
:-LLfree vs cycle :N:1,6:
:Geometry vs cycle:N:1,7,8,9,10,11:
$$
    Ncyc    Rfact    Rfree     FOM      -LL     -LLfree  rmsBOND  zBOND rmsANGL  zANGL  rmsCHIRAL $$
$$
       0  0.4700   0.4372   0.496     76147.    3823.6   0.0178  1.542   1.964  1.255    0.644
       1  0.4489   0.4288   0.507     75636.    3808.6   0.0070  0.512   1.444  0.894    0.556
       2  0.4333   0.4229   0.527     75238.    3794.8   0.0064  0.460   1.559  0.950    0.534
       3  0.4209   0.4153   0.547     74864.    3779.8   0.0064  0.456   1.623  0.984    0.523
       4  0.4093   0.4064   0.569     74406.    3761.8   0.0065  0.460   1.647  0.998    0.518
       5  0.3955   0.3973   0.592     73955.    3744.9   0.0079  0.548   1.721  1.045    0.515
       6  0.3839   0.3900   0.613     73575.    3729.4   0.0085  0.584   1.733  1.054    0.512
       7  0.3714   0.3810   0.635     73093.    3709.9   0.0092  0.646   1.738  1.060    0.510
       8  0.3663   0.3776   0.642     72953.    3706.8   0.0090  0.619   1.780  1.083    0.509
       9  0.3624   0.3752   0.646     72826.    3702.5   0.0092  0.634   1.803  1.096    0.509
      10  0.3593   0.3733   0.649     72736.    3700.3   0.0092  0.639   1.813  1.102    0.508
 $$
 $TEXT:Result: $$ Final results $$
                      Initial    Final
           R factor    0.4700   0.3593
             R free    0.4372   0.3733
     Rms BondLength    0.0178   0.0092
      Rms BondAngle    1.9641   1.8130
     Rms ChirVolume    0.6441   0.5082
 $$
 Harvest: NO PNAME_KEYWRD given - no deposit file created
 Refmac:  End of Refmac_5.8.0267
Times: User:      23.0s System:    0.0s Elapsed:     0:23
</pre>
</html>
```

Check the variation of R and Rfree

Check the geometry statistics

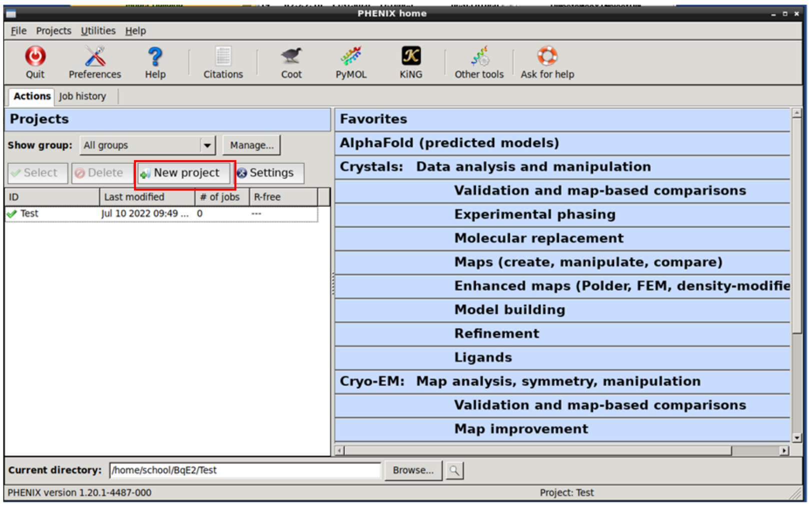
Time for some Cooting…

Refining the structure with PHENIX – creating a new project

Phenix.refine runs much slower but is much more flexible than **REFMAC** for crystal structure refinement.

- Exit ccp4i or start a new shell terminal
- Start the PHENIX GUI

  `% phenix & ↵`



Show group: All groups

ID	Last modified	# of jobs	R-free
Test	Jul 10 2022 09:49 ...	0	---

Current directory: /home/school/BqE2/Test

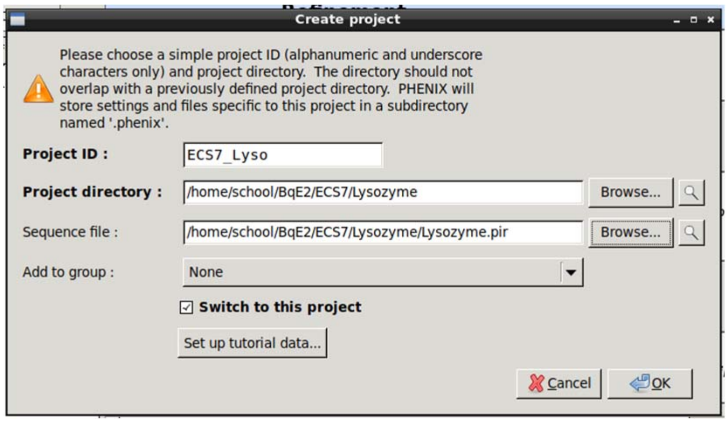
PHENIX version 1.20.1-4487-000

Project: Test

- You will probably have already a project named **Test** on your computer, but you need to create a new one for the present tutorial in the same directory where the tutorial files are located
- Click on the **New project** button

Refining the structure with PHENIX – creating a new project

- Fill out the information as shown and click on the OK button



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 : ECS7_Lyso

Project directory : /home/school/BqE2/ECS7/Lysozyme

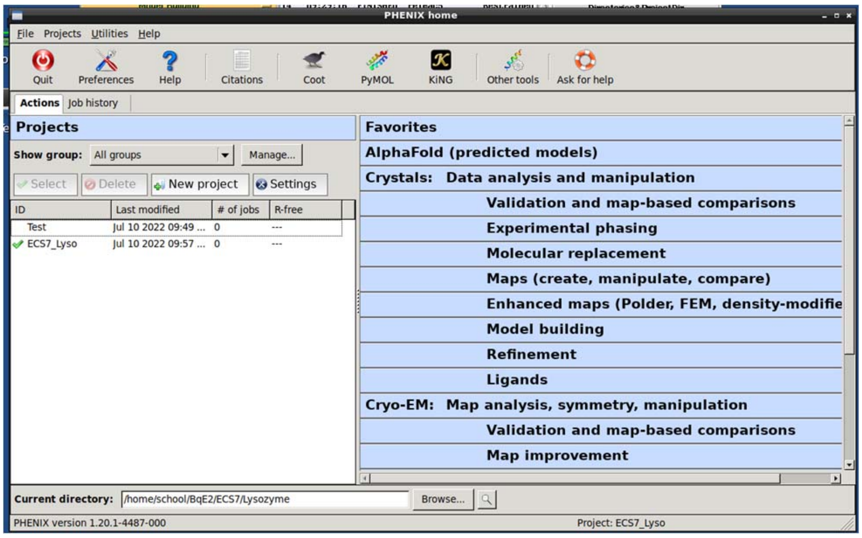
Sequence file : /home/school/BqE2/ECS7/Lysozyme/Lysozyme.pir

Add to group : None

☑ **Switch to this project**

Refining the structure with PHENIX – creating a new project

- You now have created your PHENIX project for this tutorial session



ID	Last modified	# of jobs	R-free
Test	Jul 10 2022 09:49 ...	0	---
✓ ECS7_Lyso	Jul 10 2022 09:57 ...	0	---

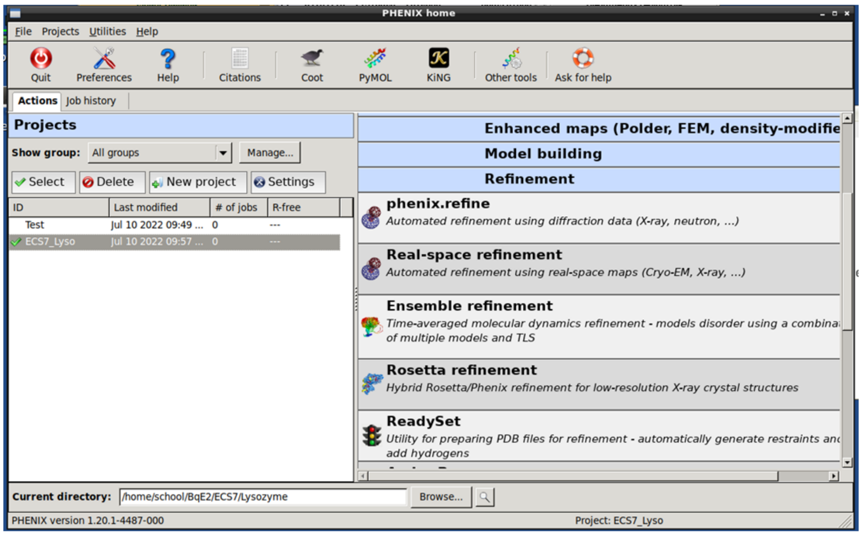
Current directory: /home/school/BqE2/ECS7/Lysozyme

PHENIX version 1.20.1-4487-000

Project: ECS7_Lyso

Refining the structure with PHENIX – adding hydrogen atoms

- Click on the **Refinement bar** to open it and click on the **ReadySet** button to open its GUI



ID	Last modified	# of jobs	R-free
Test	Jul 10 2022 09:49 ...	0	---
ECS7_Lyso	Jul 10 2022 09:57 ...	0	---

phenix.refine
Automated refinement using diffraction data (X-ray, neutron, ...)

Real-space refinement
Automated refinement using real-space maps (Cryo-EM, X-ray, ...)

Ensemble refinement
Time-averaged molecular dynamics refinement - models disorder using a combina of multiple models and TLS

Rosetta refinement
Hybrid Rosetta/Phenix refinement for low-resolution X-ray crystal structures

ReadySet
Utility for preparing PDB files for refinement - automatically generate restraints an add hydrogens

Current directory: /home/school/BqE2/ECS7/Lysozyme

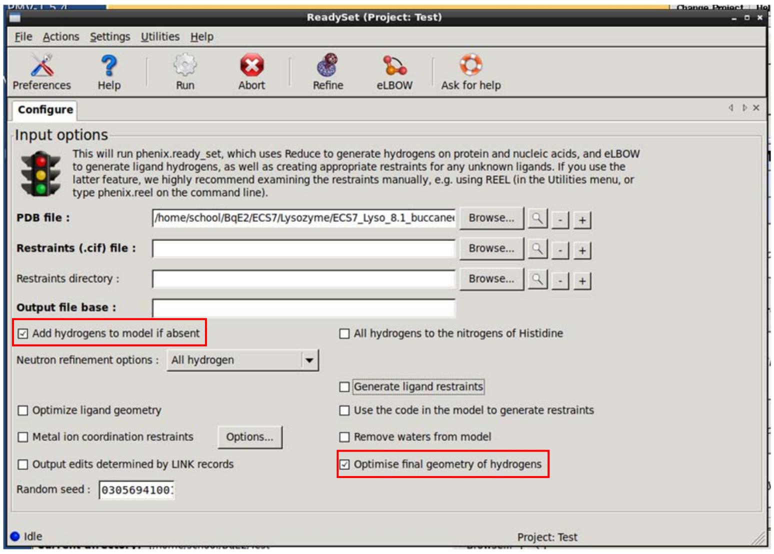
PHENIX version 1.20.1-4487-000

Project: ECS7_Lyso

Refining the structure with PHENIX – adding hydrogen atoms

- Click on the **Refinement bar** to open it and click on the **ReadySet** button to open its GUI (it takes a few seconds!)

- Load the PDB file from REFMAC and COOT in the PDB file window
- Make sure only the indicated check boxes are ticked
- Click on the Run button.



Input options

This will run phenix.ready_set, which uses Reduce to generate hydrogens on protein and nucleic acids, and eLBOW to generate ligand hydrogens, as well as creating appropriate restraints for any unknown ligands. If you use the latter feature, we highly recommend examining the restraints manually, e.g. using REEL (in the Utilities menu, or type phenix.reel on the command line).

PDB file : /home/school/BqE2/ECS7/Lysozyme/ECS7_Lyso_8.1_buccane

Restraints (.cif) file :

Restraints directory :

Output file base :

☑ Add hydrogens to model if absent
☐ All hydrogens to the nitrogens of Histidine

Neutron refinement options : All hydrogen

☐ Generate ligand restraints
☐ Optimize ligand geometry
☐ Use the code in the model to generate restraints
☐ Metal ion coordination restraints
☐ Remove waters from model
☐ Output edits determined by LINK records
☑ Optimise final geometry of hydrogens

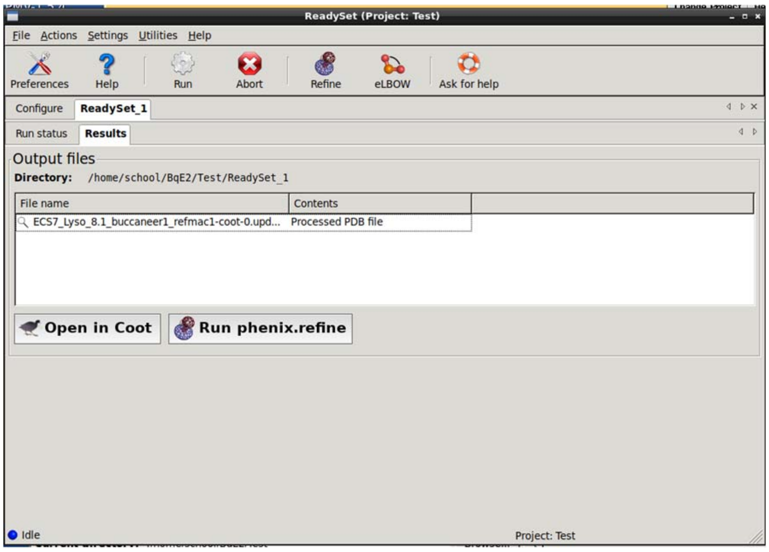
Random seed : 0305694100

Idle

Project: Test

Refining the structure with PHENIX – adding hydrogen atoms

- After a few seconds the job will be finished, and you can view the file in COOT

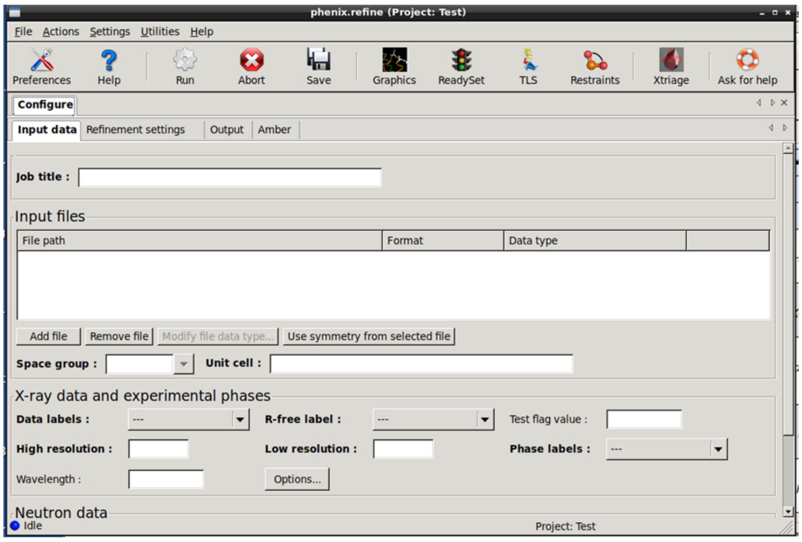


Directory: /home/school/BqE2/Test/ReadySet_1

File name	Contents
ECS7_Lyso_8.1_buccaneer1_refmac1-coot-0.upd...	Processed PDB file

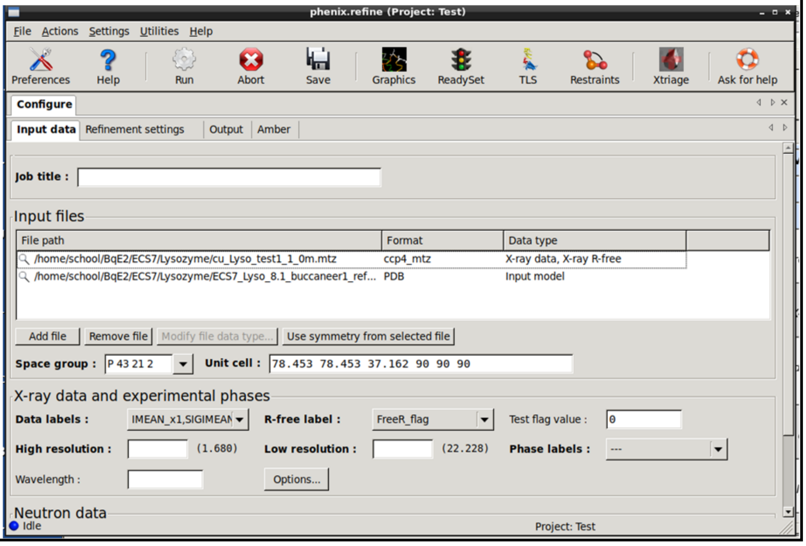
Refining the structure with PHENIX – running phenix.refine

- Now click on the **File** menu and select **close** (you can also click on the **x**) and click on the **phenix.refine** button



Refining the structure with PHENIX – running phenix.refine

- We need to load a **reflection** and a **coordinate** file using the **add file** button
- PHENIX can **identify** the file type **automatically** and complains if it cannot recognize the input file.
  - Notice that the reflection and PDB files have been correctly identified and that the crystal information has been read in.
  - PHENIX uses the **diffraction intensities** in the refinement by default.



Job title :

Input files

File path	Format	Data type
/home/school/BqE2/ECS7/Lysozyme/cu_Lyso_test1_1_0m.mtz	ccp4_mtz	X-ray data, X-ray R-free
/home/school/BqE2/ECS7/Lysozyme/ECS7_Lyso_8.1_buccaneer1_ref...	PDB	Input model

Space group : P 43 21 2 **Unit cell :** 78.453 78.453 37.162 90 90 90

X-ray data and experimental phases

Data labels : IMEAN_x1,SIGIMEAN **R-free label :** FreeR_flag Test flag value : 0

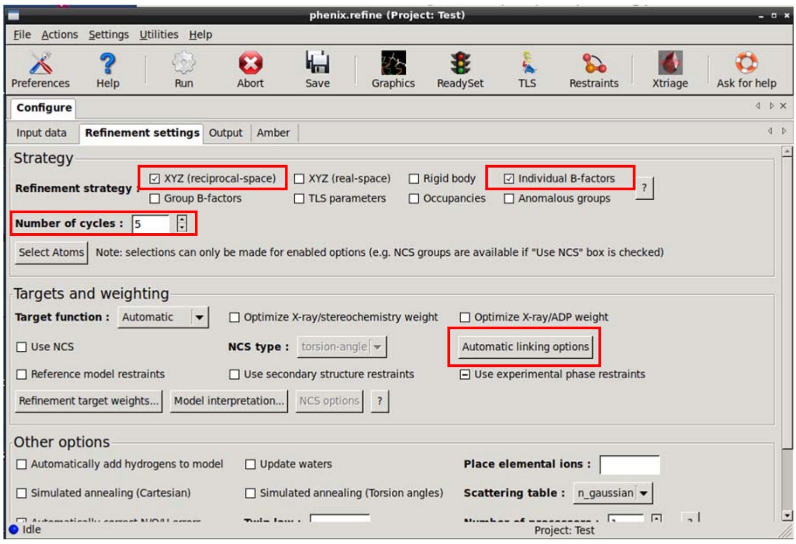
High resolution : (1.680) **Low resolution :** (22.228) **Phase labels :** ---

Wavelength :

Project: Test

Refining the structure with PHENIX – running phenix.refine

- Click on the **Refinement Settings** tab, make sure that only the **XYZ (reciprocal space)** and **Individual B-factors** boxes are checked and change the **Number of cycles** to **5**.
- Next, click on the **Automatic linking options** button.



Strategy

Refinement strategy : ☑ XYZ (reciprocal-space) ☐ XYZ (real-space) ☐ Rigid body ☑ Individual B-factors ☐ Group B-factors ☐ TLS parameters ☐ Occupancies ☐ Anomalous groups

Number of cycles : 5

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

Other options

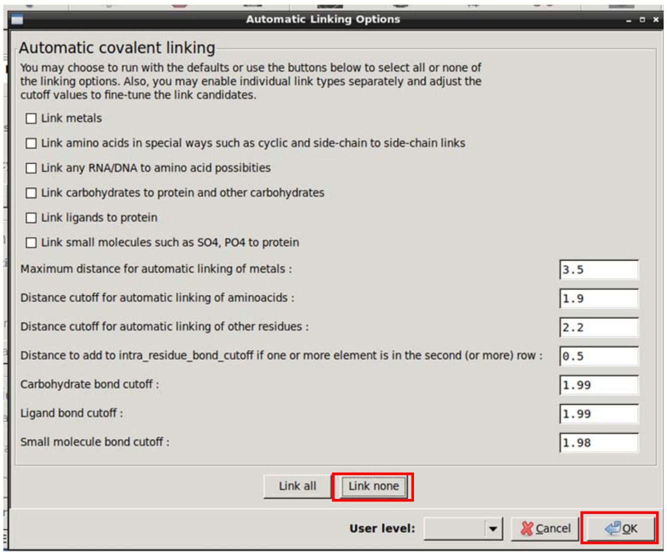
☐ Automatically add hydrogens to model ☐ Update waters **Place elemental ions :**

☐ Simulated annealing (Cartesian) ☐ Simulated annealing (Torsion angles) **Scattering table :** n_gaussian

Project: Test

Refining the structure with PHENIX – running phenix.refine

- Click on the **Link none** then on the **OK** buttons to exit this menu.



Automatic covalent linking

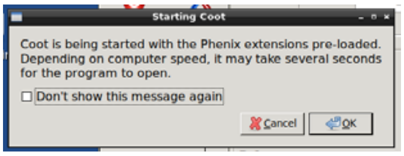
You may choose to run with the defaults or use the buttons below to select all or none of the linking options. Also, you may enable individual link types separately and adjust the cutoff values to fine-tune the link candidates.

☐ Link metals
☐ Link amino acids in special ways such as cyclic and side-chain to side-chain links
☐ Link any RNA/DNA to amino acid possibilities
☐ Link carbohydrates to protein and other carbohydrates
☐ Link ligands to protein
☐ Link small molecules such as SO4, PO4 to protein

Maximum distance for automatic linking of metals : 3.5
Distance cutoff for automatic linking of aminoacids : 1.9
Distance cutoff for automatic linking of other residues : 2.2
Distance to add to intra_residue_bond_cutoff if one or more element is in the second (or more) row : 0.5
Carbohydrate bond cutoff : 1.99
Ligand bond cutoff : 1.99
Small molecule bond cutoff : 1.98

Refining the structure with PHENIX – running phenix.refine

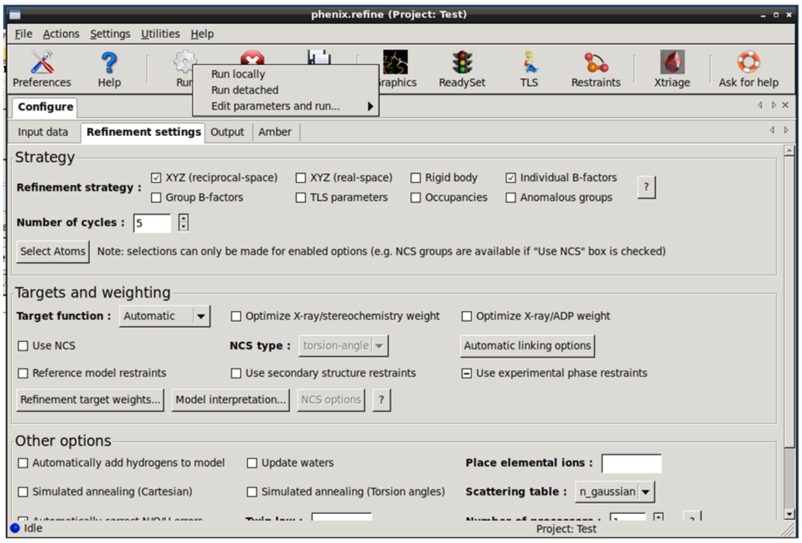
- Back on the main menu, click on **Run** and then on **Run locally** from the drop-down menu.
- A Coot window will open, and the refinement will start



Coot is being started with the Phenix extensions pre-loaded. Depending on computer speed, it may take several seconds for the program to open.

☐ Don't show this message again

- The Coot window will be updated during the refinement
- You can move around the molecule but any changes you make to the model will not be saved



Strategy

Refinement strategy : ☑ XYZ (reciprocal-space) ☐ XYZ (real-space) ☐ Rigid body ☑ Individual B-factors ☐ Group B-factors ☐ TLS parameters ☐ Occupancies ☐ Anomalous groups

Number of cycles : 5

Note: selections can only be made for enabled options (e.g. NCS groups are available if "Use NCS" box is checked)

Targets and weighting

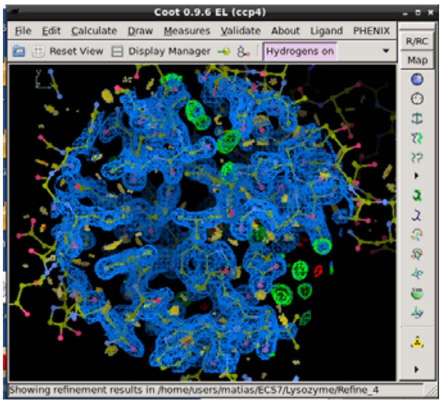
Target function : Automatic
☐ Optimize X-ray/stereochemistry weight
☐ Optimize X-ray/ADP weight
☐ Use NCS
NCS type : torsion-angle
☐ Reference model restraints
☐ Use secondary structure restraints
☐ Use experimental phase restraints

Other options

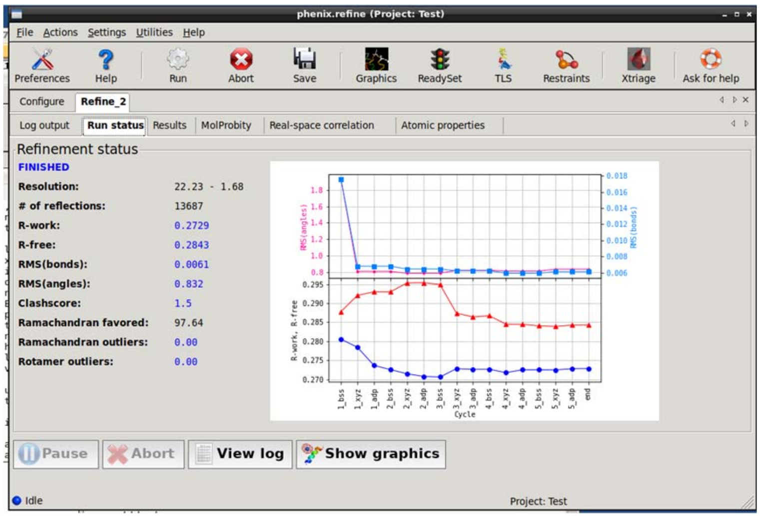
☐ Automatically add hydrogens to model
☐ Update waters
Place elemental ions :
☐ Simulated annealing (Cartesian)
☐ Simulated annealing (Torsion angles)
Scattering table : n_gaussian

Refining the structure with PHENIX – running phenix.refine

- The Run status tab shows the refinement statistics.
- Due to some reason, the R-work and the R-free will often increase in the first 2-3 cycles and sharply decrease afterwards



Showing refinement results in /home/users/matias/ECS7/Lysozyme/Refine_4



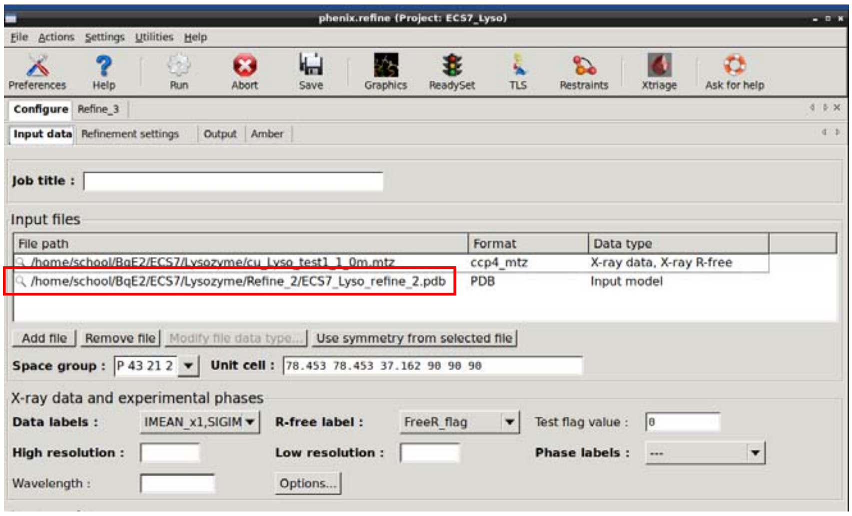
Refinement status

FINISHED

Resolution:	22.23 - 1.68
# of reflections:	13687
R-work:	0.2729
R-free:	0.2843
RMS(bonds):	0.0061
RMS(angles):	0.832
Clashscore:	1.5
Ramachandran favored:	97.64
Ramachandran outliers:	0.00
Rotamer outliers:	0.00

Refining the structure with PHENIX – running phenix.refine

- The 2F$_o$-F$_c$ map should look a lot better, but you can still go through the whole structure again to fix any problems you may notice. You will notice many green spheres of electron density (positive F$_o$-F$_c$ peaks). These are **water molecules**.
- In the next PHENIX run, we´re going to add water molecules automatically to the structure.
- Use as input coordinates the output coordinate file from the previous run (or the one from Coot).



Job title :

Input files

File path	Format	Data type
/home/school/BqE2/ECS7/Lysozyme/cu_Lyso_test1_1_0m.mtz	ccp4_mtz	X-ray data, X-ray R-free
/home/school/BqE2/ECS7/Lysozyme/Refine_2/ECS7_Lyso_refine_2.pdb	PDB	Input model

Space group : P 43 21 2 **Unit cell :** 78.453 78.453 37.162 90 90 90

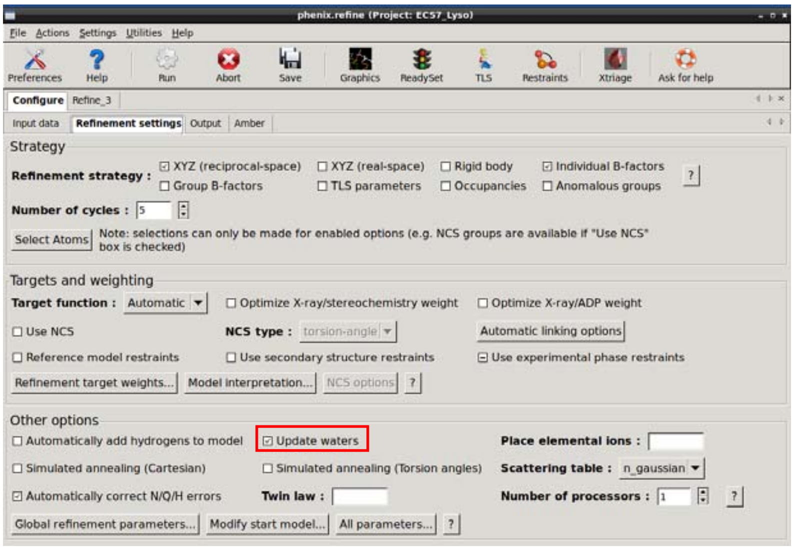
X-ray data and experimental phases

Data labels : IMEAN_x1,SIGIM **R-free label :** FreeR_flag Test flag value : 0

High resolution : **Low resolution :** **Phase labels :** ---

Wavelength :

Refining the structure with PHENIX – running phenix.refine



Strategy

Refinement strategy : ☑ XYZ (reciprocal-space) ☐ XYZ (real-space) ☐ Rigid body ☑ Individual B-factors ☐ Group B-factors ☐ TLS parameters ☐ Occupancies ☐ Anomalous groups

Number of cycles : 5

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

Other options

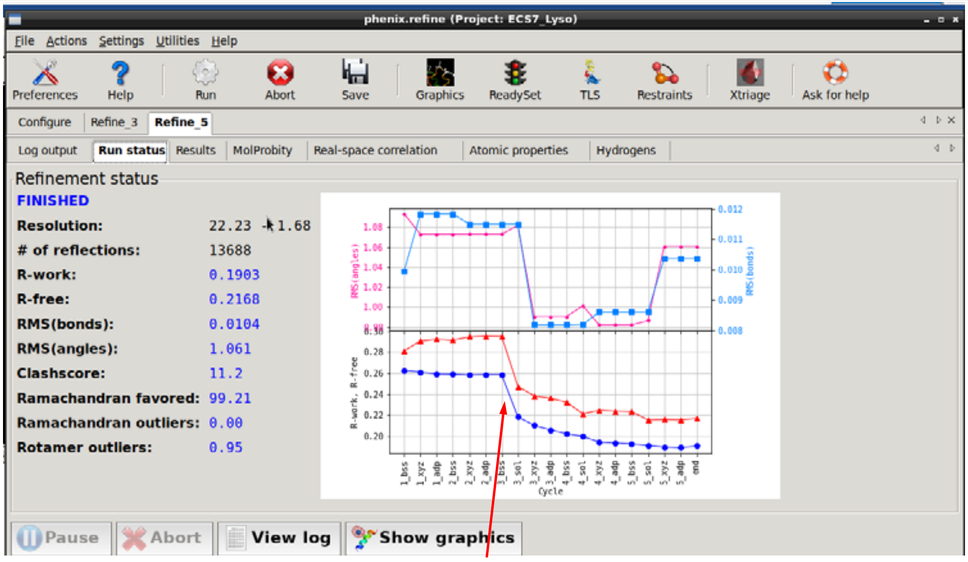
☐ Automatically add hydrogens to model ☑ Update waters **Place elemental ions :**

☐ Simulated annealing (Cartesian) ☐ Simulated annealing (Torsion angles) **Scattering table :** n_gaussian

☑ Automatically correct N/Q/H errors **Twin law :** **Number of processors :** 1

- Keep the same Refinement settings as before but check the **Update waters** box.

Refining the structure with PHENIX – running phenix.refine



Refinement status

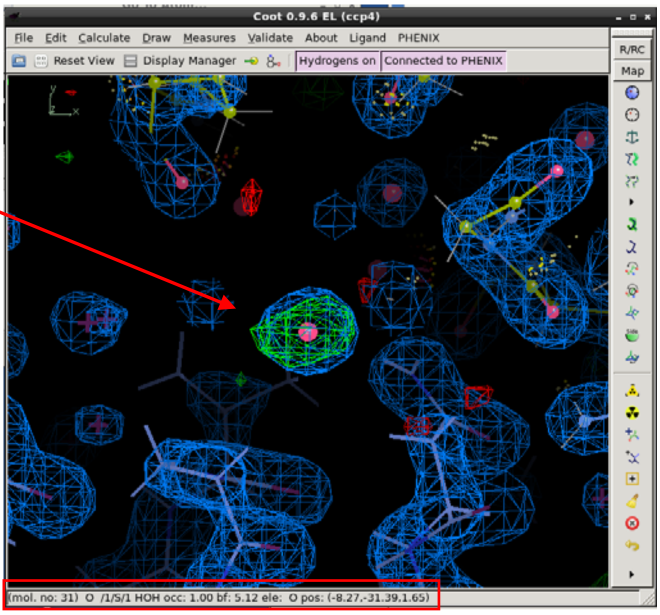
FINISHED

Resolution: 22.23 - 1.68
# of reflections: 13688
R-work: 0.1903
R-free: 0.2168
RMS(bonds): 0.0104
RMS(angles): 1.061
Clashscore: 11.2
Ramachandran favored: 99.21
Ramachandran outliers: 0.00
Rotamer outliers: 0.95

- Note the large decrease in R-work and R-free when the first group of water molecules are added.

Refining the structure with PHENIX – running phenix.refine

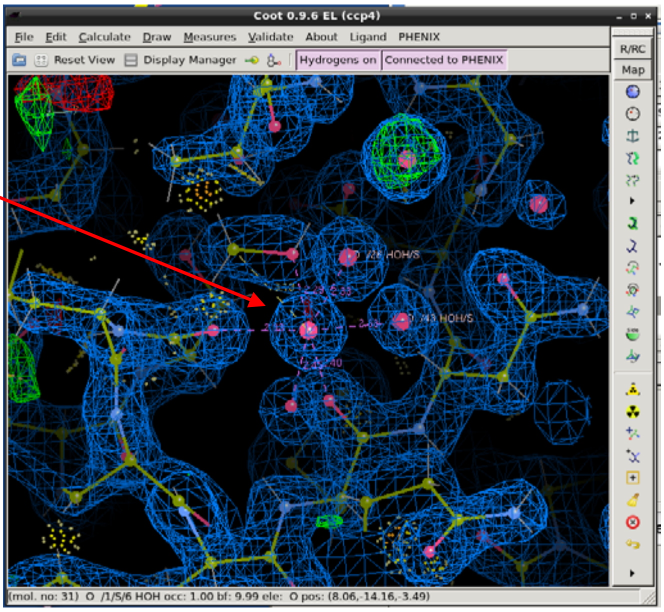
Is this a water molecule?



```
(mol. no: 31)  O  /1/S/1 HOH occ: 1.00 bf: 5.12 ele:  O pos: (-8.27,-31.39,1.65)
```

Refining the structure with PHENIX – running phenix.refine

Is this a water molecule?

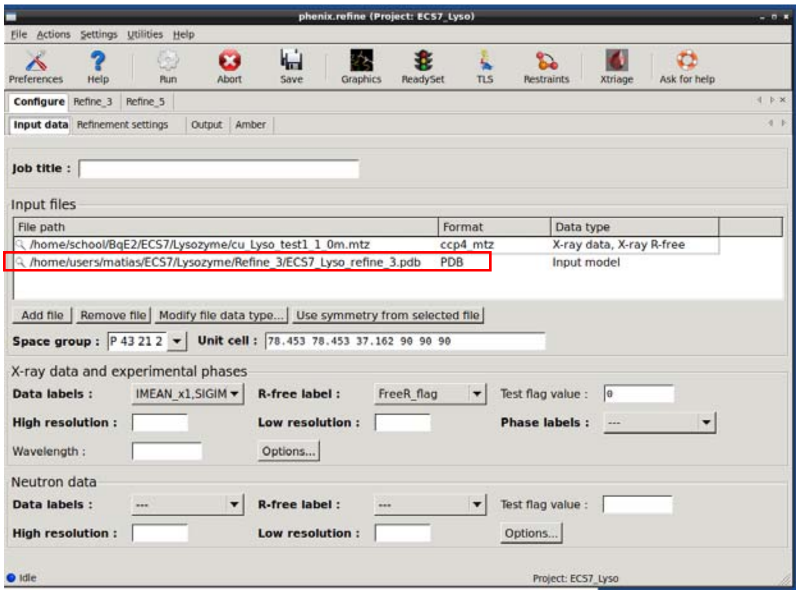


```
(mol. no: 31)  O  /1/S/6 HOH occ: 1.00 bf: 9.99 ele:  O pos: (8.06,-14.16,-3.49)
```

Refining the structure with PHENIX – running phenix.refine

Next, we can try TLS refinement of anisotropic thermal motion parameters.

First, load the final **coordinate file** after the last Coot edit. This file contains only the **isotropic** thermal motion parameters.



Job title :

Input files

File path	Format	Data type
/home/school/BqE2/ECS7/Lysozyme/cu_Lyso_test1_1_0m.mtz	ccp4_mtz	X-ray data, X-ray R-free
/home/users/matias/ECS7/Lysozyme/Refine_3/ECS7_Lyso_refine_3.pdb	PDB	Input model

Space group : P 43 21 2 Unit cell : 78.453 78.453 37.162 90 90 90

X-ray data and experimental phases

Data labels : IMEAN_x1,SIGIM R-free label : FreeR_flag Test flag value : 0

High resolution : Low resolution : Phase labels : ---

Wavelength :

Neutron data

Data labels : --- R-free label : --- Test flag value :

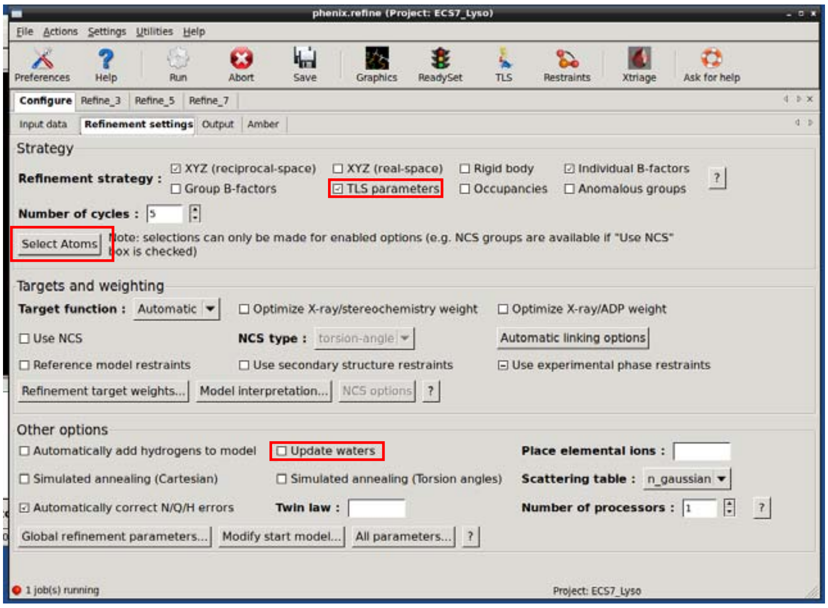
High resolution : Low resolution :

● Idle Project: ECS7_Lyso

Refining the structure with PHENIX – running phenix.refine

Next, in the **Refinement settings** tab check the **TLS parameters** box and click the **Select Atoms** button.

For a proper comparison we should uncheck the **Update waters** box.



Strategy

Refinement strategy : ☑ XYZ (reciprocal-space) ☐ XYZ (real-space) ☐ Rigid body ☑ Individual B-factors ☐ Group B-factors ☑ TLS parameters ☐ Occupancies ☐ Anomalous groups

Number of cycles : 5

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

Other options

☐ Automatically add hydrogens to model ☐ Update waters Place elemental ions :

☐ Simulated annealing (Cartesian) ☐ Simulated annealing (Torsion angles) Scattering table : n_gaussian

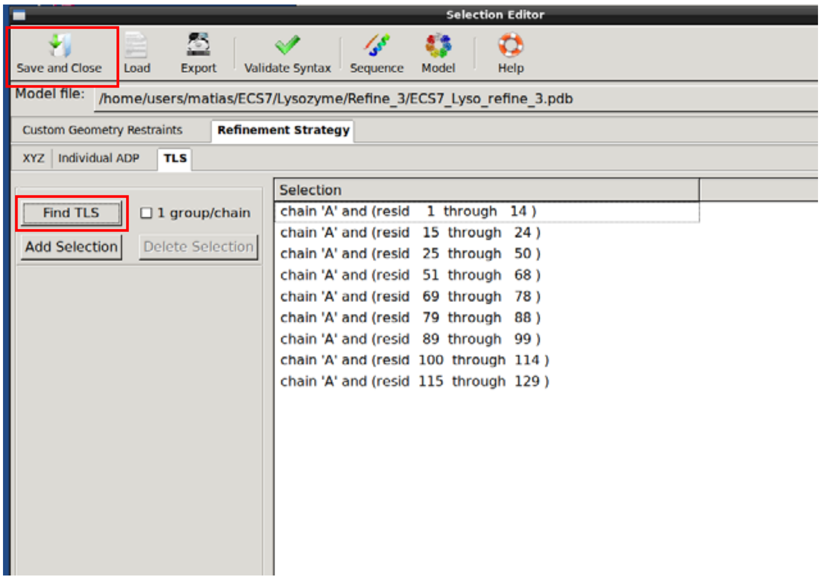
☑ Automatically correct N/Q/H errors Twin law : Number of processors : 1

● 1 job(s) running Project: ECS7_Lyso

Refining the structure with PHENIX – running phenix.refine

Now, in the **Refinement Strategy** tab select **TLS** and then click on the **Find TLS** button.

When the list of rigid body groups appears, click on the **Save and Close** button to exit this menu.



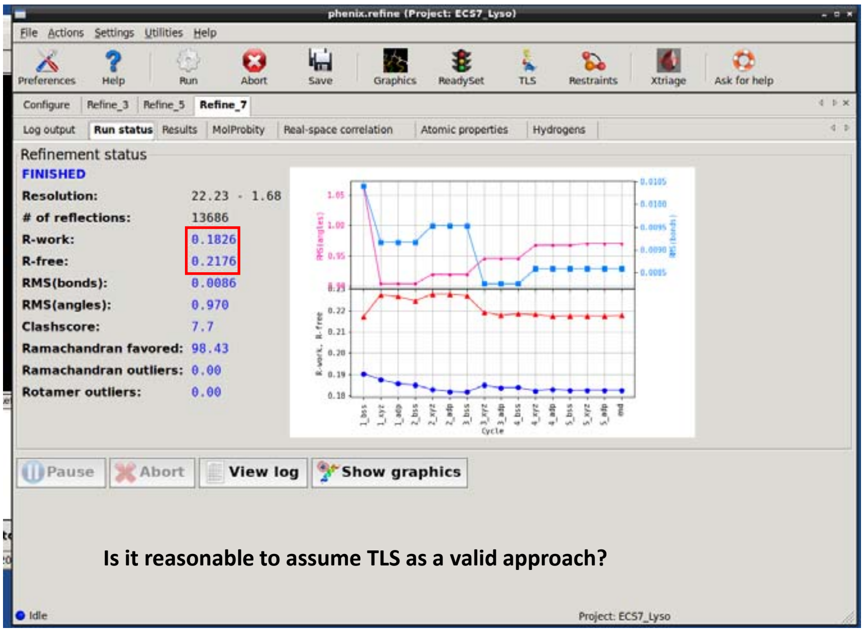
Model file: /home/users/matias/ECS7/Lysozyme/Refine_3/ECS7_Lyso_refine_3.pdb

Selection
chain 'A' and (resid 1 through 14)
chain 'A' and (resid 15 through 24)
chain 'A' and (resid 25 through 50)
chain 'A' and (resid 51 through 68)
chain 'A' and (resid 69 through 78)
chain 'A' and (resid 79 through 88)
chain 'A' and (resid 89 through 99)
chain 'A' and (resid 100 through 114)
chain 'A' and (resid 115 through 129)

Refining the structure with PHENIX – running phenix.refine

R-work and R-free w/o TLS:

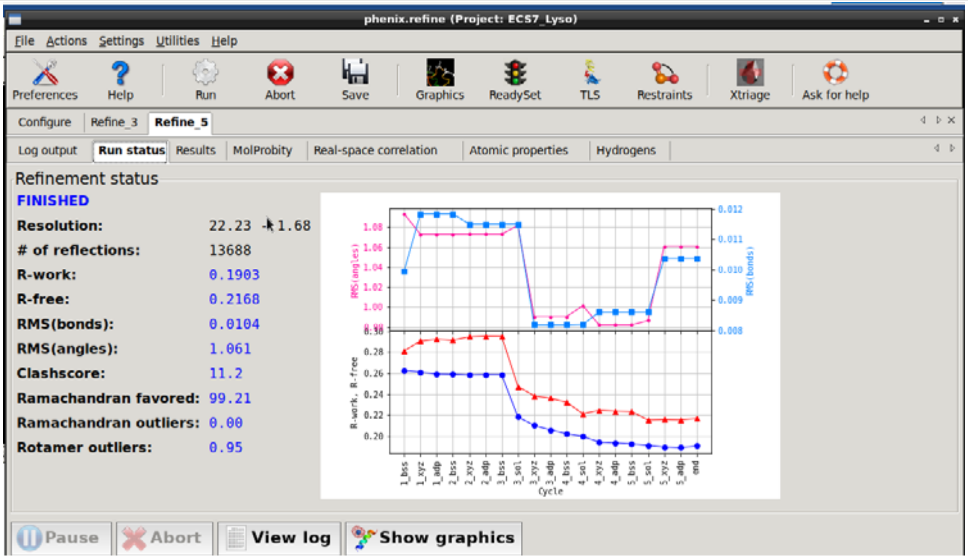
```
0.1903
0.2168
```



Refinement status
FINISHED
Resolution: 22.23 - 1.68
# of reflections: 13686
R-work: 0.1826
R-free: 0.2176
RMS(bonds): 0.0086
RMS(angles): 0.970
Clashscore: 7.7
Ramachandran favored: 98.43
Ramachandran outliers: 0.00
Rotamer outliers: 0.00

Is it reasonable to assume TLS as a valid approach?

Refining the structure with PHENIX – validation



Refinement status

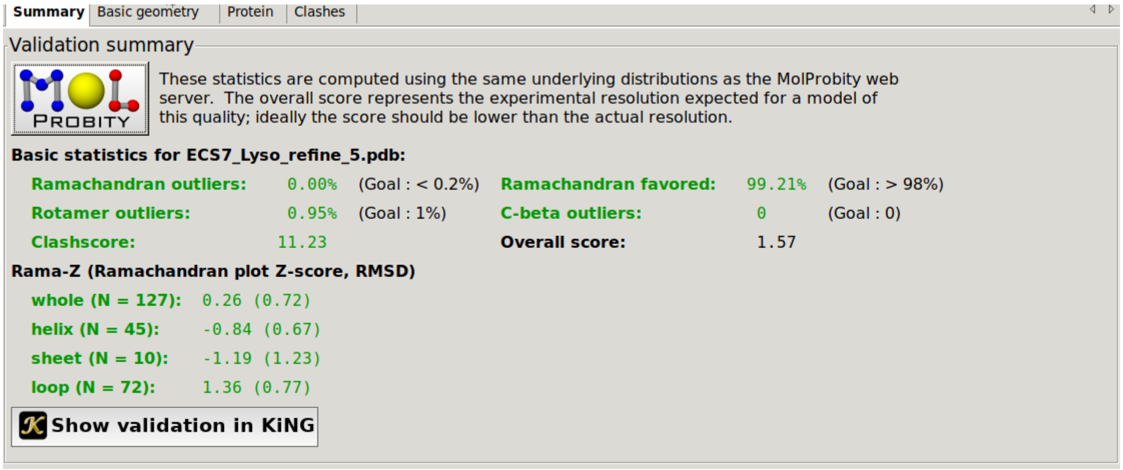
FINISHED

Resolution:	22.23 - 1.68
# of reflections:	13688
R-work:	0.1903
R-free:	0.2168
RMS(bonds):	0.0104
RMS(angles):	1.061
Clashscore:	11.2
Ramachandran favored:	99.21
Ramachandran outliers:	0.00
Rotamer outliers:	0.95

- Once we're satisfied that the structure cannot be improved any further, it's time to look at the Validation tools.

Refining the structure with PHENIX – validation

- The Validation summary gives an overall view of the validation statistics, but we can inspect the details in each of the tabs **Basic Geometry**, **Protein** and **Clashes**.



Validation summary

These statistics are computed using the same underlying distributions as the MolProbity web server. The overall score represents the experimental resolution expected for a model of this quality; ideally the score should be lower than the actual resolution.

Basic statistics for ECS7_Lyso_refine_5.pdb:

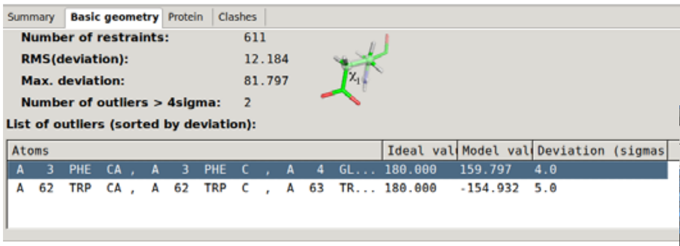
Ramachandran outliers:	0.00%	(Goal : < 0.2%)	**Ramachandran favored:**	99.21%	(Goal : > 98%)
Rotamer outliers:	0.95%	(Goal : 1%)	**C-beta outliers:**	0	(Goal : 0)
Clashscore:	11.23		**Overall score:**	1.57	

Rama-Z (Ramachandran plot Z-score, RMSD)

whole (N = 127):	0.26 (0.72)
helix (N = 45):	-0.84 (0.67)
sheet (N = 10):	-1.19 (1.23)
loop (N = 72):	1.36 (0.77)

Refining the structure with PHENIX – validation

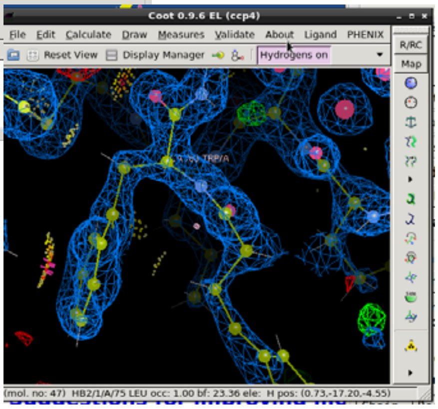
- Everything is fine, except for slightly distorted **peptide dihedral angles** between PHE3 and GLY4, and between TRP62 and TRP63.



Number of restraints: 611
RMS(deviation): 12.184
Max. deviation: 81.797
Number of outliers > 4sigma: 2
List of outliers (sorted by deviation):

```
Atoms                                          Ideal val Model val Deviation (sigmas
A   3  PHE  CA ,  A   3  PHE  C  ,  A   4  GL... 180.000   159.797   4.0
A  62  TRP  CA ,  A  62  TRP  C  ,  A  63  TR... 180.000   -154.932  5.0
```

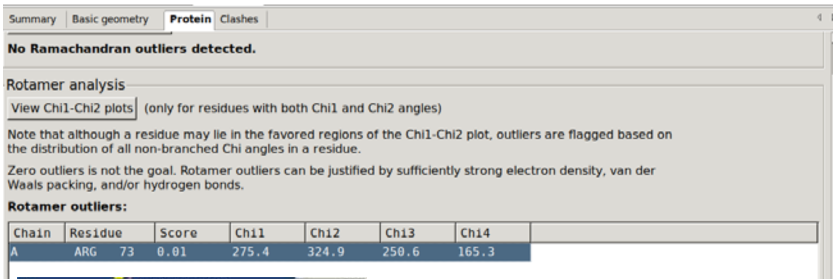
- However, this unusual geometry is **validated** by the electron density map.



(mol. no: 47) HB2/1/A/75 LEU occ: 1.00 bf: 23.36 ele: H pos: (0.73,-17.20,-4.55)

Refining the structure with PHENIX – validation

- Another flagged issue is a bad rotamer (side-chain conformation) at ARG73



No Ramachandran outliers detected.

Rotamer analysis

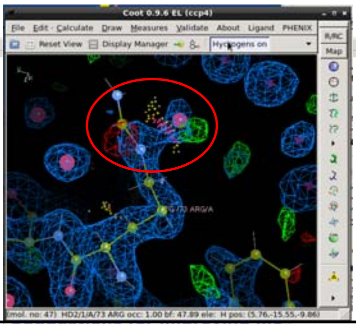
View Chi1-Chi2 plots (only for residues with both Chi1 and Chi2 angles)

Note that although a residue may lie in the favored regions of the Chi1-Chi2 plot, outliers are flagged based on the distribution of all non-branched Chi angles in a residue.

Zero outliers is not the goal. Rotamer outliers can be justified by sufficiently strong electron density, van der Waals packing, and/or hydrogen bonds.

Rotamer outliers:

```
Chain  Residue    Score   Chi1    Chi2    Chi3    Chi4
A         ARG  73  0.01   275.4   324.9   250.6   165.3
```

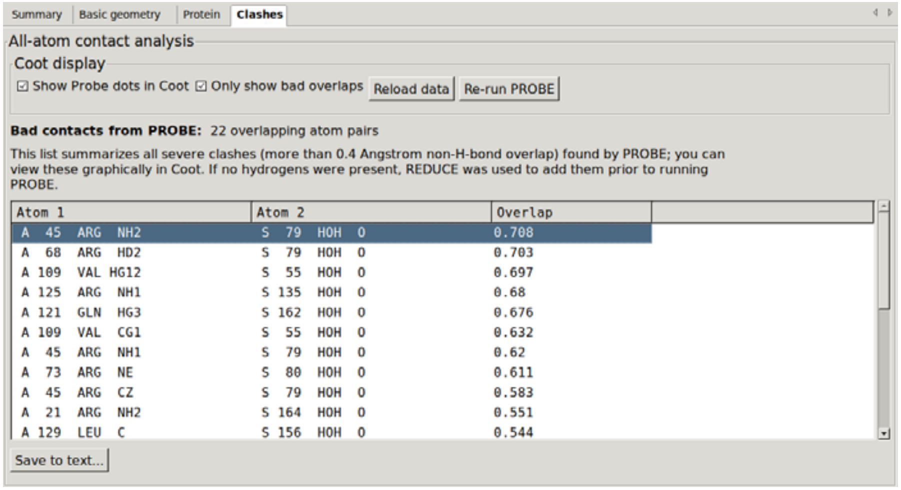


(mol. no: 47) HD2/1/A/73 ARG occ: 1.00 bf: 47.89 ele: H pos: (5.76,-15.55,-9.86)

- We can see that the guanidinium group is in the wrong place and that there is a water molecule where one of the nitrogen atoms should be. We should **delete the water molecule**, **adjust the side chain** and **run another refinement** to fix this issue.

Refining the structure with PHENIX – validation

- Finally, there are several “bad contacts” listed. Many involve hydrogen bonds but not all. It may be worth while to inspect these to see if they can be eliminated by tweaking the structure, but it will very likely be impossible to get rid of all of them.



Bad contacts from PROBE: 22 overlapping atom pairs

This list summarizes all severe clashes (more than 0.4 Angstrom non-H-bond overlap) found by PROBE; you can view these graphically in Coot. If no hydrogens were present, REDUCE was used to add them prior to running PROBE.

Atom 1	Atom 2	Overlap
A 45 ARG NH2	S 79 HOH O	0.708
A 68 ARG HD2	S 79 HOH O	0.703
A 109 VAL HG12	S 55 HOH O	0.697
A 125 ARG NH1	S 135 HOH O	0.68
A 121 GLN HG3	S 162 HOH O	0.676
A 109 VAL CG1	S 55 HOH O	0.632
A 45 ARG NH1	S 79 HOH O	0.62
A 73 ARG NE	S 80 HOH O	0.611
A 45 ARG CZ	S 79 HOH O	0.583
A 21 ARG NH2	S 164 HOH O	0.551
A 129 LEU C	S 156 HOH O	0.544