Tutorial

I - MOLECULAR REPLACEMENT PHASING

Resolution of the structure of hen lysozyme from diffraction of Hen lysozyme crystal and a coordinates file of Bovine lysozyme :

Hen lysozyme sequence :

>sp|P00698|LYSC_CHICK Lysozyme C OS=Gallus gallus GN=LYZ PE=1 SV=1 MRSLLILVLCFLPLAALGKVFGRCELAAAMKRHGLDNYRGYSLGNWVCAAKFESNFNTQA TNRNTDGSTDYGILQINSRWWCNDGRTPGSRNLCNIPCSALLSSDITASVNCAKKIVSDG NGMNAWVAWRNRCKGTDVQAWIRGCRL

Bovine lysozyme sequence:

>sp|P04421|LYSC_BOVIN Lysozyme C OS=Bos taurus GN=LYZ1 PE=1 SV=2 MKALVILGFLFLSVAVQGKVFERCELARTLKKLGLDGYKGVSLANWLCLTKWESSYNTKA TNYNPSSESTDYGIFQINSKWWCNDGKTPNAVDGCHVSCRELMENDIAKAVACAKHIVSE QGITAWVAWKSHCRDHDVSSYVEGCTL

Sequence associated to PDB coordinates file of Bovine lysozyme:

>2Z2F:A|PDBID|CHAIN|SEQUENCE KVFERCELARTLKKLGLDGYKGVSLANWLCLTKWESSYNTKA TNYNPSSESTDYGIFQINSKWWCNDGKTPNAVDGCHVSCSELMENDIAKAVACAKHIVSE QGITAWVAWKSHCRDHDVSSYVEGCTL

```
Hen and bovine lysozymes sequence alignment :
>2z2f A Lysozyme C-2; stomach lysozyme, 1,4-beta-N-acetylmuramidase C, bacteriolytic
enzyme, hydrolase; 1.50A {Bos taurus} SCOP: d.2.1.2
Length = 129 Score = 154 bits (389), Expect = 1e-38,
                                                             Method: Composition-based
stats. Identities = 71/130 (54%), Positives = 96/130 (73%), Gaps = 2/130 (1%)
Query: 19 KVFGRCELAAAMKRHGLDNYRGYSLGNWVCAAKFESNFNTQATNRNTDG-STDYGILQIN 77
          KVF RCELA +K+ GLD Y+G SL NW+C K+ES++NT+ATN N
                                                           STDYGI OIN
Sbjct: 1 KVFERCELARTLKKLGLDGYKGVSLANWLCLTKWESSYNTKATNYNPSSESTDYGIFOIN 60
Query: 78 SRWWCNDGRTPGSRNLCNIPCSALLSSDITASVNCAKKIVSDGNGMNAWVAWRNRCKGTD 137
          S+WWCNDG+TP + + C++ CS L+ +DI +V CAK IVS+ G+ AWVAW++ C+ D
Sbjct: 61 SKWWCNDGKTPNAVDGCHVSCSELMENDIAKAVACAKHIVSE-QGITAWVAWKSHCRDHD 119
Query: 138 VQAWIRGCRL 147
          V +++ GC L
Sbjct: 120 VSSYVEGCTL 129
```

You will now calculate the phase of the observed amplitudes Fobs(hkl), by the molecular replacement method using the software **MOLREP** from CCP4i package. The interface of the software opens by clicking on **Molrep** :



The amplitudes and their associated errors will be extracted from the file « lyso-1-1-A-IF.mtz » that you must enter on line « MTZ in ». The phase information will be calculated from the PDB file "**2Z2F.pdb** ", line « Model In ». The « solution » PDB file will be named **« 2Z2F_molrep1.pdb »**, as indicated on line « Coords out ».

	😑 🕙 🔿 🕅 Molrep – Molecular Replacement Initial parameters from /Users/ml161111/Desktop/E	
	Enter input MTZ file name (HKLIN)	Help
	This interface is for version 11.2 of Molrep	
	Job title Lyso_Taurus	
	Do molecular replacement — performing rotation and translation function —	-
	Get input structure factors from MTZ file 🛁	
	Input fixed model	
	🔟 Multi-copy search	
	Use sequence	
C	MTZ in TP-Oleron - Iyso-1-1-A-IF.mtz Browse Vi	ew
	Use _ Intensities Enter input MTZ file name	
	FPSIGFPSIGF	
	Model in TP-Oleron - 2Z2F.pdb Browse M	ew
C	Coords out TP-Oleron - 222F_molrep1.pdb Browse M	ew
	Experimental Data (Resolution,ANISO,DIFF,BADD,INVER,DSCALE,)	
	The Model (SIM,COMPL,SURF,NMR,NCSM,DSCALEM)	
	Search Parameters (NMON,NP,NPT,PST,STICK,LOCK,)	
	Infrequently Used Parameters (MODE,SAPTF,RAD,PACK,SCORE,LMIN,NOSG)	
		1
	🗾 Run 💴 Save or Restore 💷 Close	

Press Run to start the calculation.

When the calculation is done, the word FINISHED appears in the CCP4i Gui.

000 🛛	CCP4 Program	Suite 6.4.0 CCP4	nterface 2.2.1 r	unning on is1519	72.intra.cea.fr Project: TP-C	leron		
Edit information or notebook entry	/ associated with	selected job(s)					Change Project	Help
Program List	- 1	11:1 :28	FINISHED	m lrep	Lyso_Tauru 🛆	Directo	ories&ProjectDir	1
Mapmask								
Mapmask (fft)						VI	ew Any File	
Mapmask (patterson)						View Files fro	m Job 🗧	- [4]
					1	0	Databasa.	

You can open the output file by double click on the job line Two tabs are going to allow you to analyze the results of your calculation :

	00					4	qtRView 1	1.09 – Job	1: Lyso_Ta	aurus					
	ا ا	PDF		?										×	0
	Print Pl	DF/PS	Refresh	molrep	CCP4									Preferences	Exit
	Results	Log File	Help												
7	1						Job 1	: Lyso_	_Taurus	;					
	P Run P Ma A. J.	I Of MOL Please cite: OLREP: an an Vagin, A. Teply Appl. Cryst. (REP(cc utomated p vakov. 1997) 30, 1	rogram for 1 022-1025.	18/ 5/2015 nolecular replace	at 11:19:0 ^{ment}	2						_	_	
	🔽 Inpu	ıt Files	_	_	_	_	_	_	_	_	_	_	_	_	
	► R	leflection c	lata	Vie	wHKL										
	D M	lodel struc	ture		coot c	cp4mg	Display								
	🔽 Outp	out Files	_	_	_	_	_	_	_	_	_	_	_	_	_
	▶ S	tructure			Coot C	cp4mg	Display								

The Results section allows you to open the input and output with COOT, ccp4mg or Display.

● ● → qtRView 1.09 - Job 1: Lyso_Taurus ↓	Preferences	0 Exit
Results Log File Help		
ACCPERTING CONTINUETROE 2.2.1 ECCPERTING CONTINUETROE 2.2.1 ECCPERTING TO BAY 2015 11:19:02 ACCPERTING THE ANY 2015 ACCPERTING THE ACCPERTING TH		
### CCP4 6.4: MOLARP(ccp4) version 11:2.05: ### CCP4 6.4: MOLARP(ccp4) version 11:2.05: ####################################		
Plasas reference: Collaborative Computational Project, Number 4. 1994. "The CDPA Suite: Programs for Frotein Crystallography". Acta Cryst. D50, 760-763. as well as any specific reference in the program write-up.		
970X71me6fcrcon1.85 comment 85 MGREPs a matchaskad program for molecular replacement A.Vagin,A.Teplyakody. Ogram. 3. Appl. Cryst. (1997) 30, 1022-1023.		
\$\$ \$SUMMARY :Reference1: \$\$ MOLREP: \$\$!TEXT:Reference1: \$\$		
MOLREP /Vers 11.2.05; 31.07.2013/		

The Log file contains the information regarding the calculation itself.

Questions

- 1. Which parameters will allow to determine the correctness of **MOLREP** solution ? / *Quels paramètres vont permettre de déterminer si la solution trouvée par MOLREP est correcte ?*
- 2. What is the reliability factor Rfac of the solution found by MOLREP? / *Quel est le facteur de confiance Rfact de la solution trouvée par MOLREP* ?
- 3. What contrast has the solution found by MOLREP? / *Quel contraste a la solution trouvée par MOLREP*?
- 4. Open the output file structure with **COOT**, and visualize the solution. What can tell you that the solution is probably correct (or wrong)? *Ouvrez le fichier de sortie avec* **COOT** *et visualisez la solution. Qu'est-ce qui peut vous indiquer que la solution est coorecte (ou fausse)*?

<u>II - REFINEMENT OF MOLECULAR REPLACEMENT SOLUTION</u> <u>WITH REFMAC</u>

1) Before running the refinement with Refmac ...

Check that a **free Rfactor flag** (FreeRflag) has been set in your mtz data file. If you have followed the XDS data processing tutorial, the FreeRflag has already been added to the mtz file (loot at the uniquemtz.out log file).

If you have used imosflm, you should add the FreeRflag as follow:

- Check whether the FreeRflag is present in the mtz file:

type in the terminal window:

mtzdump hklin lyso-1-1-A-IF.mtz

go

if FreeRflag is absent, type: freerflag hklin lyso-1-1-A-IF.mtz hklout lyso-1-1-A-IFfree.mtz << eof freerfrac 0.05 end eof

and check again: mtzdump hklin lyso-1-1-A-IFfree.mtz go

if the FreeRflag is there now, lyso-1-1-A-IFfree.mtz is the file to use for the next steps.

2) Refinement with Refmac

Based on molecular replacement solution, you can initiate REFMAC :



As previously, the amplitudes and their associated errors will be extracted from the file « lyso-1-1-A-IF.mtz » that you must enter on line « MTZ in ». The phase information will be directly calculated from the MOLREP solution PDB file "2Z2F_molrep1.pdb ", line refined model PDB « Model In ». After refinement, the file will be « 2Z2F_molrep1_refmac1.pdb », as indicated on line « PDB out ». The amplitudes, and electron density map coefficient will be in the binary file « lyso-1-1-A-IF refmac1.mtz », as indicated on line MTZ out.

O O 🛛 🕅 🛛 Run Refmac5		
ob title Refinement of MOLREP lysozyme solution		
o restrained refinement - using no prior phase information - input		
Input fixed TLS parameters		
no 🔤 twin refinement		
lse Prosmart: no - (le	w resolution ref	inement
Run Coot:findwaters to automatically add/remove waters to refined structure		
ITZ in D-Oleron 🖂 lyso-1-1-A-IF.mtz	Browse	View
P F - Sigma SIGF		-
ITZ out P-Oleron - Iyso-1-1-A-IF refmac1.mtz	Browse	View
DB in D Oleron 222E moleron and	Browse	1.60.00
	Diowse	VIEW
DB out P-Oleron222F_molrep1_refmac1.pdb	Browse	View
IB in TP-Oleron - Merge LIE	INS Browse	View
Dutput lib TP-Oleron _ 222F_molrep1.cif	Browse	View
lefmac keyword file TP-Oleron -	Browse	View
ata Kanustina		· · · ·
anna rus rus ang		-
atun Roomatric Roctrainte		-
adar dan Cristiania		-
Lies sutematically appended local NCS metraints		_
lo NCC restraints		
e nos resumins de contenuy denned		
Edit list		estraint
xternal Restraints		
Aonitoring and Output Options		
Calina		
Run - Save or Restore -	Close	

Press Run to start the calculation.

The REEMAC	ioh will ha	RUNNING first	thon	FINISHED	
пе кегмас	JOD WIII DE	KUNNING III SU	, then	LINISUED	

00	X CCP4 Program	Suite 6.4.0 CCP4	nterface 2.2.1 r	unning on is15197	2.intra.cea.fr Project: TP-O	leron
List of jobs for project. Double-	click on a job display:	s the log file, shift	-double click ren	uns the job.		Change Project Help
Program List	- 2	10:28:2	RUNNING	refmac5	Refinement 🛆	Directories&ProjectDir
ProDrg		18 May 15	THICUT	molrep	Lyso_Tauru	View Any File
Professs					-	
ProSMART						View Files from Job
		Suite C 4.0 CCB4	Interferen 2.2.1		12 later car fr. Decisity TD.	News
000	X CCP4 Program	Suite 6.4.0 CCP4	Interface 2.2.1	running on is1519/	2.intra.cea.fr Project: IP-0	Dieron Dielost Hele
					-	Change Project Help
Program List	- 2	10:28:5	FINISHED	refmacb	Refinement 🛆	Directories&ProjectDir
ProDrg	1	18 May 1.		molrep	Lyso_Tauru	View Any File
Professs						

As previously, you can open the output file by double click on the job line. There is also one Results section and one Log File section in the output file.

Pergeta Log Fie	Job 2: Refinement of MOLREP lysozyme solution 5 at 10:26/23 Infortures* Infortures* Name: RANNetwork, M.D. Wore, F.Long and A.A. Vagin, (2011) um Läutihood Method.*	
Pesult Initial Final A Error 0.5022 0.4 A Error 0.5021 0.4 Constant of the error of the er	0.7 0.6 0.7 0.7 0.6 0.7 0.7 0.7 0.7 0.7 0.7 0.7 0.7	
Cycle 1. Rfactor analysis, F distribution v r Cycle 1. Fom(cost[DelPhil)>acentric, centri Cycle 11. Rfactor analysis, F distribution v r Cycle 11. Rfactor analysis, F distribution v r Cycle 11. Fom(cost[DelPhil)=acentric, cent P Gradet analysis, stats va cycle Input Files I Model structure Reflection data For Cutput Files Cutput Files I Structure and electron density	sin e, overall v resin sin coot cop4mg Display wHKL Doot cop4mg ViewHKL Display	

Press Coot to open the output files (pdb and mtz files) in coot.



<u>Questions</u>

- 1. What parameters inform on the refinement quality?
- 2. What appears in the Coot window?
- 3. Based on the lysozyme sequences alignment, can you correct the refined structure in Coot?

3) Model building with Coot

The minimum to know about Coot:

Coot is fairly "intuitive": most of the necessary action can be selected from the icons on the right icon bar of the Coot window: just put the mouse arrow on the icon to know what it does.

Before to start model building

Check out the contouring level of electron density maps: {2Fobs – Fcalc} maps (FWT, PHWT) should be contoured at 1 rms {Fobs – Fcalc} residual maps (DELFWT, PHDELWT) should be contoured at +3 or -3 rms

Clic on "Display Manager" and select the scroll button corresponding to the desired map; the use the middle wheel of your mouse to increase or decrease the level.



000	🗴 Display Manager	
Maps All		
1 Iyso-1-1-A-IF_refmac2.mtz FWT PHWT	🗹 Display 💿 Scroll Poperties Delete Map	^
2 Iyso-1-1-A-IF_refmac2.mtz DELFWT PHDELWT	Display O Scroll Properties Delete Map	
] [4]		Ŧ
Molecules All		
0 2Z2F_molrep1_refmac2.pdb	Display Active Bonds (Colour by Atom)	^
		Ŧ
	💥 Close	1

- Once you have made the desired structural change in coot, you can save the new coordinates of the model ("file" menu) and run refmac again.

- You do cycle coot and refmac steps until you have interpreted all the significant electron density, with a good fit between your model and the electron density.

At some point (but you might not reach it during the practical), you may need to look at the refinement parameters submenu in Refmac:

000	🔀 Run Refmac5			
		4		H
Job title Refmac_cycle2				
Do restrained refinement usin	no prior phase information	input		
_ Input fixed TLS parameters				
no twin refinement				
Use Prosmart:	no	(low r	esolution refi	nement)
Run libg to generate external restraint	s (DNA/RNA) automatically 📃			
Run Coot:findwaters to automatically add/r	remove waters to refined structure			
MTZ in Refmac-lyso 🛁 <mark>lyso-1-1-A-IF_refm</mark> a	ac1.mtz		Browse	View
FPF	Sigma	SIGF		
MTZ out Refmac-lysolyso-1-1-A-IF_refm	nac2.mtz		Browse	View
PDB in Refmac-lyso = 222F_molrep1_refm	ac1.pdb		Browse	View
PDB out Refmac-lyso - 222F_molrep1_ref	mac2.pdb		Browse	View
LIB in Refmac-lyso 🖃 🔽		Merge LIBINs	Browse	View
Output lib Refmac-lyso _ 2Z2F_molrep1_re	efmac1.cif		Browse	View
Refmac keyword file Refmac-lyso 🖃			Browse	View
Data Harvesting				
Refinement Parameters				
Setup Geometric Restraints				
Setup Non-Crystallographic Symmetry (NCS) I	Restraints			
External Restraints				
Monitoring and Output Options				
Scaling				
Geometric parameters				
Dum	Source av Dead		Class	
- Run -	Save or Rest		Cluse	

0 0 0 🕅 🕅 Refmac5		
		He
b title Refmac_cycle2		
p restrained refinement using no prior phase information input		
Input fixed TLS parameters		
no twin refinement		
se Prosmart: (low res	solution refin	ement)
Run libg to generate external restraints (DNA/RNA) automatically		
Run Coot:findwaters to automatically add/remove waters to refined structure		
TZ in Refmac-lyso — lyso-1-1-A-IF_refmac1.mtz	Browse	View
· F Sigma SIGF		
IZ out Refmac-lyso lyso-1-1-A-IF_refmac2.mtz	Browse	View
JB in Refmac-lyso - 2Z2F_molrep1_refmac1.pdb	Browse	View
DB out Refmac-lyso - 222F_molrep1_refmac2.pdb	Browse	View
B in Refmac-lyso - Merge LIBINs	Browse	View
tput lib Refmac-lyso [222F_molrep1_refmac1.cif	Browse	View
efmac keyword file Refmac-lyso 🛁	Browse	View
ita Harvesting		
ninement Parameters		
10 cycles of maximum likelihood restrained refinement		
e hydrogen atoms: generate all hydrogens and output to coordinate file		
Resolution range from minimum 55.847 to 1.101		
use automatic waveshad at use experimental signals to weight Xray terms		
fine isotronic – temperature factors		
Eve isotropic		
		-
verall ICA Bostrointe		
mixed (isotropic/anisotropic)		
onitoring Bira - that Ontion		
aling		
ometric parameters		
Dura de Destante de La companya de la	Gasa	
Save or Restore	Close	

You may have to modify:

- the number of cycles: put more than 10 if your Rfactor is not stable after 10 cycles

- the B factor (temperature factor) : isotropic (default, one B factor per atom), overall (one B factor for the entire model, recommended for low resolution data, i.e. $d_{max} > 3.2$ Å), anisotropic B factor (6 parameters per atom, model the anisotropy of atom motion, useful for high resolution data, i.e. $d_{max} < 1.3$ Å)