X-ray crystallography practical

Oleron 2017

Last update: 19 June 2017

Table of Contents

Schedule of the X-ray crystallography practical	1
Data collection on a Gd derivative on beamline FIP	2
Data reduction with iMosflm	3
Start iMosflm	3
Indexation	3
Integration and scaling	5
Data reduction with XDS package	6
Introduction	6
Default xds parameter files	8
Run all step in a row	8
Re-run scaling	8
xscale and xdsconv	9
Data reduction with XDS in graphic/automated mode	11
Graphic mode, using xdsGUI	11
Automated data reduction	12
Quick SAD phasing with Phaser in ccp4i	14
Quick SAD phasing in ccp4i2	18
Run the SHELX automated pipeline	18
Results analysis	20
Automated SAD phasing with Phaser	23
SAD phasing, step by step, with ccp4i	25
MR	34
Complete model	34
Partial/remote model	34
Refinement / construction	34
Complete model	34
Partial/remote model	34

Schedule of the X-ray crystallography practical

Tuesday 20 June, 9:15AM-12:45AM (1h45 + 1h30)

9h00-9h15 : sponsors

Présentation du TP, constitution des groupes (~15 min)

Data reduction (~30 min)

Données : BM30A-2014-11-19 (native, 164 frames, 1.1 A) lyso poule 15-05-13-lyso/lyso-Gd_SAD (anomales Gd, 300 frames, 1.65 A) Soft : xds, xdsGUI ou mosflm Support : tutorial_xds ou tutorial_iMOSFLM

Phasage par MR (~30 min)

Données : BM30A-2014-11-19 , divers modèles : lyso incomplet de poule ou modèle complet "basse" homologie (boeuf) Soft : phaser Support : ?

Phasage SAD (~30 min)

Données : lyso-Gd_SAD Soft : ccp4 Support : tutorial_SAD, tutorial_SAD_bis

Pause café (11h00-11h15)

Construction des parties manquantes / affinement (~45 min)

Données : BM30A-2014-11-19, modèle lyso incomplet Soft : coot, refmac Support : ?

Construction des parties manquantes / affinement (~45 min)

Données : lyso-Gd_SAD, modèle lyso incomplet, modèle ligand+Gd Soft : coot, refmac Support : tutorial SAD, tutorial SAD bis

Repas (12h45)

Data collection on a Gd derivative on beamline FIP

The fluorescence of Gd was measured with a Roentec MCA at the Gd LIII edge. Raw data are in Edge/lyso_1_Gd1 (columns 5 and 7)

```
and the plot vs Energy in
```

Edge/lyso_1_Gd1.jpg



The spectrum was processed with Chooch. Final drawing of calculated f and f' is in Edge/final.jpg



Based on that, beam energy was tuned to 7242.6 eV, and a single-wavelength dataset was collected (300 frames, 1 deg each). Frames (compressed with bzip2) are named img/e000_prefix_1_00xxx.img.bz2

Data reduction with iMosflm

Start iMosflm

The main steps are:

In "Settings \rightarrow Experiment settings", change select "Reverse direction of spindle rotation" In "Session \rightarrow Add images..." open the images (uncompress images) first.

	iMosflm version 7.2.1, 21st September 2015	~ ~ &
ngs		Help
♣157.59 第157.80 ↔11	4.20	m
Images		
A Lattice I	11.21, 11.21, 38.10, 90.00, 90.00, 90.00 p43212	8
Mosaicity	0 404	
Mosaic block size	100	
E Sector e000 prefi	r 1 ####### :0.00->99.00	
O Matrix	e000 prefix 1 00001	
O Image 1	\notin (r):0.00 - 1.00, \notin (x):-0.01, \notin (y):-0.07, \notin (z):-	0.05
Image 2	\notin (r):1.00 - 2.00, \notin (x):-0.01, \notin (y):-0.07, \notin (z):-	0.05
- Image 3	\notin (r):2.00 - 3.00, \notin (x):-0.01, \notin (y):-0.07, \notin (z):-	0.05
- Image 4	\notin (r):3.00 - 4.00, \notin (x):-0.01, \notin (y):-0.07, \notin (z):-	.0.05
- Image 5	∉ (r):4.00 - 5.00, ∉ (x):-0.04, ∉ (y):-0.07, ∉ (z):-	·0.05
-O Image 6	∉ (r):5.00 - 6.00, ∉ (x):-0.04, ∉ (y):-0.06, ∉ (z):-	.0.08
🔵 Image 7	∉ (r):6.00 - 7.00, ∉ (x):-0.04, ∉ (y):-0.07, ∉ (z):-	.0.07
-O Image 8	∉ (r):7.00 - 8.00, ∉ (x):-0.04, ∉ (y):-0.07, ∉ (z):-	.0.07
🔵 Image 9	∉ (r):8.00 - 9.00	
Image 10	∉ (r):9.00 - 10.00	
🗌 🔵 Image 11	∉ (r):10.00 - 11.00	
Image 12	∉ (r):11.00 - 12.00	
Image 13	∉ (r):12.00 - 13.00	
Image 14	\notin (r):13.00 - 14.00	
Image 15	∉ (r):14.00 - 15.00	
Image 16	$\not\in$ (r):15.00 - 16.00	
Image 17	$\not\in$ (r):16.00 - 17.00	
Image 18	$\mathcal{E}(\mathbf{r}): 17, 00 = 18, 00$	
Image 19	∉ (r):18.00 - 19.00	
Image 20	$\mathcal{L}(\mathbf{r}): 19.00 - 20.00$	
Image 21	$(\mathbf{r}): 20.00 - 21.00$	
Trage 22	$(\mathbf{r}):$ 22.00 - 22.00	
Trage 24	$(\mathbf{r}):22.00 = 23.00$	
Thage 24	(±).23.00 - 24.00 (±).24.00 - 25.00	
Trage 25	$f(x) \cdot 25.00 = 26.00$	
Tmage 20	$a(r) \cdot 26.00 = 27.00$	
Tmage 28	$z(r) \cdot 27 \ 00 = 28 \ 00$	
Tmage 29	\neq (r) · 28 00 - 29 00	
Trage 20	# /~\.20.00 20.00	
- Image 26 - Image 27 - Image 28 - Image 29 - Image 29		∉ (r):25.00 - 26.00 ∉ (r):26.00 - 27.00 ∉ (r):27.00 - 28.00 ∉ (r):28.00 - 28.00 ∉ (r):28.00 - 29.00 ⊄ (r):28.00 - 29.00

Indexation

Run "Indexing". By default images 2 and 91 are used. You can use more by replacing "2, 91" by "2-6, 91-95", for an exemple. Click on "Index" to redo the indexing.

Select the correct space-group: nb 96, $P4_{3}2_{1}2$



then run "Cell Refinement"



Integration and scaling

Select "Integration"



Then run "QuickSymm" (quick scaling to check the symmetry) and "QuickScale" (scaling) Check the Rmerge and Rmeas

Notably, compare Rmerge or Rmeas calculated either on:

all I+ and I- (includes the centrosymmetry (Fridel's law))

or

within I+ or I- (excludes the centrosymmetry (Fridel's law))

DeltaAnom correlation between half-sets tells you whether the anomalous signal is significant: You divide your data into two sets, and compare the anomalous difference I^+ - I^- in the two sets. If the correlation is high, the signal is significant.

You can look also at the other figures on the table. All the integrated and scaled data are stored in the file ctruncate_A5_1__00002-unique.mtz

This file contains h, k, l, FreeRflag, F, SigF, Dano, SigDano, F(+), SigF(+), F(-), SigF(-), Isym, Imean, SigImean I(+), SigI(+), I(-), SigI(-)

For the following steps, you can use the **ctruncate_A5_1__00002-unique.mtz** file.

Data reduction with XDS package

Introduction

XDS is a suite of programs dedicated to the reduction of macromolecular crystallography data. The suite of programs include:

xds: data processing, from images to unmerged h,k,l,Intensities, sigma(Intensities)
xscale: scaling and merging Intensities from either one or several data sets.
xdsconv: converts reflection data files as obtained from xds or xscale into various formats required by software packages for crystal structure determination like CCP4, CNS (X-PLOR), or SHELX.
2cbf: converts a detector image file to CBF format. (not often used)

merge2cbf: converts a series of detector image files to CBF format. (not often used)

cellparm: used to determine the mean of the cell parameters obtained from processing several data sets from the same crystal form. (not often used).

Only xds, xscale and xdsconv will be used here.

xds requires

- diffraction images

- a parameter file called XDS.INP that contains all the necessary information regarding the experimental setup.

Most of the time, an XDS.INP file is generated automatically when your launch a data collection at a synchrotron. However, the file XDS.INP needs some editing during the data processing, but only a few input parameters require to be looked at. See the commented XDS.INP file for further details.

The whole data processing includes 7 steps define in the JOB= command line. Each step generate a log file named with the .LP suffix.

XYCORR: computes a table of spatial correction values for each pixel: allow to precisely localise each pixel of the detector. Fully automatic, to be done once. files created:

X-CORRECTIONS.cbf Y-CORRECTIONS.cbf XYCORR.LP

INIT: determines an inital background for each detector pixel and finds the trusted region of the detector surface. Needs 5 to 10 images to run properly (look at BACKGROUND_RANGE= command). To be done once.

files created:

BKGINIT.cbf BLANK.cbf GAIN.cbf INIT.LP

COLSPOT: collects strong diffraction spots from a specified subset of the data images (see

SPOT_RANGE= command). files created: FRAME.cbf SPOT.XDS COLSPOT.LP

IDXREF: interprets observed spots by a reciprocal lattice and refines all diffraction parameters (cell dimensions, orientation matrix, crystal-detector distance, etc ...).

files created: XPARM.XDS IDXREF.LP

DEFPIX: defines the trusted region of the detector, recognizes and removes shaded areas, and eliminates regions outside the resolution range defined by the user. files created:

BKGPIX.cbf ABS.cbf DEFPIX.LP

XPLAN: helps planning data collection. Tells you what data to collect in order to get the most complete data set. Only useful when at the synchrotron beamline, before launching the data collection.

files created:

XPLAN.LP

INTEGRATE: collects 3-dimensional profiles of all reflections occuring in the data images and estimates their intensities

files created:

INTEGRATE.HKL INTEGRATE.LP

CORRECT: corrects intensities for decay, absorption and variations of detector surface sensitivity, reports statistics of the collected data set and refines the diffraction parameters using all observed spots.

files created:

ABSORP.cbf DECAY.cbf DX-CORRECTIONS.cbf DY-CORRECTIONS.cbf GX-CORRECTIONS.cbf GY-CORRECTIONS.cbf MODPIX.cbf GXPARM.XDS XDS_ASCII.HKL CORRECT.LP

The different steps are presented in a series of directories, for sake of clarity. In practice, they can be performed in a single directory by successive modification of the input files and running the XDS package programs at the command line.

Default xds parameter files

Go to the directory xds_step0_default-XDSINP

You will find there the initial parameter file for xds, as created automatically by the beamline control software (simplified version)

xds_step0/XDS.INP
as well as a fully commented one
 xds_step0/XDS.INP_sav

Run all step in a row

Go to directory **xds_step1_all** (783 sec elapsed time on a HP ElitBook 840 i5 notebook, with bzipped frames; 252 sec elapsed time with uncompressed frames)

As we expect anomalous signal, the Friedel mates will differ. So uncomment the line

FRIEDEL'S_LAW=FALSE

in XDS.INP. Then, run xds at the command line (or xds_par for the paralleled version).

Re-run scaling

Go to the directory xds_step2_CORRECT (60 sec elapsed time)

At the previous step, xds automatically figures out the Bravais symmetry, and picked up space group P422 as a representative. To check for extinctions (helices), just select space group 96 (P4(3)2(1)2) and enter refined cell parameters in XDS.INP

SPACE_GROUP_NUMBER=96 UNIT_CELL_CONSTANTS= 77.268 77.268 38.704 90.000 90.000 90.000

and select only the final scaling step of the processing (CORRECT)

JOB= CORRECT

Then run

xds

In CORRECT.LP, check for low intensity of reflections that should be absent (marked with "*") in the list above lines

AVERAGE INTENSITY FOR207 REFLECTIONS WHICH SHOULDBE SYSTEMATICALLY ABSENT IS0.2% OF MEAN INTENSITY

xscale and xdsconv

Go to the directory xds_step3_XSCALE-XDSCONV (5.6 sec elapsed time)

Optional: use XSCALE for final scaling, merging of several dataset. Create the XSCALE.INP input file with the following lines. Not necessary for a single dataset, apart to redefine boundaries of resolution shells.

OUTPUT_FILE=XSCALE.HKL INPUT_FILE=XDS_ASCII.HKL

and run

xscale
at the command line (alternatively: xscale_par).

Use XDSCONV to generate reflection files in CCP4 FP/DANO format (F, SigF, Dano, SigDano):

INPUT_FILE=XSCALE.HKL XDS_ASCII OUTPUT_FILE=temp_ccp4.hkl CCP4 FRIEDEL'S LAW=FALSE

and run

xdsconv at the command line.

XDSCONV generates the input file F2MTZ.INP needed by f2mtz (CCP4 package) for the final conversion to binary mtz format. To run the CCP4 programs f2mtz just type the command:

f2mtz HKLOUT temp_ccp4.mtz < F2MTZ.INP

Use XDSCONV again to generate reflection files in CCP4 F+/F- format (F, SigF, F+, SigF+, F-, SigF-):

INPUT_FILE=XSCALE.HKL XDS_ASCII OUTPUT_FILE=temp_ccp4_f.hkl CCP4_F FRIEDEL'S_LAW= FALSE GENERATE_FRACTION_OF_TEST_REFLECTIONS=0.05

and run again xdsconv at the command line

Then run CCP4 program f2mtz f2mtz HKLOUT temp_ccp4_f.mtz < F2MTZ.INP

Then, to run cad (to convert indices to the CCP4-asymmetric unit), cad HKLIN1 temp_ccp4.mtz HKLIN2 temp_ccp4_f.mtz HKLOUT Lyso-Gd_SAD.mtz <<EOF LABIN FILE 1 E1=FP E2=SIGFP E3=DANO E4=SIGDANO E5=ISYM LABIN FILE 2 E1=F(+) E2=SIGF(+) E3=F(-) E4=SIGF(-) E5=FreeRflag END EOF

Data reduction with XDS in graphic/automated mode

Graphic mode, using xdsGUI

In xds_xdsGUI Run xdsgui at the command line

₽			XE	SGUI 2017-	01-25 running	g in /home/da	ta/Oleron201	7/lyso-Gd_15N	/lay2013/xds_	dsGUI			~	^ 😣
Menu Hel	р													
Projects	Frame	XDS.INP	XYCORR	INIT	COLSPOT	IDXREF	DEFPIX	INTEGRATE	CORRECT	tools	statistics	XDSCONV	XSCALE	
Folder w	rith XDS c	onfigurati	ion and ou	tput files										
Default is th	e current dire	ectory. The tit	le bar of the XI	OSGUI winde	ow shows the	currently used	d folder.							
		Lo	ad recent proje	ect										
Choose o	r create new f	folder												
/home/data/(Oleron 2017/I	lyso-Gd 15M	av2013/vds v	deGLIT										

From there, all steps described above can be performed, starting with the edition of the XDS.INP parameter file

nu Heln			AL	56012017	-01-251011111	g in momerua	ta/Oteron20	1770950-00_151	way2013/Xus_	kuseoi			Ŷ
rojects	Frame	XDS.INP	XYCORR	INIT	COLSPOT	IDXREF	DEFPIX	INTEGRATE	CORRECT	tools	statistics	XDSCONV	XSCALE
					Save			Run	xds			stop xds	
			DATA COI	LECTIO	N STRATE	GY (XPLA	N) ====:						
RIEDEL TARTIN	I'S_LAW=	FALSE	00 s	TARTIN	G_FRAME=	1							
			SELECTIO	NOFD	ATTA TMAG	RS =====				_			
NAME_T	EMPLATE	OF_DAT	A_FRAMES=	/img	/e000_pr	efix_1_?	????.imo	a a					
DATA_R	ANGE=1	300	! Numbe	ers of	first an	d last d	ata imag	ge collec	ted				
SPOT R	ANGE=1	20	J !Numbe	and 1	ast data	image n	umber fo	ge for bai or finding	a spots				
									9 - 2				
			== INDEX1	NG PAR	AMETERS					50			
INDEX_	_ORIGIN=	:000											
======		:==== C)	RITERIA F	OR ACC	EPTING R	EFLECTIO	NS ====	==========		-			
VALUE_	RANGE_F	OR_TRUS	TED_DETEC	TOR_PI	XELS= 70	00 30000							
TRUSTE	D_REGIO	N= 0.0	1.41 ANGE=50 0										
INCHOL	E_RESOL	OTION_N	AUGE-30.0	0.0									
		= INTEG	RATION AN	ID PEAK	PROFILE	PARAMET	ERS ===:			0			
	(INTEGR	TOT (TOT)	CHANCE DE	AM ORT	ENTATION	CELL							
REFINE	- 5 000	(ATE) =DI	STANCE BE	ANI OIL	5111111 2011	CLLL							

and then running xds (the "Run XDS" button). Log files are available in the "XYCORR", "INIT", ... windows, with graphical display of the

statistics:



extra statistics are available runing xdsstat in the "statistics" window:



Access to xscale and xdsconv is provided through the corresponding windows.

Automated data reduction

A series of automated scripted layers are available for automated data reduction with xds, such as

<u>xdsme</u>

```
<u>xdsapp</u>
Just run from the top directory
xdsapp -a -cmd --fried=false
```

The dataset will be automatically detected and processed (parameters from the header of images). Alternatively, xdsapp is also a graphic interface for xds.

<u>Xia</u>

Quick SAD phasing with Phaser in ccp4i

Launch ccp4i and define a new project with lyso-Gd_SAD/ccp4_SAD as working directory

Select "Phaser SAD Pipeline" (button highlighted in blue in Figure 1) (~330 sec elapsed time). Then enter the following parameters:

- reflexion file: xds_step3_XSCALE-XDSCONV/Lyso-Gd_SAD.mtz (field highlighted in blue in Figure 2)

- sequence in fasta format directory ccp4_MR (field highlighted in green in Figure 2)

- heavy atom type: GD (field highlighted in red in Figure 2)

- wavelength: 1.7119 (field highlighted in orange in Figure 2)



Figure 1

🗣 💿 Maximum Likelihood Experimental Phasing Initial parameters from /home/data/15-05-13-lyso/lyso-Gd_S 📀	0
	Help
Job title [No title given]	
Mode for experimental phasing Single-wavelength anomalous dispersion (SAD) 💴	
Phaser SAD pipeline	
Run SHELXD — before Phaser	
🔳 Run Parrot (density modification) after Phaser 🔳 Run Buccaneer (model building) after Parrot	
Define data	
MTZ in Full path/home/data/15-05-13-lyso/lyso-Gd_SAD/xds_step3/Lyso-Gd_>rowse View	
Crystal unknown belonging to Project unknown	
Dataset name unknown140515	
F(+) F(+) SIGF(+) SIGF(+)	-
F(-) F(-) SIGF(-) SIGF(-)	
EPEED FreeBflag	-
Resolution 38.635 A to 1.648 A: Wavelength 1.7119	
Space group read from mtz file 'P 43 21 2' : I Enantiomorph choice Both	
Enter scattering from fluorescence scan (default is to calculate f' and f" from wavelength)	
Define atoms	
Atom sites run ShelxC/D -	
Find 4 heavy atoms of type GD	
Crystal contains cluster compound	
LLG-map completion on: all atom types	
Shelx parameters	1
Composition of the asymmetric unit	
Total scattering determined by components in asymmetric unit	
Component #1 protein 🔤 sequence file 💻 Number in asymmetric unit 1	
SEQ file Full path	
Edit list 📃 Define another component	
Accessory parameters	
Output control	IJ
Expert parameters	1
Run 🔤 Save or Restore 💻 Close	



 $=> \sim 80\%$ of residues built automatically

Upon completion of the job, and to analyze the log file, select the "Phaser_EP" job in the list (button highlighted in blue in Figure 3). Then, from the "View Files from Job", select "View Job Results (new style)" (button highlighted in red in Figure 3)



Figure 3

Experimental map, sub-structure of anomalous atoms and model can be displayed with Coot: Run

coot

at the command line, and load pdb files (button highlighted in blue in Figure 4) and mtz files (button highlighted in red in Figure 4) as listed below:





Sub structure of Gd atoms is in lyso-Gd_SAD/LysoGd_SAD_3.1.pdb Experimental map is in lyso-Gd_SAD/LysoGd_SAD_3.1.mtz and after automated density modification with parrot lyso-Gd_SAD/LysoGd_SAD_3_parrot.mtz The model built automatically (80% of the residues) is available in lyso-Gd_SAD/LysoGd_SAD_3_buccaneer.pdb

This model and the experimental map are good enough to start manual building.

Alternatively, run Refmac for a first refinement and manual rebuilt with LysoGd_SAD_3_buccaneer.pdb as pdb input file.

Quick SAD phasing in ccp4i2

From the directory ccp4i2_SAD, launch ccp4i2 at the command line

ccp4i2

Then Start a new crystallography project and define a project directory.



Run the SHELX automated pipeline

In the present section, the automated pipeline for SAD phasing is described. This pipeline is based on SHELX and Refmac5/Buccaneer.

From the ccp4i2 window, in job list, select Experimental phasing, and then the SHELX pipeline.

💠 🖡							CCP4-	7.0.036	Project Viewer: Gd	~ ^ 😣
File Edit	History Ui View in Coot Project direc Job/File	view in itilities	Projects	Help	() Help	Evaluation R=0.33 Rfree	CCP4- Clone job Clone job → ● → ● → ● → ● → ● → ● → ● → ●	7.0.036 Run Impor Integr X-ray Experi Bioinf Molec Model Refine Ligano Valida Expor Reflec Coord Devel	Project Viewer: Gd Run on server t merged data, sequences, alignments or coordinates ate X-ray images data reduction and analysis mental phasing prmatics including model preparation for Molecular Replacement utar Replacement building and Graphics ment Is tion and analysis tion and analysis tion data tools inate data tools oper tools	~ ~ 🛇
<						,			New job Cancel]

Then provide

- the sequence
 - file **193L.fasta**, to be renamed as .seq
 - directory lyso-Gd_15May2013/ccp4_MR
- the crystallography data
 - file temp_ccp4_f.mtz
 - directory lyso-Gd_15May2013/xds_step3_XSCALE-XDSCONV

File Edit	History Utilities Projects Help			~ ^ 0
Task menu	View in Cool View in CCP4mg Export MTZ Help	Bibliography	😺 🥵 🥳 Clone job Run Run on server	
Job list	Project directory		Job 2: Automated structure solution - SHELXC/D/E phasing and building The job is Pending	
Q	Job/File	Evaluation	Input Results Comments	
	• 2 SHELX		Input Data Important Options Advanced Options	
> ? •	▲ 1SHELX	R=0.33 Rfree	Job title SHELX Image: Shell of the sequence Image: Start pipeline with Substructure detection image: and end with Model building image: Start pipeline with Substructure detection image: and end with Model building image: Start pipeline with Substructure detection image: Start pipeline with Substructure atom: Se Number of substr. atoms in asymmetric unit: Anomalous data (Friedet pairs) Imput unmerged/merged SCA/XDS/SHELX format Image: Reflections Interp.ccp4_f: unknown140515 imported by image: Image: None image: Imput anomalous data #2 (MAD) Imput native observations (Crystal #2) Imput native observations (Crystal #2) Imput native observations image: Detection image: Detectine image: Detection image: Detection image: D	

and mention that the data were collected at the peak of fluorescence.

Then press **Run** to chain the following steps:

- Substructure determination (SHELXD)
- Phasing, Density Modification, first model (SHELXE)
- Refinement (Refmac5) / Model building (Buccaneer)

Results analysis

Substructure determination:

The substructure consists in 4 Gd atoms, 2 of them with >10 sigma density in the 2mfo-Dfc density map upon refinement, at the interface between 2 molecules, close to residues 62-63.



Model building and refinement:

- 3 cycles of refinement, with final FOM of 0.818, Rfree of 37.5% and Rfactor of 32.3%
- 114 residues built automatically.



Results displayed in ccp4i2 for the model building/refinement step.

The 2mfo-Dfc map, sub-model and refined model can be displayed with coot (press the **Manual COOT** button).



Partial model displayed in the 2mfo-Dfc contoured at 1sigma.Extra density is visible around the chelating HPDO3A for manual building.



Refined model displayed in the 2mfo-Dfc contoured at 1sigma.

Automated SAD phasing with Phaser

Instead of SHELXE, Phaser can be used for phasing, using the substructure established by SHELXD.

To do so, in the **Experimental phasing** menu, select **SAD phasing from heavy atom sites** – **Phaser**.

Provide the requested information (wavelength: 1.711888 Ang; MW: 14300).

File Edit History Utili	ies Projects:	Help										
2 2	3	R	?	427	Q)	*	*					
Task menu View in Coot V	/iew in CCP4mg	Export MTZ	Help	Bibliography	Clone job	Run Run	on server					
Job list Project director	Y					Job 3: SAL	D phasing from	n heavy atom si	ites - PHASER		The job is Pen	ding
Q Job/File				Evaluation		Input	Results	Comments				
• • 3 SAD phasi	ng from heavy at	om sites - PHAS	ER			Input dat	a Keywords					
>- ☆・ 				R=0.33 Rfree=0.	38	Job ti	tle SAD phasir	ng from heavy ato	m sites - PHASE	R		
						~	Use data from jo	b 1 Automated	l structure soluti	on - SHELXC/D	'E phasing aı 🗸	as input below
						Experim	iental data					
							Reflections	1 temp ccp4 1	f: unknown1405	15 imported by i	ob 1	
											1	
						Resolu	tion range				Wavelength	1.5418
						Known	sites, phasing mo	del, or phases				
						Provid	e sites, partial mo	del or phases as:	Partial heavy	atom substructu	re 🗸	
						1	Partial HA model	1 Model coord	linates			-
						Compos	ition					
						For est	imating asymmet	ric unit contents:	Provide an es	timate of the mo	lecular weight o	f protein and \checkmark
						Moleci	ılar weight (Da) o	of protein in the A	su			
						Molecu	ılar weight (Da) o	of nucleic acid				
						Cycles	of heavy atom mo	odel completion	1			
						Search 1	or pure anomalo	us scatterers)			
							Show list Ele	ments to search fo	or			
<					>							

SAD phasing, step by step, with ccp4i

Launch

ccp4i and define a new project with mosflm (the directory in which you have processed the data) as working directory

Step 1: Calculate the anomalous difference Patterson map and check the presence of peaks.

$$\mathbf{P}_{\rm H}(\vec{u}) = \sum_{\rm h,k,l} (|F_{\rm PH}(\vec{s})| - |F_{\rm PH}(-\vec{s})|)^2 \exp[-2i\pi \vec{u} \cdot \vec{s}]$$

Go to menu / sub-menus: Experimental Phasing / Heavy Atom Location / Generate Patterson Map

autor patterson map in CCP4 format If Run FFT to generate anomalous difference Patterson map in CCP4 format It List peaks to file It List peaks to file integration format If List peaks to file If Espace-no-TimeMachine/Phelma-2016/16-11-30-TP-Phelma-201 Browse Mew F1 F(+) SigmaF1 SIGF(+) F(-) F(-) SigmaF2 SIGF(-) Map TEMPORARY Lyso-Gd_SAD_patterson1.map Browse Mew Mew Peak coord Lyso-Id_SAD_patterson1.map Browse Mew Define Map Browse View Define Map Browse View Scale amplitudes for set 1 and set 2 Exclude reflections Exclude reflections Exclude reflections F Exclude reflections with difference between F1 and F2 > Exclude reflections - parameters for Set 1 and Sta1 Sta1 F less than n * sigmaF where n is 3.0 3.0 F A F F Sta1 F F Sta1 F F Sta1 F F Sta1 F Sta1 F Sta1 <th>lob title</th> <th>nattoreau</th> <th></th> <th></th> <th></th> <th></th> <th></th>	lob title	nattoreau					
MTZ in Full path //Espace-no-TimeMachine/Phelma-2016/16-11-30-TP-Phelma-201 Wew F1 F(*) SigmaF1 SIGF(*) F2 F(-) SigmaF2 SIGF(-) Map Erowse View Peak coord Iyso-1 Lyso-Gd_SAD_peaks1.pdb Browse View Define Map Scale amplitudes for set 1 and set 2 Exclude reflections Exclude reflections Exclude reflections F Exclude reflections ifference between F1 and F2 > Exclude reflections - parameters for set 1 and set 2 F less than n * sigmaF where n is 3.0 3.0 3.0 Scale F F absolute value less than Image for set 1 Image for set 1 Image for set 1 Image for set 1 Set 2	■ Run ■ List _ Plot	FFT to generate peaks to file default Harker	anomalous difference	Patterson -	f peaks in map	format	
F1 F(+) SigmaF1 SIGF(+) F2 F(-) SigmaF2 SIGF(-) Map TEMPORARY Lyso-Gd_SAD_patterson1.map Browse View Peak coord lyso-1 Lyso-Gd_SAD_peaks1.pdb Browse View Define Map Scale amplitudes for set 1 and set 2 Exclude Reflections Exclude Reflections Exclude reflections Exclude reflections Exclude reflections Files than n * sigmaF where n is 3.0 3.0 3.0 F absolute value less than	MTZ in	Full path 🧹	Espace-no-Time Machine	e/Phelma-2016/16-11	-30-TP-Phelma-2	01 Browse	View
F2 F(-) SigmaF2 SIGF(-) Map TEMPORARY Lyso-Gd_SAD_patterson1.map Browse View Peak coord lyso-1 Lyso-Gd_SAD_peaks1.pdb Browse View Define Map Scale amplitudes for set 1 and set 2 Exclude for set 1 and set 2 Extent * asymmetric unit \checkmark or range x y z Exclude reflections Exclude reflections Exclude reflections * Set 1 and set 2 Set 1 Set 2 Set 2 F less than n * sigmaF where n is 3.0 3.0 Sa.0	FI		F(+)	- SigmaF1	SIGF(+	+)	
Map TEMPORARY Lyso-Gd_SAD_patterson1.map Browse View Peak coord lyso-1 Lyso-Gd_SAD_peaks1.pdb Browse View Define Map Scale amplitudes for set 1 and set 2 Image: Constraint of the set	FZ		F(-)	SigmaF2	SIGF(-	.)	_
Peak coord Iyso-Gd_SAD_peaks1.pdb Browse View Define Map Scale amplitudes for set 1 and set 2 Image: Scale amplitudes for set 1 Image: Scale amplitudes for set 2 Image: Scale amplitude for for set 1 Image: Scale amplitude for for set 2 Image: Scale amplitude for for set 2 Image: Scale amplitude for for set 2 Image: Scale amplitude for for for set 1 Image: Scale amplitude for for for set 1 Image: Scale amplitude for for for for for for for for for set 1 Image: Scale amplitude for	Мар ТЕ	MPORARY - Ly	so-Gd_SAD_patterson1.	map		Browse	View
Define Map Image: Constraint of the set of the	Peak co	ord lyso-1	Lyso-Gd_SAD_peak	cs1.pdb		Browse	View
Exclude reflections with difference between F1 and F2 > Exclude reflections - parameters for set 1 and set 2 F less than n * sigmaF where n is 3.0 F absolute value less than F absolute value greater than	Scale a	map mplitudes for set 1	and set 2				-
Resolution less than 34.511 A or greater than 1.900 A	Scale a Extent <i>Exclude</i>	map mplitudes for set 1	and set 2 international and set 3 international and se	У	z		
Infrequently Used Patterson Options	Scale a Extent Exclude Exclude Fles Fles Fab Rese	mpitudes for set 1 asymmetric un <i>Reflections</i> ude reflections with <i>reflections - paran</i> is than n * sigmaF visolute value less the solute value greate polution less than 3	and set 2 and set 2 and or range x and difference between F1 seters for where n is tan r than 4.511 A or greater th	and F2 >	z	set 2 3.0	*
Peak Search Details	Scale a Extent Exclude Exclude Exclude Fles Fles Reso Reso	mplitudes for set 1 asymmetric un <i>Reflections</i> ude reflections with <i>reflections - paran</i> is than n * sigmaF to solute value less th solute value greate plution less than 3 <i>ently Used Pattersol</i>	and set 2 and set 3 and 3	and F2 >	z z	set 2 3.0	F

select "anomalous difference Patterson"

MTZ in: ctruncate_A5_1__00002-unique.mtz or Lyso-Gd_SAD.mtz

check that F1 stands for F+ and F2 for F-

In "Exclude Reflections" you can chose the resolution range, indicate a threshold in the ratio signal-to-noise, etc.

The list of peaks in the anomalous diffrence Patterson is stored in Lyso-Gd_SAD_peaks1.pdb (Peak coord).

Edit this file with any text editor:

CRYST1	77.	170	77.	.17	0 38.3	360 90.0	0 90.00	90.00 1	2 4/m 1	m m
SCALE1		0.01	2958	- 0	.000000	-0.00000	0	-0.00000		
SCALE2	-	0.00	0000	0	.012958	-0.00000	0	0.00000		
SCALE3		0.00	0000	- 0	.000000	0.02606	9	-0.00000		
ATOM	1	OW	WAT	Х	1	0.000	0.000	0.00013	37.191	37.19
ATOM	17	OW	WAT	Х	2	25.077	0.000	14.622	9.35	9.35
ATOM	13	OW	WAT	Х	3	38.585	35.227	10.790	9.16	9.16
ATOM	14	OW	WAT	Х	4	38.585	10.136	12.893	6.86	6.86
ATOM	12	OW	WAT	Х	5	15.234	13.308	9.606	6.63	6.63
ATOM	12	OW	WAT	Х	6	13.308	15.234	9.606	6.63	6.63
ATOM	6	OW	WAT	Х	7	2.715	3.095	4.514	6.40	6.40
ATOM	18	OW	WAT	Х	8	25.208	25.208	16.120	6.22	6.22
ATOM	4	OW	WAT	Х	9	23.342	23.342	3.180	5.87	5.87
ATOM	9	OW	WAT	Х	10	38.585	36.598	6.570	5.72	5.72
ATOM	20	OW	WAT	Х	11	28.584	1.994	19.180	5.71	5.71
ATOM	7	OW	WAT	Х	12	12.806	9.762	5.186	5.33	5.33
ATOM	7	OW	WAT	Х	13	9.762	12.806	5.186	5.33	5.33
ATOM	16	OW	WAT	Х	14	11.822	15.680	14.051	5.29	5.29
ATOM	16	OW	WAT	Х	15	15.680	11.822	14.051	5.29	5.29
ATOM	8	OW	WAT	Х	16	29.094	29.094	5.656	5.27	5.27
ATOM	21	OW	WAT	Х	17	21.835	3.112	19.180	5.26	5.26
ATOM	3	OW	WAT	Х	18	26.280	26.280	1.214	5.25	5.25
ATOM	11	OW	WAT	Х	19	9.264	12.258	9.493	4.64	4.64
ATOM	11	OW	WAT	Х	20	12.258	9.264	9.493	4.64	4.64
ATOM	5	OW	WAT	Х	21	38.585	16.886	3.847	4.64	4.64
ATOM	10	OW	WAT	Х	22	13.464	36.192	8.329	4.62	4.62
ATOM	10	OW	WAT	Х	23	36.192	13.464	8.329	4.62	4.62
ATOM	19	OW	WAT	Х	24	22.978	28.723	17.844	4.23	4.23
ATOM	19	OW	WAT	Х	25	28.723	22.978	17.844	4.23	4.23
ATOM	15	OW	WAT	Х	26	26.327	26.327	13.383	3.52	3.52
ATOM	2	OW	WAT	Х	27	15.437	38.585	0.000	3.18	3.18
END										

You should have peaks above 4 sigmas.

<u>Step 2</u>: determine the position of Gd atoms using Shelxd program. Basically calculates the coordinates of Gd atoms from the anomalous difference Patterson map.

Go to menu / sub-menus: Experimental Phasing / Heavy Atom Location / ShelxS - Heavy Atom Search

NOTE This tasi	k uses th	e Shelx j	program whic	ch is not distribu	ted with CCP	4 - click Help f	or details	He
Job title Shelx	S 2							
Try to find hea	vy atom:	s by P	atterson sea	rch 🔤				
nput format is	мт	Z file	_ using	anomalous	data			
MTZ in Full p	ath <	/Espa	ce-no-Time	Machine/Phelma-	2016/16-11-3	0-TP-Phelma	Browse	View
FP		FP		Sigma		SIGFP		_
DP		DANG		SIGDP		SIGDANO		-
Cell Parameter.	5	\sim						
Space group F	43 21 2		and lattice ty	rpe primitiv		Wavelength 1.	54178	
Set cell a 77.1	7 b	77.17	c 38.36	alpha 90.0	beta 90.	0 gamma	90.0	
Cell ESDs a 0.	001	b 0.001	c 0.001	alpha 0.0	beta 0	.0 gamn	na 0.0	-
Search for 4) a	toms of	Gd	_	,			
Exclude Reflec	tions							
Shelx Patterso	n Search	Paramet	ters					
Try 10	superpos	sition ve	ctors					
Enter 'known'	vectors:							
					Ed	it list 📮	Add Ve	ector
_			_					
	Run		-	Save o	Restore	-	Close	

MTZ in: ctruncate_A5_1__00002-unique.mtz or Lyso-Gd_SAD.mtz

check that FP stands for FP and DP for DANO

Guess how many Gd you could have bound to one lysozyme molecule (You can slightly overestimate), and specify that you are looking for Gd.

10 superpositions vectors (number of Patterson peaks that will analysed as Gd-Gd interatomic vectors, you can try from 4 to about 10, not critical in normal cases)

Look at the log file (old style): Two Gd found:

Patterson vector superposition minimum function for E ShelxS 2
Patt. sup. on vector 1 0.1222 0.1606 0.2497 Height 6. Length 18.28
Maximum = 8.96, minimum = -98.87 highest memory used = 15501 /673654
0.1 seconds CPU time
500 Superposition peaks employed, maximum height 75.1 and minimum height 19.1 on atomic number scale
Heavy-Atom Location for E ShelxS 2
9562 reflections used for structure factor sums
Solution 1 CFOM = 1.35 PATFOM = 99.9 Corr. Coeff. = 11.6 SYMFOM = 99.9
Shift to be added to superposition coordinates: 0.4209 0.0618 0.0471
Name At.No. x y z s.o.f. Minimum distances / PATSMF (self first)
GD1 68.4 0.1857 0.5136 0.7047 1.0000 22.47 3.0

.27
2.7
4800

Step 2 (alternative)

Use already known phases to locate Gd atoms.

advantages: more precise localization of anomalous scatterers (especially when you have several atoms to locate), no need for very accurate phases

drawbacks: need for phases (either from molecular replacement or from other heavy atoms derivatives)

Here, we can use the phases from the molecular replacement solution done with the bovine lysozyme model (lyso-gd_3.1.mtz contain phases information from the molecular replacement solution)

Go to menu / sub-menus: Map & mask Utilities / Run fft - Create Map

lob title fft		
Run FFT — to generate anomalous — map		
Dutput map in CCP4 - format to cover all atoms in PDB file -		
Do peak search of map		
Plot map sections with no coordinates -		
ITZ in Full path a-2016/16-11-30-TP-Phelma-2016/andrea/ccp4/lyso-gd_3.1.m	tz Browse	View
PHI PHIC - Weight FO	M	
DANO DANO – Sigma SIGD.	ANO	_
PDB file Full path a-2016/16-11-30-TP-Phelma-2016/andrea/ccp4/lyso-gd_3.1.	b Browse	View
Aap out Iyso-1 — Iyso-gd_3.1.map	Browse	View
Peak coords lyso-1 _ <mark>lyso-gd-fft.pdb</mark>	Browse	View
Define Map		
Exclude set of FreeR reflections defined by MTZ label FreeR flag	-	
Extending Man		
Border in Angstrom around atoms 5		_
Exclude Reflections		
Peak Search Details		
Coordinates of peaks will be saved in coordinate file & plotted		
Also output peaks in fractional coordinates		
Search for peaks greater than 3.0 sigma 🔟 Search for negative peaks		
lumber of peaks output to list 50		
PDB ATOM card with chain X residue WAT atom OW Bfactor occu	pancy 1.0	

PHIC are the phases from molecular replacements

FOM is the Figure of Merit, i.e. a value between 0 and 1 that quantify the quality of the phase estimates (used here has a weight)

DANO is the anomalous difference F+ - F-, associated with the sigma (error) SIGANO The pdb file is only used to determine where the map is calculated (around the model, easier for interpretation, but not mandatory, the map can be calculated on the asymmetric unit if no model is known)

The map calculated is :

$$\rho_{\mathrm{H}}(\vec{r}) = \sum_{\mathrm{h,k,l}} \mathrm{w.}(|\mathrm{F}_{\mathrm{PH}}(\vec{s})| - |\mathrm{F}_{\mathrm{PH}}(-\vec{s})|) \exp[i(\varphi_{\mathrm{calc}} - \frac{\pi}{2})] \exp[-2i\pi\vec{r}\cdot\vec{s}]$$

with $(|\mathrm{F}_{\mathrm{PH}}(\vec{s})| - |\mathrm{F}_{\mathrm{PH}}(-\vec{s})|) = \mathrm{DANO}; \mathrm{w} = \mathrm{FOM}; \varphi_{\mathrm{calc}} = \mathrm{PHIC}$

the shift of $-\pi/2$ come from the fact that we are looking at the f' contribution (the one that breaks Friedel's law), which contributes with a phase shift of $\pi/2$ (if' in the structure factor formula)

The highest peaks of this map should correspond to Gd atoms. The file lyso-1_7_peaks.ha provides the peaks in fractional coordinates:

GRID 12	20 120	64					
CELL	77.17	00 77.1	.700 38	3.3600	90.0000 90	.0000	90.0000
ATOM1	Ano	-0.1838	0.5129	0.0428	29.37	0.0 BFA	AC 20.0
ATOM2	Ano	-0.3162	0.0129	-0.2928	29.37	0.0 BFA	AC 20.0
ATOM3	Ano	-0.3162	0.0129	0.7072	29.37	0.0 BFA	AC 20.0
ATOM4	Ano	-0.0129	0.3162	-0.2072	29.37	0.0 BFA	AC 20.0
ATOM5	Ano	-0.0129	0.3162	0.7928	29.37	0.0 BFA	AC 20.0
АТОМ6	Ano	0.1838	0.4871	0.5428	29.37	0.0 BFA	AC 20.0
ATOM7	Ano	-0.1403	0.4817	-0.0746	27.04	0.0 BFA	AC 20.0
ATOM8	Ano	0.1403	0.5183	0.4254	27.04	0.0 BFA	AC 20.0
ATOM9	Ano	0.0183	0.3597	-0.3246	27.04	0.0 BFA	AC 20.0
ATOM10	Ano	0.0183	0.3597	0.6754	27.04	0.0 BFA	AC 20.0
ATOM11	Ano	-0.1421	0.4820	0.8906	5.75	0.0 BFA	AC 20.0
ATOM12	Ano	-0.1315	-0.0038	-0.2092	4.02	0.0 BFA	AC 20.0
ATOM13	Ano	-0.1315	-0.0038	0.7908	4.02	0.0 BFA	AC 20.0
ATOM14	Ano	0.1315	0.0038	0.2908	4.02	0.0 BFA	AC 20.0
ATOM15	Ano	0.0038	0.1315	-0.2908	4.02	0.0 BFA	AC 20.0
ATOM16	Ano	0.0038	0.1315	0.7092	4.02	0.0 BFA	AC 20.0
ATOM17	Ano	0.2835	0.5146	0.2851	3.88	0.0 BFA	AC 20.0
ATOM18	Ano	0.0146	0.2165	0.5351	3.88	0.0 BFA	AC 20.0

Two independent sites (the first 6 ones are related by crystallographic symmetry and have the same heights (29.37), same for the next 4 ones (heights of 27.04)

the file lyso-gd-fft.pdb provides the peaks in cartesian coordinates:

CRYST1	77.	170	77.	.17	0 38.3	360 90.0	0 90.00	90.00	P 43 2	21 2	
SCALE1		0.01	2958	- C	.000000	-0.00000	0	-0.0000	C		
SCALE2	-	0.00	0000	С	.012958	-0.00000	0	0.0000	C		
SCALE3		0.00	0000	- C	.000000	0.02606	9	-0.0000	C		
ATOM	5	OW	WAT	Х	1	-14.183	39.579	1.643	29.37	29.37	0
ATOM	5	OW	WAT	Х	2	-24.402	0.994	-11.233	29.37	29.37	0
ATOM	5	OW	WAT	Х	3	-24.402	0.994	27.127	29.37	29.37	0
ATOM	5	OW	WAT	Х	4	-0.994	24.402	-7.947	29.37	29.37	0
ATOM	5	OW	WAT	Х	5	-0.994	24.402	30.413	29.37	29.37	0
ATOM	5	OW	WAT	Х	6	14.183	37.591	20.823	29.37	29.37	0
ATOM	3	OW	WAT	Х	7	-10.831	37.174	-2.860	27.04	27.04	0
ATOM	3	OW	WAT	Х	8	10.831	39.996	16.320	27.04	27.04	0
ATOM	3	OW	WAT	Х	9	1.411	27.754	-12.450	27.04	27.04	0
ATOM	3	OW	WAT	Х	10	1.412	27.754	25.910	27.04	27.04	0
ATOM	23	OW	WAT	Х	11	-10.962	37.199	34.164	5.75	5.75	0
ATOM	6	OW	WAT	Х	12	-10.151	-0.292	-8.025	4.02	4.02	0
ATOM	6	OW	WAT	Х	13	-10.151	-0.292	30.335	4.02	4.02	0
ATOM	6	OW	WAT	Х	14	10.151	0.292	11.155	4.02	4.02	0
ATOM	6	OW	WAT	Х	15	0.292	10.151	-11.155	4.02	4.02	0
ATOM	6	OW	WAT	Х	16	0.292	10.151	27.205	4.02	4.02	0
ATOM	13	OW	WAT	Х	17	21.879	39.714	10.936	3.88	3.88	0
АТОМ	13	OW	WAT	Х	18	1.129	16.706	20.526	3.88	3.88	0

In order to avoid symmetry related peaks, you can calculated the map on the asymmetric unit:

			не
Job title fft_au			
Run FFT	- to generate anomalous - map		
Output map in CCI	P4 - format to cover asymmetric unit		
Do peak search	of map		
🔄 Plot map sectio	ons with no coordinates		
MTZ in Full noth	2.2016/16.11.30.TP. Phalma.2016/androa/con/Alvea.ord.3.1.mtz	Provine	Manu
	PUIC Weight FOM	browse	VIEW
	DANO SIGDANO	$) \rightarrow$	
Man out Iven-1	lives ad 31-au man	Provine	Manu
map our iyso-i	- iyso-gu_5n-aumap	browse	VIEW
Peak coords ly	so-1lyso-gd-fft-au.pdb	Browse	View
Define Map			
Exclude set of I	FreeR reflections defined by MTZ label FreeR_flag		
Exclude Reflection	5		
Peak Search Detail	5		
Coordinates of pea	ks will be saved in coordinate file & plotted		
🖬 Also output pea	ks in fractional coordinates		
Search for peaks g	reater than 3.0 sigma 🔄 Search for negative peaks		
Number of peaks o	autput to list 50		
PDB ATOM card w	ith chain X residue GD atom GD Bfactor ccupancy	1.0	
	Antiona		

and the files lyso-1_8_peaks.ha lyso-gd-fft-au.pdb only contains two peaks much higher than the rest (which is basically noise):

GRID 12	20 120	64											
CELL	77.170)0 77.	1700	38.	3600	90.	0000	90	.0000	90.	0000		
ATOM1	Ano	0.8162	0.51	29	0.0428		30.	70	0.0	BFAC	20.0		
ATOM2	Ano	0.4817	0.85	97	0.0746		28.	26	0.0	BFAC	20.0		
АТОМЗ	Ano	0.4962	0.63	15	0.0408		4.	20	0.0	BFAC	20.0		
ATOM4	Ano	0.9854	0.78	35	0.0351		4.	05	0.0	BFAC	20.0		
ATOM5	Ano	0.6936	0.31	70	0.0308		4.	01	0.0	BFAC	20.0		
АТОМб	Ano	0.2161	0.34	38	0.0380		3.	96	0.0	BFAC	20.0		
CRYST1	77.1	L70 77	170	38.	360 90	0.00	90.0	0 9	90.00	P 43	21 2		
SCALE1	(0.012958	3 -0.00	0000	-0.000	0000)	- 0 .	.0000	0			
SCALE2	- (0.00000	0.01	2958	-0.000	0000)	0.	.0000	0			
SCALE3	(0.00000	-0.00	0000	0.026	5069		- 0 .	.0000	0			
ATOM	16	GD GI) X 1		62.98	37	39.579	-	1.643	30.70	30.7	C	0
ATOM	23	GD GI) X 2		37.17	74	66.339	2	2.860	28.26	28.2	5	0
ATOM	17	GD GI) X 3		38.29	93	48.736	-	1.565	4.20	4.2	C	0
ATOM	11	GD GI) X 4		76.04	11	60.464	-	1.346	4.05	4.0	5	0
ATOM	9	GD GI) X 5		53.52	23	24.464	-	1.182	4.01	4.0	1	0
ATOM	10	GD GI) Х б		16.67	78	26.533	-	1.457	3.96	3.9	5	0

Here, the peaks of the map gives you directly the coordinates of the Gd atoms.

Further analysis should show that these two peaks match the two peaks found by difference Patterson map, if you take into account the possible symmetries.

<u>Step 3:</u> calculating the experimental phases using the anomalous signal and the two gadolinium sites determined previously.

A) the lyso-gd-fft-au.pdb needs some editing to be interpreted properly (it will be saved in lyso-gd-fft-aus.pdb):
just keep the two main sites
put the occupancy to 1.0 (it will be refined anyway afterwards)
replace o by GD in the last column (otherwise the program will believe you have oxygen instead of Gd, and will fail)

 CRYST1
 77.170
 77.170
 38.360
 90.00
 90.00
 P 43 21 2

 SCALE1
 0.012958
 -0.000000
 -0.000000
 -0.00000

 SCALE2
 -0.000000
 0.012958
 -0.000000
 0.00000

 SCALE3
 0.000000
 -0.000000
 0.026069
 -0.00000

 ATOM
 16
 GD
 GD X
 1
 62.987
 39.579
 1.643
 1.00
 30.70
 GD

 ATOM
 23
 GD
 GD X
 2
 37.174
 66.339
 2.860
 1.00
 28.26
 GD

B)

Go to menu / sub-menus: Experimental Phasing / Automated Search & Phasing / Phaser SAD pipeline:

Pay attention to deselect run SHELXD before Phaser (you have already found where the GD atoms are)

some files are in the ccp4 directory and not in the mosflm one: just be sure where they are in you session.

Here you need F+ and F- and not DANO

The main 3 steps of the program will be:

- 1) Using Phaser to calculate phases from the data and positions of Gd atoms, using the principles described in the lectures. Positions, B-factor and occupancy of Gd atoms will be refined. Other sites will be searched and refined. (check log file)
- 2) Using Parrot for density modification procedure (solvent flatening, etc ...): should improve the quality of the phases
- 3) Using Buccaneer to build automatically a model in the resulting electron density map, using the primary sequence information provided.

Joh title Ev		He
oon ade jes	perimental Phasing	
Mode for ex	perimental phasing Single-wavelength anomalous dispersion (SAD) -	
Phaser SAL	pipeline	
Run S	HELXD before Phaser	
Run Pa	rrot (density modification) after Phaser 🔳 Run Buccaneer (model building) after Parrot	
Define data		Π.
MTZ in F	ull path 🛁 6/16-11-30-TP-Phelma-2016/andrea/ccp4/Lyso-Gd_SAD.mtz Browse Vie	w
Crystal un	known belonging to Project unknown	
Dataset na	ne unknown061216	
F(+)	F(+)	
F(-)	F(-)	
FP	FP SIGFP SIGFP	
-REER	an 24 511 A to 1 900 At Wavelength 1 712445	
	a mod from min file (D 42 21 2) + Transformenth choice Deth	
space grou	stread from intz life P 45 21 2 ; in Enandomorph choice Both	
_ Enter sc	attering from huorescence scan (default is to calculate f and f from wavelength)	
Define aton	8	-
Atom sites	in PDB file	
PDB in	yso-1 Jyso-gd-fft-aus.pdb Browse Vie	W
Change :	scattering type to	
Crystal	contains cluster compound	
LLG-ma	ap completion on: all atom types 💴	
Compositio	n of the asymmetric unit	
Fotal scatte	ering determined by components in asymmetric unit	
	al protain cogrange file Number in sourcestain unit 1	
Component	*i protein	
Component SEQ file	iuli path Ima-2016/16-11-30-TP-Phelma(2016/andrea/ccp4/2vb1.seq Browse Me	w
Component SEQ file	Ima-2016/16-11-30-TP-Phelma 2016/andrea/ccp4/2vb1.seq Browse Vie Edit list _ Define another component	w nt
Component SEQ file Accessory ;	Full path Ima-2016/16-11-30-TP-Phelma@016/andrea/ccp4/2vb1.seq Rrowse Me Edit list Define another component parameters	nt
Component SEQ file Accessory ; Output com	Full path Ima-2016/16-11-30-TP-Phelma@016/andrea/ccp4/2vb1.seq Mercese Merces	w nt ⊐

the experimental phases are in lyso-1_10_parrot.mtz. The model automatically constructed by buccaneer is in lyso-1_10_buccaneer.pdb

Files with 'hand' are also created in cases where the correct hand is not known. Here, the space group was set properly, since it was already known.

Step 4: Use coot to look at maps and models

open coot

open the molecular replacement solution with bovine lysozyme in ccp4 directory: lyso-gd_3.1.pdb menu File / Open coordinates

open molecular replacement maps in ccp4 directory: lyso-gd_3.1.mtz menu File / Auto Open MTZ

open the automatically built model in mosflm directory: lyso-1_10_buccaneer.pdb menu File / Open coordinates

Open experimental phases map in mosflm directory: lyso-1_10_parrot.mtz menu File / Open MTZ, mmCIF, fcf or phs... and chose parrot.F_phi.F as amplitudes parrot.F_phi.phi as phases adjust the contour to 1 sigma

Compare the different maps. Focus on regions were the sequence of bovine and Hen lysozyme are different.

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 QIN Sbjct: 1 KVFERCELARTLKKLGLDGYKGVSLANWLCLTKWESSYNTKATNYNPSSESTDYGIFQIN 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



MR

under construction

Complete model

Partial/remote model

Refinement / construction

under construction

Complete model

Partial/remote model