



# XChem Data Processing XCE & PanDDA

## 2024

---



# Working directory



**You will have a proposal number starting with lb, e.g.:**

- lb13385
- **For each target/screen you will have a visit number, e.g.:**
  - lb13385-1
- **You will end up with visits assigned to both:**
  - Lab34: labxchem
  - The beamline: I04-1
- **For data analysis you should be working in the processing subdirectory of your labxchem visit**

# Working directory structure

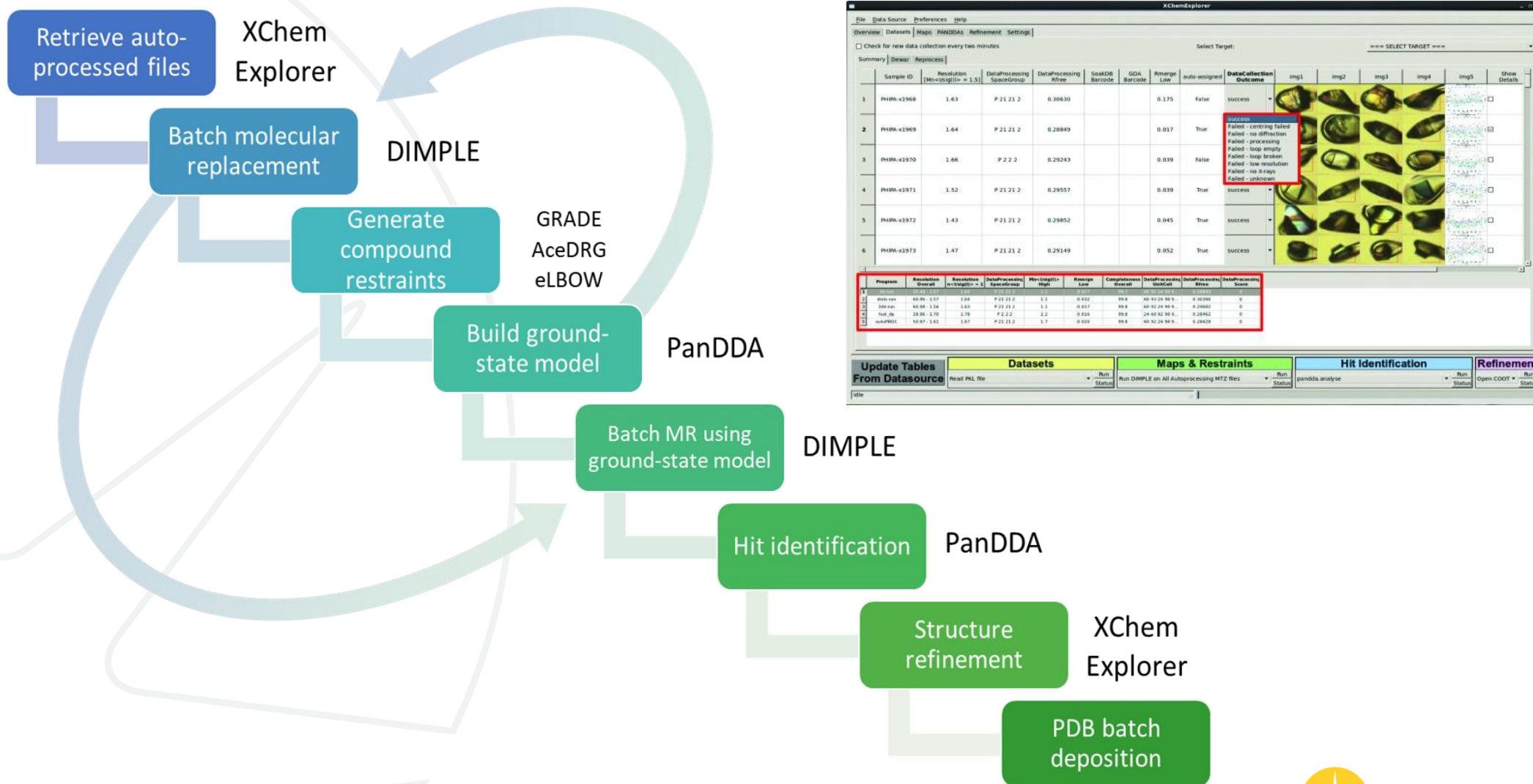
- analysis • Data collection results and pandda analysis
- beamline • Links to beamline visit directories (*obsolete – do not use*)
- database • SQLite datafile (and backups)
- group\_deposition • Files for PDB deposition
- html • Directory for lab work (soakDB, echo, shifter)
- lab36 • pdb of a good reference model
- reference
- tmp
- prepareVisit.log
- xce.log

# Useful linux commands



- Setup useful commands (**do this first**):
  - `cd /dls/labxchem/data/proposal/visit/processing/`
  - `source /dls/science/groups/i04-1/software/XChem/xchempaths.sh`
- `xchempaths.sh` will set paths for these commands:
  - `preparevisit` - to create the subfolders needed for XChem
  - `tserver` - to launch a windows remote desktop from linux
  - `xce` - to launch XChemExplorer
    - **Needs to be run under the 'processing' folder**
  - `csv2ispyb` - to automatically load the data collection information in iSPyB
    - **Needs to be run under the 'processing' folder**
- Checking the status of jobs on the cluster (type into terminal):
  - “`ssh wilson`” – connect to the DLS Slurm Cluster; must be done first
  - “`sacct`” – display jobs
  - “`watch sq.sh -u yourfedid -nf`” – watch jobs

# Data Analysis Workflow



XChem Explorer

Overview | Datasets | Maps | PANDDAs | Refinement | Settings

Summary | Details | Reprocess

Sample ID	Resolution [Max-avg(1) = 1.51]	DataProcessing SpaceGroup	DataProcessing Rfree	ScaleDB Barcode	GDA Barcode	Image Low	Auto-assigned	DataCollection Outcome	img1	img2	img3	img4	img5	Show Details
1	PH00-43968	1.63	P 21 21 2	0.30630		0.175	False	success						<input type="checkbox"/>
2	PH00-43969	1.64	P 21 21 2	0.28849		0.017	True	success						<input type="checkbox"/>
3	PH00-43970	1.66	P 2 2 2	0.29243		0.039	False	Failed - centring failed Failed - no diffraction Failed - processing Failed - loop empty Failed - low resolution Failed - no X-rays Failed - unknown						<input type="checkbox"/>
4	PH00-43971	1.52	P 21 21 2	0.29557		0.039	True	success						<input type="checkbox"/>
5	PH00-43972	1.43	P 21 21 2	0.29852		0.045	True	success						<input type="checkbox"/>
6	PH00-43973	1.47	P 21 21 2	0.29149		0.052	True	success						<input type="checkbox"/>

Program	Resolution Overall	Resolution (CryoEM)	DataProcessing SpaceGroup	Max-avg(1)	Rmerge Low	Completeness Overall	DataProcessing SpaceGroup	DataProcessing Rfree	DataProcessing Score
1	1.63	1.63	P 21 21 2	1.1	0.037	99.8	P 21 21 2	0.30630	0.00000
2	1.64	1.64	P 21 21 2	1.1	0.017	99.8	P 21 21 2	0.28849	0.00000
3	1.66	1.66	P 2 2 2	1.1	0.039	99.8	P 2 2 2	0.29243	0.00000
4	1.52	1.52	P 21 21 2	1.1	0.039	99.8	P 21 21 2	0.29557	0.00000
5	1.43	1.43	P 21 21 2	1.1	0.045	99.8	P 21 21 2	0.29852	0.00000
6	1.47	1.47	P 21 21 2	1.1	0.052	99.8	P 21 21 2	0.29149	0.00000

Update Tables From DataSource | Datasets | Maps & Restraints | Hit Identification | Refinement

# XChem jargon and experimental philosophies



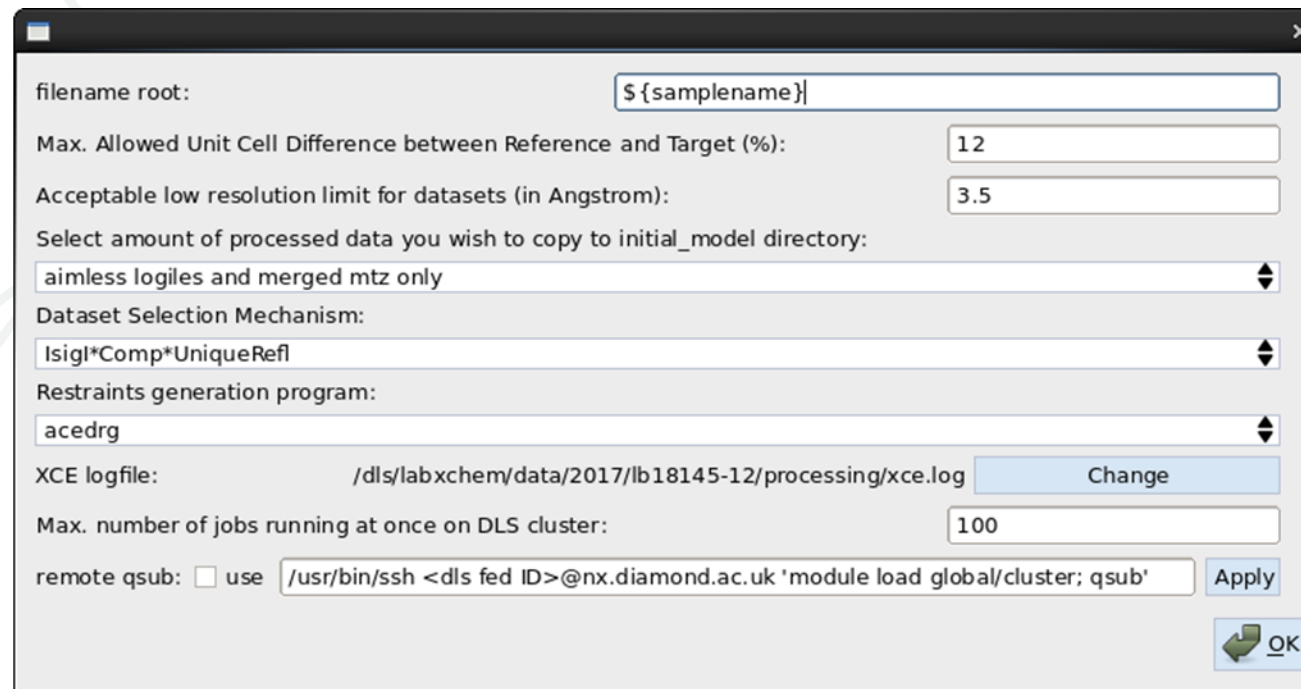
- Reference model = Dimple/MR model = PanDDA input model = **ground-state model**
- PanDDA model = ligand model = **bound-state model**
- Ensemble model = ground-state model + bound-state model
  
- The ensemble model is usually the one refined, particularly with low occupancy fragments
- The bound-state model will be the one you will update in the XCE refinement Coot window and the one which will be deposited on the PDB

# XCE Preferences

Dimple reference model  
selection criteria →

Datasets tab  
options }

Restrains generation  
program options →



filename root:

Max. Allowed Unit Cell Difference between Reference and Target (%):

Acceptable low resolution limit for datasets (in Angstrom):

Select amount of processed data you wish to copy to initial\_model directory:

Dataset Selection Mechanism:

Restrains generation program:

XCE logfile:

Max. number of jobs running at once on DLS cluster:

remote qsub:  use

# XCE Settings



File Datasource Preferences Deposition Proasis Help Labels

Overview Datasets Maps PANDDAs Refinement Deposition Settings

**Project Directory: - REQUIRED -**  
/dls/labxchem/data/2019/lb24544-1/processing/analysis/model\_building

**Reference Structure Directory: - OPTIONAL -**  
/dls/labxchem/data/2019/lb24544-1/processing/reference

**Data Source: - REQUIRED -**  
/dls/labxchem/data/2019/lb24544-1/processing/database/soakDBDataFile.sqlite

**Data Collection Directory: (e.g. /dls/i04-1/data/2017/lb18145-70) -**  
/dls/i04-1/data/2019/lb24544-3   
 Read Agamemnon data structure

**CCP4\_SCR Directory: - OPTIONAL -**  
/dls/labxchem/data/2019/lb24544-1/processing/tmp

**PANDDAs directory: - OPTIONAL -**  
/dls/labxchem/data/2019/lb24544-1/processing/analysis/panddas

**HTML export directory: - OPTIONAL -**  
/dls/labxchem/data/2019/lb24544-1/processing

**Group deposition directory: - OPTIONAL -**  
/dls/labxchem/data/2019/lb24544-1/processing/group\_deposition

Paths should look like this.

If not, this is because you haven't opened XCE in your processing directory!

Project directory is: **/analysis/model\_building**



**For UDC visits, select "Read Agamemnon data structure" before setting directory**

Manually set the data collection directory (/dls/i04-1/...)

Now, XCE can link your SoakDB data to the x-ray diffraction data



# Running jobs on SLURM cluster



Whenever you launch a group of jobs on the SLURM cluster, you will need to provide your FedID password for authentication.

# Data source tab: Overview of your experiments



Sample ID	Compound ID	Smiles	Visit	Resolution [Mn<math>\langle\sigma\sigma\rangle\langle\sigma\sigma\rangle <= 1.5]	Refinement Rfree	Data Collection Date	Puck	PuckPosition	Ligand Confidence
1	NUDT21A-x0060		lb18145-14	3.22	0.26199	2017-06-28 12:18:37	DL5593	1	None
2	NUDT21A-x0061		lb18145-14			2017-06-28 12:20:26	DL5593	2	
3	NUDT21A-x0062		lb18145-14			2017-06-28 12:22:52	DL5593	3	
4	NUDT21A-x0063		lb18145-14	3.88		2017-06-28 12:24:04	DL5593	4	
5	NUDT21A-x0064		lb18145-14	n/a	0.31977	2017-06-28 12:26:58	DL5593	5	None
6	NUDT21A-x0065		lb18145-14			2017-06-28 12:28:10	DL5593	6	
7	NUDT21A-x0066		lb18145-14	2.45	0.27973	2017-06-28 12:30:17	DL5593	7	None
8	NUDT21A-x0067		lb18145-14			2017-06-28 12:31:33	DL5593	8	
9	NUDT21A-x0068		lb18145-14			2017-06-28 12:33:19	DL5593	9	
10	NUDT21A-x0069		lb18145-14	3.01	0.33435	2017-06-28 12:36:05	DL5593	10	None
11	NUDT21A-x0070		lb18145-14	2.71	0.29731	2017-06-28 12:37:48	DL5593	11	None
12	NUDT21A-x0071		lb18145-14	2.05	0.25401	2017-06-28 12:39:56	DL5593	12	None
13	NUDT21A-x0072		lb18145-14			2017-06-28 14:10:01	DL5593	13	
14	NUDT21A-x0073		lb18145-14	7.12		2017-06-28 12:43:59	DL5593	14	
15	NUDT21A-x0074		lb18145-14			2017-06-28 12:46:29	DL5593	15	
16	NUDT21A-x0075		lb18145-14	8.29		2017-06-28 12:48:05	DL5593	16	
17	NUDT21A-x0076		lb18145-14	3.44	None	2017-06-28 12:01:39	DF045	1	None
18	NUDT21A-x0077		lb18145-14			2017-06-28 12:04:22	DF045	2	
19	NUDT21A-x0078		lb18145-14			2017-06-28 12:06:01	DF045	3	
20	NUDT21A-x0079		lb18145-14	3.40	0.40750	2017-06-28 12:07:31	DF045	4	None
21	NUDT21A-x0080		lb18145-14	2.40	0.25742	2017-06-28 12:09:43	DF045	5	None
22	NUDT21A-x0081		lb18145-14	1.81	0.26781	2017-06-28 12:12:32	DF045	6	None
23	NUDT21A-x0082		lb18145-14	3.88		2017-06-28 12:13:21	DF045	7	
24	NUDT21A-x0083		lb18145-14	2.20	0.26296	2017-06-28 12:15:05	DF045	8	None
25	NUDT21A-x0084		lb18145-14	1.89	0.26273	2017-06-28 12:16:38	DF045	9	None
26	NUDT21A-x0044		lb18145-14				DL5524	1	
27	NUDT21A-x0045		lb18145-14				DL5524	2	
28	NUDT21A-x0046		lb18145-14				DL5524	3	
29	NUDT21A-x0047		lb18145-14				DL5524	4	
30	NUDT21A-x0048		lb18145-14				DL5524	5	
31	NUDT21A-x0049		lb18145-14				DL5524	6	
32	NUDT21A-x0050		lb18145-14				DL5524	7	
33	NUDT21A-x0051		lb18145-14				DL5524	8	
34	NUDT21A-x0052		lb18145-14				DL5524	9	
35	NUDT21A-x0053		lb18145-14				DL5524	10	
36	NUDT21A-x0054		lb18145-14				DL5524	11	
37	NUDT21A-x0055		lb18145-14				DL5524	12	
38	NUDT21A-x0056		lb18145-14				DL5524	13	
39	NUDT21A-x0057		lb18145-14				DL5524	14	
40	NUDT21A-x0058		lb18145-14				DL5524	15	
41	NUDT21A-x0059		lb18145-14				DL5524	16	



LabVisit  
 LibraryName  
 Smiles  
 Compound ID  
 CrystalPlate  
 CrystalWell  
 ProteinName  
 CompoundConcentration  
 SolventFraction  
 SoakingTime  
 SoakDB Barcode  
 Visit  
 Data Collection Date

DataCollection Outcome  
 GDA Barcode  
 Path to diffraction image files  
 Program  
 DataProcessing SpaceGroup  
 Resolution High  
 Resolution [Mn</sig(I)> = 1.5]  
 Rmerge Overall  
 Rmerge Low  
 Rmerge High  
 Mn</sig(I)> High  
 Completeness Overall  
 Completeness Low

Useful for evaluating the solvent characterisation

File Datasource Preferences Deposition Proasis Help

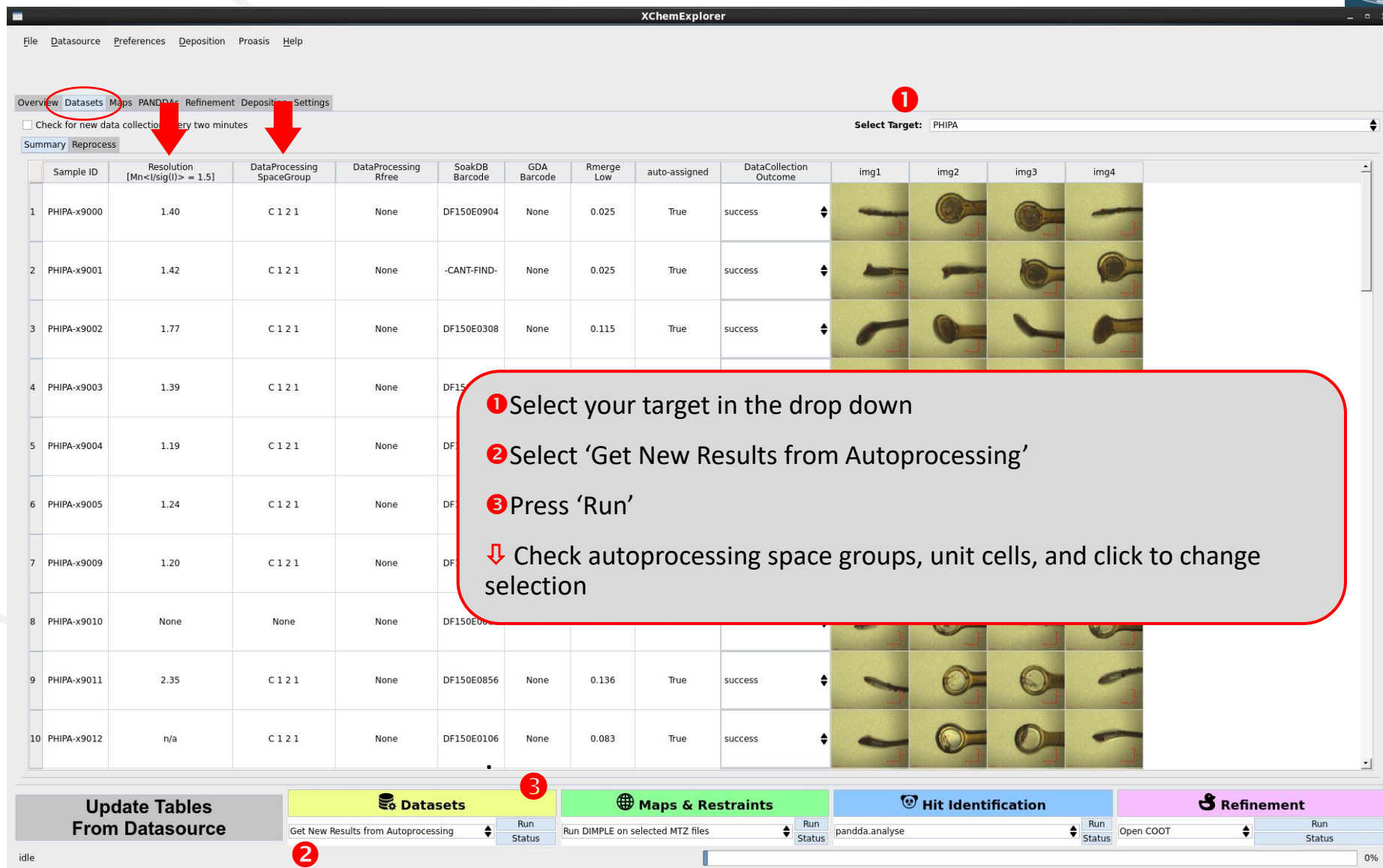
Overview Datasets Maps PANDDAs Refinement Deposition Settings

Data Source Summary

Sample ID	LibraryName	IventFracti	SoakingTime	Resolution [Mn</sig(I)> = 1.5]	
20	PHIPA-x9019	DMSO(1hr)	0	01:16:32	n/a
21	PHIPA-x9020	DMSO(1hr)	5	01:17:13	1.80
22	PHIPA-x9021	DMSO(1hr)	5	01:17:54	n/a
23	PHIPA-x9022	DMSO(1hr)	5	01:19:17	n/a
24	PHIPA-x9023	DMSO(1hr)	5	01:20:35	1.92
25	PHIPA-x9024	DMSO(1hr)	10	01:22:05	n/a
26	PHIPA-x9025	DMSO(1hr)	10	01:22:38	n/a
27	PHIPA-x9026	DMSO(1hr)	10	01:23:16	1.76
28	PHIPA-x9027	DMSO(1hr)	10	01:24:22	1.80
29	PHIPA-x9028	DMSO(1hr)	20	01:25:09	1.54
30	PHIPA-x9029	DMSO(1hr)	20	01:25:50	n/a
31	PHIPA-x9030	DMSO(1hr)	20	01:26:35	1.86
32	PHIPA-x9031	DMSO(1hr)	20	01:27:13	1.81
33	PHIPA-x9032	DMSO(3hr)	5	03:02:04	n/a
34	PHIPA-x9033	DMSO(3hr)	5	03:03:28	1.38
35	PHIPA-x9034	DMSO(3hr)	5	03:04:40	n/a
36	PHIPA-x9035	DMSO(3hr)	10	03:05:13	1.18
37	PHIPA-x9036	DMSO(3hr)	10	03:06:09	1.79
38	PHIPA-x9037	DMSO(3hr)	10	03:07:36	n/a
39	PHIPA-x9038	DMSO(3hr)	20	03:08:23	n/a
40	PHIPA-x9039	DMSO(3hr)	20	03:08:52	1.80
41	PHIPA-x9040	DMSO(3hr)	20	03:09:14	n/a
42	PHIPA-x9041	DMSO(3hr)	20	03:09:42	1.27
43	PHIPA-x9042	DMSO(3hr)	5	03:12:31	1.72
44	PHIPA-x9043	DMSO(3hr)	5	03:13:07	1.87
45	PHIPA-x9044	DMSO(3hr)	5	03:13:36	n/a
46	PHIPA-x9045	DMSO(3hr)	5	03:14:13	2.25
47	PHIPA-x9046	DMSO(3hr)	10	03:14:51	n/a

Update Tables

# Datasets tab: Load datasets













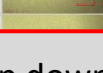
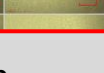


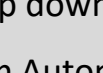
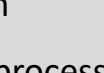
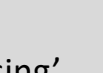

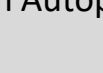
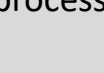
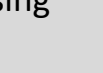


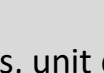
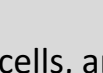

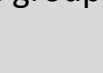
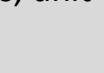
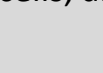
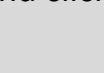










File Datasource Preferences Deposition Proasis Help

Overview **Datasets** Maps PANDPA Refinement Deposition Settings

Check for new data collection every two minutes

Select Target: PHIPA

Sample ID	Resolution [Mn</sig(l)> = 1.5]	DataProcessing SpaceGroup	DataProcessing Rfree	SoakDB Barcode	GDA Barcode	Rmerge Low	auto-assigned	DataCollection Outcome	img1	img2	img3	img4
1 PHIPA-x9000	1.40	C 1 2 1	None	DF150E0904	None	0.025	True	success				
2 PHIPA-x9001	1.42	C 1 2 1	None	-CANT-FIND-	None	0.025	True	success				
3 PHIPA-x9002	1.77	C 1 2 1	None	DF150E0308	None	0.115	True	success				
4 PHIPA-x9003	1.39	C 1 2 1	None	DF150E0604	None	0.025	True	success				
5 PHIPA-x9004	1.19	C 1 2 1	None	DF150E0804	None	0.025	True	success				
6 PHIPA-x9005	1.24	C 1 2 1	None	DF150E0404	None	0.025	True	success				
7 PHIPA-x9009	1.20	C 1 2 1	None	DF150E0604	None	0.025	True	success				
8 PHIPA-x9010	None	None	None	DF150E0604	None	0.025	True	success				
9 PHIPA-x9011	2.35	C 1 2 1	None	DF150E0856	None	0.136	True	success				
10 PHIPA-x9012	n/a	C 1 2 1	None	DF150E0106	None	0.083	True	success				

Update Tables From Datasource






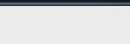
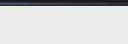
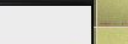


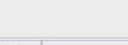
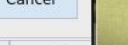

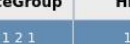

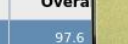

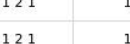

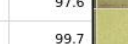



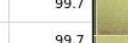













**Datasets** **Maps & Restraints** **Hit Identification** **Refinement**

Get New Results from Autoprocessing Run Status Run DIMPLE on selected MTZ files Run Status pandda.analyse Run Status Open COOT Run Status

idle 0%


**1** Select your target in the drop down  
**2** Select 'Get New Results from Autoprocessing'  
**3** Press 'Run'  
↓ Check autoprocessing space groups, unit cells, and click to change selection

# Datasets tab: Load datasets

Sample ID	Resolution [Mn<math>\langle I \rangle \rangle \ge 1.5]	DataProcessing SpaceGroup	DataProcessing Rfree	SoakDB Barcode	GDA Barcode	Rmerge Low	auto-assigned	DataCollection Outcome	img1	img2	img3	img4
1 PHIPA-x9000	1.40 	C 1 2 1										
2 PHIPA-x9001	1.42	C 1 2 1										
3 PHIPA-x9002	1.77	C 1 2 1										
4 PHIPA-x9003	1.39	C 1 2 1										
5 PHIPA-x9004	1.19	C 1 2 1										
6 PHIPA-x9005	1.24	C 1 2 1										
7 PHIPA-x9009	1.20	C 1 2 1										
8 PHIPA-x9010	None	None										
9 PHIPA-x9011	1.35	C 1 2 1										

Sample ID	Visit	Run	Program	Resolution Overall	Resolution High	DataProcessing SpaceGroup	Mn<math>\langle I \rangle \rangle \ge High	Rmerge Low	Complete Overa
1 PHIPA-x9000	lb18145-97	PHIPA-x9000_1_	3dii-run	40.38 - 1.35	1.35	C 1 2 1	1.1	0.025	97.6
2 PHIPA-x9000	lb18145-97	PHIPA-x9000_1_	3dii-runC121	40.38 - 1.35	1.35	C 1 2 1	1.1	0.025	97.6
3 PHIPA-x9000	lb18145-97	PHIPA-x9000_1_	dials-run	40.43 - 1.36	1.36	C 1 2 1	1.3	0.188	99.7
4 PHIPA-x9000	lb18145-97	PHIPA-x9000_1_	dials-run-remove-blanks	40.43 - 1.36	1.36	C 1 2 1	1.3	0.188	99.7
5 PHIPA-x9000	lb18145-97	PHIPA-x9000_1_	dials-runC121	40.42 - 1.36	1.36	C 1 2 1	0.7	0.388	99.7
6 PHIPA-x9000	lb18145-97	PHIPA-x9000_1_	autoPROC	40.38 - 1.21	1.21	C 1 2 1	0.5	0.027	90.7

 By clicking on the sample row, you can choose the autoprocesing

XCE has automatically selected the “best” one (you can specify the selection mechanism used by going to *Preference* → *Edit Preferences*)

You can also manually select the preferred autoprocesing result from the list

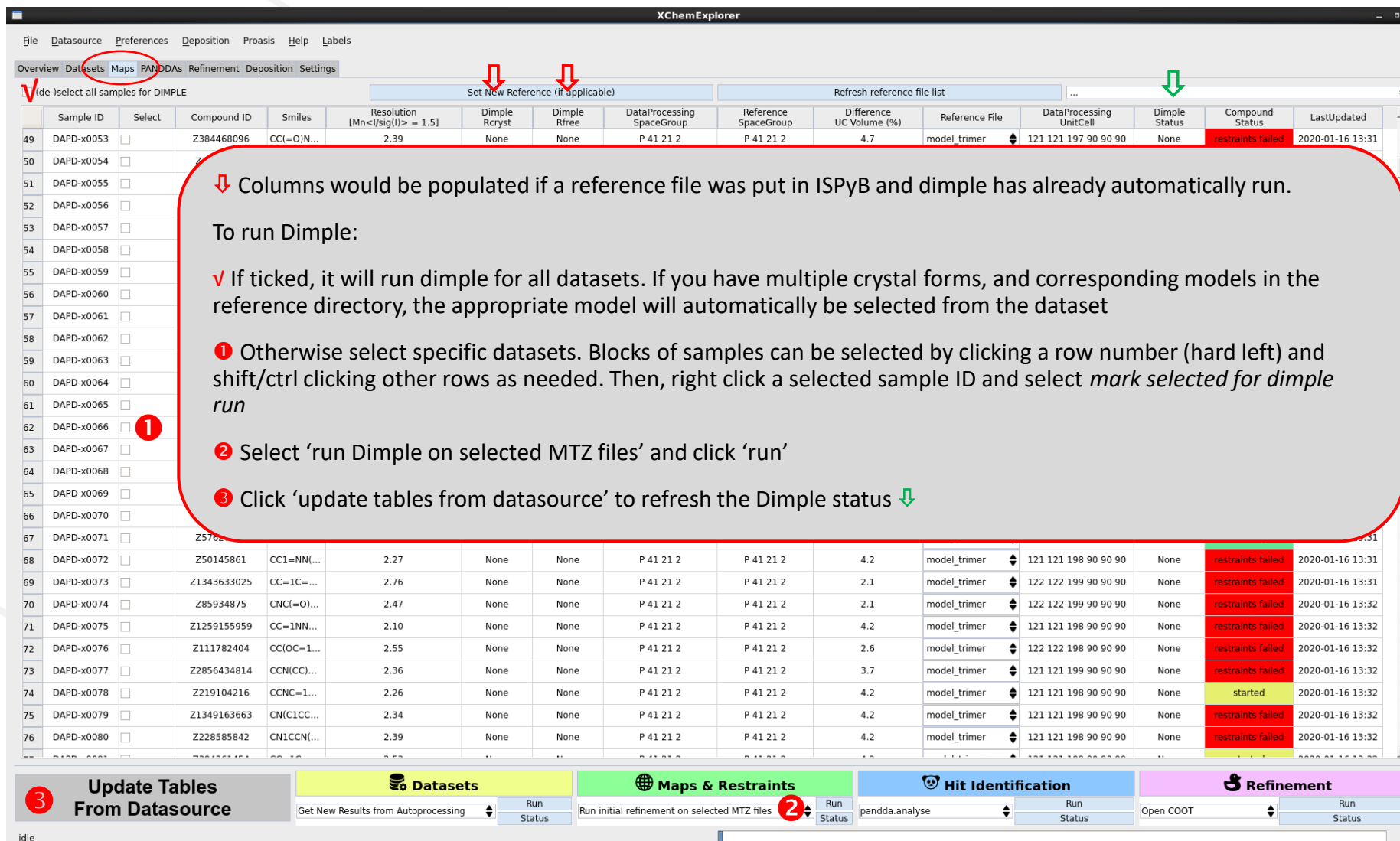
Click on Update Datasource to push the changes in the database

# Reference Model



- Use a model that you are confident best represents your crystal system as used in the XChem experiment:
  - Used previously for solving by molecular replacement
  - Containing all waters, cofactors, ligands
  - Intrinsic ligands and cofactors need PDB official three letter codes, or codes that are not 'LIG' or 'DRG'

# Maps tab: Running Dimple for MR (“Run initial refinement”)



Overview **Maps** PANDAS Refinement Deposition Settings

✓(de-)select all samples for DIMPLE

Set New Reference (if applicable) Refresh reference file list

Sample ID	Select	Compound ID	Smiles	Resolution [Mn</sig(I)> = 1.5]	Dimple Rcryst	Dimple Rfree	DataProcessing SpaceGroup	Reference SpaceGroup	Difference UC Volume (%)	Reference File	DataProcessing UnitCell	Dimple Status	Compound Status	LastUpdated
49	<input type="checkbox"/>	Z384468096	CC(=O)N...	2.39	None	None	P 41 21 2	P 41 21 2	4.7	model_trimer	121 121 197 90 90 90	None	restraints failed	2020-01-16 13:31
50	<input type="checkbox"/>													
51	<input type="checkbox"/>													
52	<input type="checkbox"/>													
53	<input type="checkbox"/>													
54	<input type="checkbox"/>													
55	<input type="checkbox"/>													
56	<input type="checkbox"/>													
57	<input type="checkbox"/>													
58	<input type="checkbox"/>													
59	<input type="checkbox"/>													
60	<input type="checkbox"/>													
61	<input type="checkbox"/>													
62	<input checked="" type="checkbox"/>													
63	<input type="checkbox"/>													
64	<input type="checkbox"/>													
65	<input type="checkbox"/>													
66	<input type="checkbox"/>													
67	<input type="checkbox"/>	Z576...												
68	<input type="checkbox"/>	Z50145861	CC1=NN(...	2.27	None	None	P 41 21 2	P 41 21 2	4.2	model_trimer	121 121 198 90 90 90	None	restraints failed	2020-01-16 13:31
69	<input type="checkbox"/>	Z1343633025	CC=1C=...	2.76	None	None	P 41 21 2	P 41 21 2	2.1	model_trimer	122 122 199 90 90 90	None	restraints failed	2020-01-16 13:31
70	<input type="checkbox"/>	Z85934875	CNC(=O)...	2.47	None	None	P 41 21 2	P 41 21 2	2.1	model_trimer	122 122 199 90 90 90	None	restraints failed	2020-01-16 13:32
71	<input type="checkbox"/>	Z1259155959	CC=1NN...	2.10	None	None	P 41 21 2	P 41 21 2	4.2	model_trimer	121 121 198 90 90 90	None	restraints failed	2020-01-16 13:32
72	<input type="checkbox"/>	Z111782404	CC(OC=1...	2.55	None	None	P 41 21 2	P 41 21 2	2.6	model_trimer	122 122 198 90 90 90	None	restraints failed	2020-01-16 13:32
73	<input type="checkbox"/>	Z2856434814	CCN(CC)...	2.36	None	None	P 41 21 2	P 41 21 2	3.7	model_trimer	121 121 199 90 90 90	None	restraints failed	2020-01-16 13:32
74	<input type="checkbox"/>	Z219104216	CCNC=1...	2.26	None	None	P 41 21 2	P 41 21 2	4.2	model_trimer	121 121 198 90 90 90	None	started	2020-01-16 13:32
75	<input type="checkbox"/>	Z1349163663	CN1CC...	2.34	None	None	P 41 21 2	P 41 21 2	4.2	model_trimer	121 121 198 90 90 90	None	restraints failed	2020-01-16 13:32
76	<input type="checkbox"/>	Z228585842	CN1CCN(...	2.39	None	None	P 41 21 2	P 41 21 2	4.2	model_trimer	121 121 198 90 90 90	None	restraints failed	2020-01-16 13:32

Columns would be populated if a reference file was put in ISPyB and dimple has already automatically run.

To run Dimple:

- ✓ If ticked, it will run dimple for all datasets. If you have multiple crystal forms, and corresponding models in the reference directory, the appropriate model will automatically be selected from the dataset
- 1 Otherwise select specific datasets. Blocks of samples can be selected by clicking a row number (hard left) and shift/ctrl clicking other rows as needed. Then, right click a selected sample ID and select *mark selected for dimple run*
- 2 Select 'run Dimple on selected MTZ files' and click 'run'
- 3 Click 'update tables from datasource' to refresh the Dimple status ↓

3 Update Tables From Datasource

Update New Results from Autoprocessing Run Status

Maps & Restraints Run initial refinement on selected MTZ files Run Status

Hit Identification pandda.analyse Run Status

Refinement Open COOT Run Status

# Check jobs are running

- To check the status of jobs on the cluster, type into terminal:
  - “ssh wilson” – connect to the DLS Slurm Cluster; must be done first
  - “sacct” – display jobs
  - “watch sq.sh -u *yourfedid* -nf” – watch jobs

```
Every 2.0s: sq.sh -u ill13029 -nf
```

```
ill13029's queue
```

```
No jobs running.
```

```
No jobs pending.
```

```
Previous 10 jobs (last fortnight):
```

JobID	'Job Name'	#N #C	Start Time	Run Time	Status
9363997	'xce_buster'	1 1c	Jul 1st 08:38	40m 20s	Completed
9368556	'xce_buster'	1 1c	Jul 1st 12:58	36m 21s	Completed
9370433	'xce_buster'	1 1c	Jul 1st 13:50	33m 4s	Completed
9372944	'xce_buster'	1 1c	Jul 1st 14:55	24m 30s	Completed



# Maps tab: Creating the ligands restraints

1

XChemExplorer

File Data Source Preferences Deposition Help

Overview Datasets Maps PANDDAS Refinement Deposition Settings

✓ (de-)select all samples for DIMPLE

Set New Reference (if applicable)

Sample ID	Select	Compound ID	Smiles	Resolution n<1/sig(I)> = 1	Dimple Rcryst	Dimple Rfree	DataProcessing SpaceGroup	Reference SpaceGroup	Difference UC Volume (%)	Reference File	DataProcessing UnitCell	Dimple Status	Compound Status	LastUpdated
1	✓	FMOOA0008...	Cc1nc2ccc(c(...	1.48	0.21304	0.24324	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	finished	restraints failed	2017-01-17 1...
2	✓	FMOOA0008...	Cc1nc2ccc(c(...	1.31	None	None	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	running	restraints failed	2017-01-17 1...
3	✓	FMOOA0008...	Cc1nc2ccc(c(...	1.46	None	None	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	running	restraints failed	2017-01-17 1...
4	✓	FMOOA0008...	Cc1nc2ccc(c(...	1.23	None	None	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	running	restraints failed	2017-01-17 1...
5	✓	FMOOA0007...	c1ccc2c(c1)n...	1.15	None	None	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	running	restraints failed	2017-01-17 1...
6	✓	FMSOA00140...	c1ccc2c(c1)n...	1.44	0.21881	0.24943	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	finished	running	2017-01-17 1...
7	✓	FMOOA0007...	Cc1nc2cccc...	1.56	0.21405	0.24522	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	finished	running	2017-01-17 1...
8	✓	XST00000832b	c1ccc2c(c1)n...	1.56	0.21831	0.25413	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	finished	running	2017-01-17 1...
9	✓	FMOOA0007...	c1ccc2c(c1)n...	1.81	0.21946	0.26333	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	finished	running	2017-01-17 1...
10	✓	XST00000560c	c1ccc2c(c1)n...	1.66	0.21508	0.24978	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	finished	running	2017-01-17 1...
11	✓	FMOOA0008...	c1ccc2c(c1)n...	2.00	0.21879	0.27986	P 41 21 2	P 43 21 2	0.0	JMJD2DA.ref	72 72 151 90...	finished	running	2017-01-17 1...
12	✓	JMJD2DA-x1719										finished	running	2017-01-17 1...
13	✓	JMJD2DA-x1717										finished	running	2017-01-17 1...
14	✓	JMJD2DA-x1716										finished	running	2017-01-17 1...
15	✓	JMJD2DA-x1715										finished	running	2017-01-17 1...
16	✓	JMJD2DA-x1708										finished	running	2017-01-17 1...
17	✓	JMJD2DA-x1707										finished	running	2017-01-17 1...
18	✓	JMJD2DA-x1699										finished	running	2017-01-17 1...
19	✓	JMJD2DA-x1696										finished	running	2017-01-17 1...
20	✓	JMJD2DA-x1692										finished	running	2017-01-17 1...
21	✓	JMJD2DA-x1686										finished	running	2017-01-17 1...
22	✓	JMJD2DA-x1681										finished	running	2017-01-17 1...
23	✓	JMJD2DA-x1728										None	running	2017-01-17 1...
24	✓	JMJD2DA-x1680										None	running	2017-01-17 1...
25	✓	JMJD2DA-x1667										None	running	2017-01-17 1...
26	✓	JMJD2DA-x1666										None	running	2017-01-17 1...
27	✓	JMJD2DA-x1663										None	running	2017-01-17 1...
28	✓	JMJD2DA-x1659										None	running	2017-01-17 1...
29	✓	JMJD2DA-x1643										None	running	2017-01-17 1...
30	✓	JMJD2DA-x1637										None	running	2017-01-17 1...

2

3

Update Tables From Datasource

Datasets

Maps & Restraints

Hit Identification

Refinement

Get New Results from Autoprocessing Run Status

Create CIF/PDB/PNG file of ALL compounds Run Status

pandda.analyse Run Status

Open COOT Run Status

idle 0%

3

2

1

Go to Preferences -> Edit preferences. You will get a pop up window where you can change the program to use. You have the choice between: acedrg (default), grade and phenix.elbow

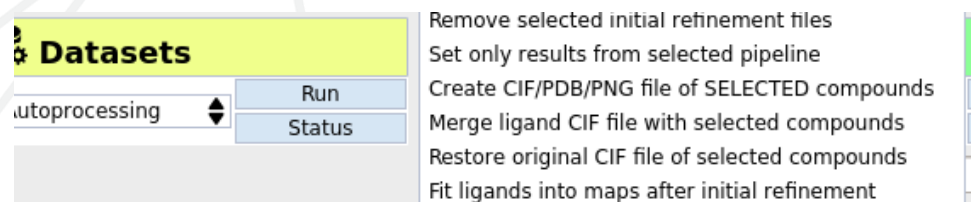
Select 'create CIF/PDB/PNG files for ALL compounds' or 'create CIF/PDB/PNG files for selected compounds' if you have selected some and click 'run'

Click 'update tables from datasource' to refresh the Compound status. If the bulk have failed, change the number of jobs submitted concurrently to the cluster in preferences.

# Merge ligand restraints from non-standard ligand



1. Open 'Preferences' menu (Edit preferences) and select the CIF file of your non-standard ligand in 'Additional CIF file for non-standard ligand'.
2. Select the samples which you want to merge in the Maps tab
3. Choose 'Merge ligand CIF file with selected compounds' and press Run.



XCE will now prepare a merged version of the file in the sample directory with the same name. It does not touch the original files in the compound subfolder.

Before you start merging: **the ligand code of the additional ligand cannot be LIG or DRG!** Both codes are reserved for ligands generated by XCE.

# Merge ligand restrains from non-standard ligand



## Restore original CIF file

In case you need/ want to restore the original CIF file:

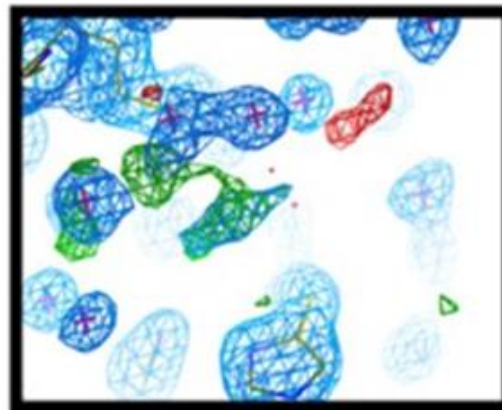
1. select the samples in the Maps tab which you want to restore (see above).
2. choose '*Merge ligand CIF file with selected compounds*' from the green action box and press *Run*.

Please note that this is not a requirement in case you want to merge another ligand. XCE will in this case first remove the old, merged CIF file, before doing the merging as described before.

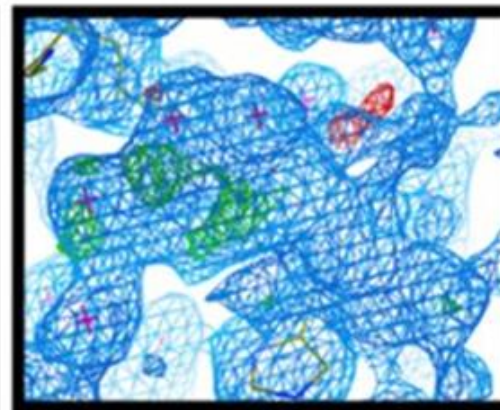
# Finding hits - Pan Density Dataset Analysis (PanDDA)



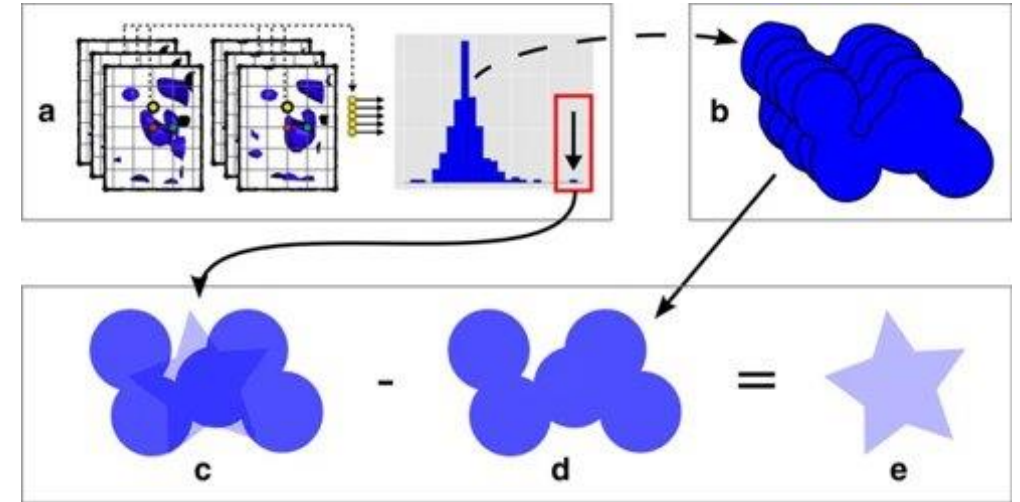
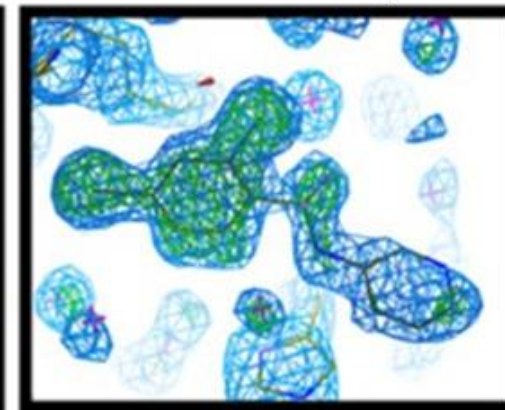
High Contour



Low Contour



PanDDA Map



# Ground-state model building



- A PanDDA pre-run allows you to build the best possible reference model: the ground-state model.

1 Click on the drop-down menu in the “Hit identification” action box

2 Select “pre-run for ground state model”.

Wait for the job to finish. This creates a subdirectory in the reference directory with all the required files.

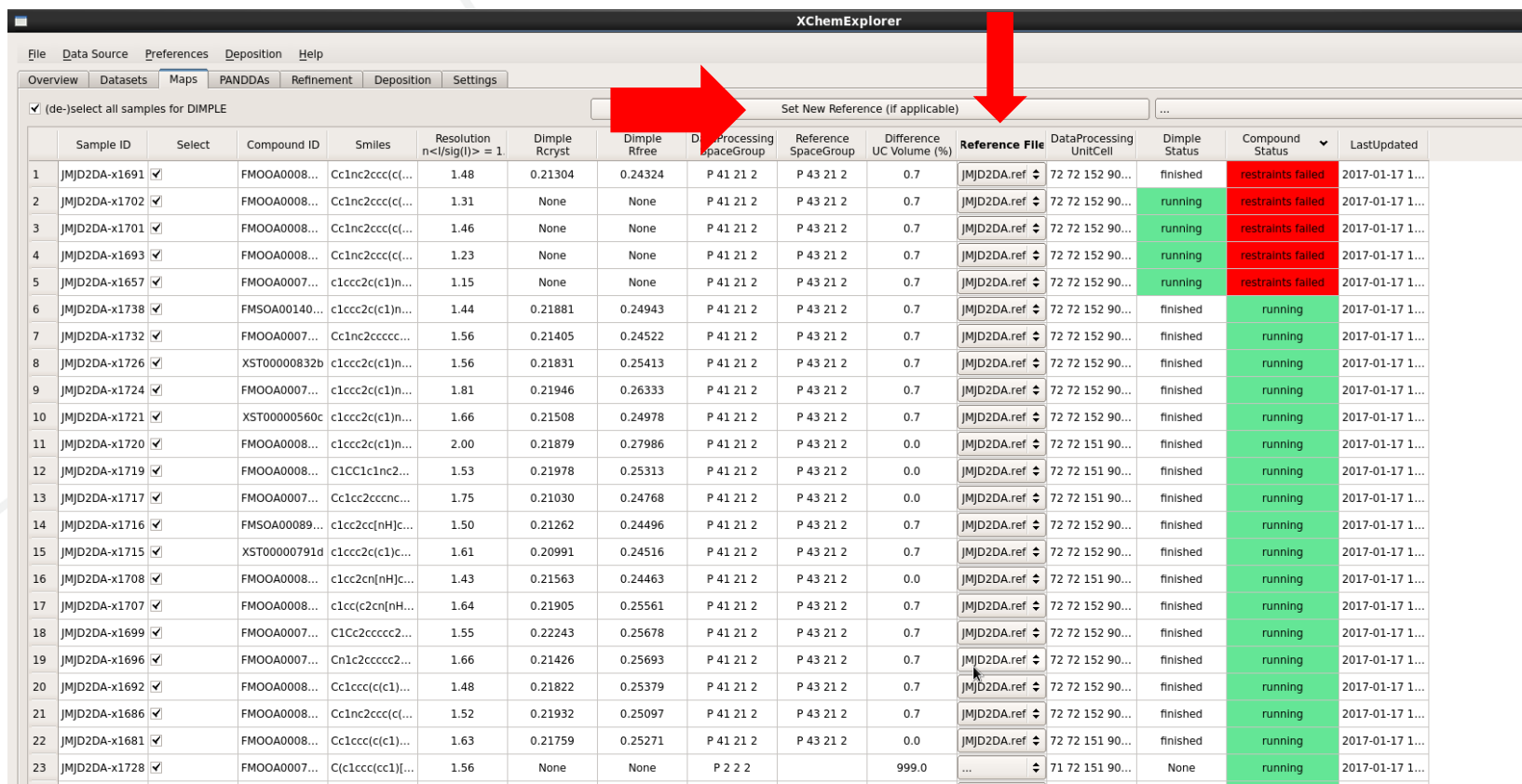
3 Once the PanDDA pre-run is done, select “build ground state model”.

Coot will open the PanDDA mean-map and the 2Fo-Fc/Fo-Fc maps loaded from the new reference/subdirectory

A screenshot of the XCHEM web interface. A dropdown menu is open, showing various actions. A red circle '1' is above the menu. A red circle '2' is next to the 'pre-run for ground state model' option, and a red circle '3' is next to the 'Build ground state model' option. The menu items include: pandda.analyse, pandda.inspect, run pandda.inspect at home, Export NEW PANDDA models, Export ALL PANDDA models, Export SELECTED PANDDA models, Show HTML summary, cluster datasets, Event Map -&gt; SF, apo -&gt; mmcif, check modelled ligands, refine ALL bound-state models with BUSTER, refine NEW bound-state models with BUSTER, refine ALL bound-state models with BUSTER (no sanity check), refine NEW bound-state models with BUSTER (no sanity check), pre-run for ground state model, and Build ground state model. The 'pandda.analyse (PanDDA2)' option is highlighted in blue at the bottom of the menu. The background shows a 'Refinement' section with a 'Run' button and a 'Status' button.

# Ground-state model building

- Remodel and refine the reference model as you wish using the PanDDA mean map in Coot.
- Re-run Dimple (XCE – Maps table) by using this ground-state model as new reference



XChemExplorer

File Data Source Preferences Deposition Help

Overview Datasets Maps PANDDAs Refinement Deposition Settings

(de-)select all samples for DIMPLE

Set New Reference (if applicable) ...

Sample ID	Select	Compound ID	Smiles	Resolution n</sig(I)> = 1	Dimple Rcryst	Dimple Rfree	Data Processing SpaceGroup	Reference SpaceGroup	Difference UC Volume (%)	Reference File	DataProcessing UnitCell	Dimple Status	Compound Status	LastUpdated
1	<input checked="" type="checkbox"/>	FMOOA0008...	Cc1nc2ccc(c(...	1.48	0.21304	0.24324	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	finished	restraints failed	2017-01-17 1...
2	<input checked="" type="checkbox"/>	FMOOA0008...	Cc1nc2ccc(c(...	1.31	None	None	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	running	restraints failed	2017-01-17 1...
3	<input checked="" type="checkbox"/>	FMOOA0008...	Cc1nc2ccc(c(...	1.46	None	None	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	running	restraints failed	2017-01-17 1...
4	<input checked="" type="checkbox"/>	FMOOA0008...	Cc1nc2ccc(c(...	1.23	None	None	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	running	restraints failed	2017-01-17 1...
5	<input checked="" type="checkbox"/>	FMOOA0007...	c1ccc2c(c1)n...	1.15	None	None	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	running	restraints failed	2017-01-17 1...
6	<input checked="" type="checkbox"/>	FMSOA00140...	c1ccc2c(c1)n...	1.44	0.21881	0.24943	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	finished	running	2017-01-17 1...
7	<input checked="" type="checkbox"/>	FMOOA0007...	Cc1nc2cccc...	1.56	0.21405	0.24522	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	finished	running	2017-01-17 1...
8	<input checked="" type="checkbox"/>	XST00000832b	c1ccc2c(c1)n...	1.56	0.21831	0.25413	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	finished	running	2017-01-17 1...
9	<input checked="" type="checkbox"/>	FMOOA0007...	c1ccc2c(c1)n...	1.81	0.21946	0.26333	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	finished	running	2017-01-17 1...
10	<input checked="" type="checkbox"/>	XST00000560c	c1ccc2c(c1)n...	1.66	0.21508	0.24978	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	finished	running	2017-01-17 1...
11	<input checked="" type="checkbox"/>	FMOOA0008...	c1ccc2c(c1)n...	2.00	0.21879	0.27986	P 41 21 2	P 43 21 2	0.0	JMJD2DA.ref	72 72 151 90...	finished	running	2017-01-17 1...
12	<input checked="" type="checkbox"/>	FMOOA0008...	C1CC1c1nc2...	1.53	0.21978	0.25313	P 41 21 2	P 43 21 2	0.0	JMJD2DA.ref	72 72 151 90...	finished	running	2017-01-17 1...
13	<input checked="" type="checkbox"/>	FMOOA0007...	Cc1cc2cccc...	1.75	0.21030	0.24768	P 41 21 2	P 43 21 2	0.0	JMJD2DA.ref	72 72 151 90...	finished	running	2017-01-17 1...
14	<input checked="" type="checkbox"/>	FMSOA00089...	c1cc2c[nH]c...	1.50	0.21262	0.24496	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	finished	running	2017-01-17 1...
15	<input checked="" type="checkbox"/>	XST00000791d	c1ccc2c(c1)c...	1.61	0.20991	0.24516	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	finished	running	2017-01-17 1...
16	<input checked="" type="checkbox"/>	FMOOA0008...	c1cc2c[nH]c...	1.43	0.21563	0.24463	P 41 21 2	P 43 21 2	0.0	JMJD2DA.ref	72 72 151 90...	finished	running	2017-01-17 1...
17	<input checked="" type="checkbox"/>	FMOOA0008...	c1cc(c2c[nH]...	1.64	0.21905	0.25561	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	finished	running	2017-01-17 1...
18	<input checked="" type="checkbox"/>	FMOOA0007...	C1Cc2cccc2...	1.55	0.22243	0.25678	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	finished	running	2017-01-17 1...
19	<input checked="" type="checkbox"/>	FMOOA0007...	Cn1c2cccc2...	1.66	0.21426	0.25693	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	finished	running	2017-01-17 1...
20	<input checked="" type="checkbox"/>	FMOOA0008...	Cc1ccc(c(c1)...	1.48	0.21822	0.25379	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	finished	running	2017-01-17 1...
21	<input checked="" type="checkbox"/>	FMOOA0008...	Cc1nc2ccc(c(...	1.52	0.21932	0.25097	P 41 21 2	P 43 21 2	0.7	JMJD2DA.ref	72 72 152 90...	finished	running	2017-01-17 1...
22	<input checked="" type="checkbox"/>	FMOOA0008...	Cc1ccc(c(c1)...	1.63	0.21759	0.25271	P 41 21 2	P 43 21 2	0.0	JMJD2DA.ref	72 72 151 90...	finished	running	2017-01-17 1...
23	<input checked="" type="checkbox"/>	FMOOA0007...	C(c1ccc(cc1)[...	1.56	None	None	P 2 2 2		999.0	...	71 72 151 90...	None	running	2017-01-17 1...

# PanDDA Workflow



- “pandda.analyse” – uses PanDDA to generate event maps
- “pandda.analyse (PanDDA2)” – uses PanDDA 2 to calculate statistical models, generate event maps, and autofit ligands
  - Many datasets, larger unit cells, and multimers can all increase run time
  - Even though ligands are autofit, you **must**: “mark events as interesting”; select “Ligand placed”; and assign a confidence level in pandda.inspect for relevant datasets
- pandda.inspect: COOT plugin to inspect, annotate and place the fragments
  - **Not traditional model refinement – do not use to refine model at large**
- Export models for refinement
- Refine models
- Deposit/disseminate data

# PanDDAs Tab: pandda.analyse



The screenshot shows the XChemExplorer interface with the 'PanDDAs' tab selected. The 'pandda.analyse' sub-tab is active. A red box highlights the 'PANDDAs' tab in the top navigation bar. A red arrow points to the 'Output directory' field, which contains the path `/001/lb27001-39/processing/analysis/panddas`. A red box with a downward arrow and text provides instructions: 'Create a timestamped 'processing/analysis/panddas\_XXX' directory, and repeat for subsequent runs.' Below this, another red box with a red circle and '1' instructs to 'Select 'pandda.analyse (PanDDA2)' and click 'run''. A context menu is open over the 'pandda.analyse (PanDDA2)' button, with the 'pandda.analyse' option highlighted and a red circle and '1' next to it. The status bar at the bottom shows 'STATUS: UNKNOWN' and a 'Ready' indicator.

↓ Create a timestamped 'processing/analysis/panddas\_XXX' directory, and repeat for subsequent runs.

1 Select 'pandda.analyse (PanDDA2)' and click 'run'

STATUS: UNKNOWN

Update Tables From Datasource

Datasets

Maps & Restraints

Refinement

Ready

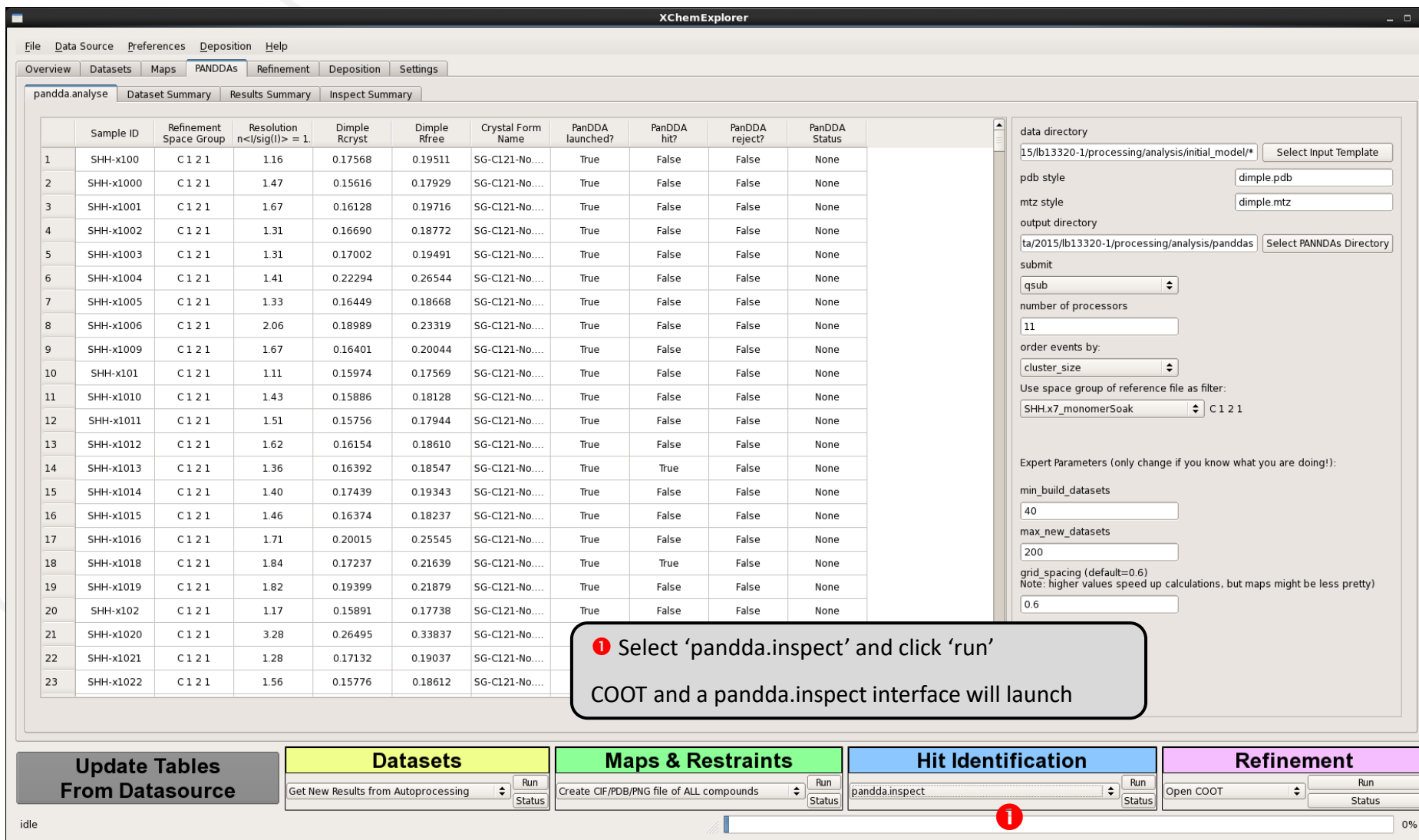
XChemExplorer

Check if jobs are running on the cluster as described previously





# PanDDAs Tab: pandda.inspect



The screenshot shows the XChemExplorer interface with the PanDDAs tab selected. The main window displays a table of sample data with columns for Sample ID, Refinement Space Group, Resolution, Dimple Rcryst, Dimple Rfree, Crystal Form Name, PanDDA launched?, PanDDA hit?, PanDDA reject?, and PanDDA Status. A callout box with a red 'i' icon contains the instruction: "Select 'pandda.inspect' and click 'run' COOT and a pandda.inspect interface will launch". The bottom of the interface features a control bar with buttons for "Update Tables From Datasource", "Datasets", "Maps & Restraints", "Hit Identification", and "Refinement". The "Hit Identification" button is highlighted with a red 'i' icon and has "pandda.inspect" selected in its dropdown menu.

Sample ID	Refinement Space Group	Resolution n<math>\langle I \rangle > = 1	Dimple Rcryst	Dimple Rfree	Crystal Form Name	PanDDA launched?	PanDDA hit?	PanDDA reject?	PanDDA Status	
1	SHH-x100	C 1 2 1	1.16	0.17568	0.19511	SG-C121-No....	True	False	False	None
2	SHH-x1000	C 1 2 1	1.47	0.15616	0.17929	SG-C121-No....	True	False	False	None
3	SHH-x1001	C 1 2 1	1.67	0.16128	0.19716	SG-C121-No....	True	False	False	None
4	SHH-x1002	C 1 2 1	1.31	0.16690	0.18772	SG-C121-No....	True	False	False	None
5	SHH-x1003	C 1 2 1	1.31	0.17002	0.19491	SG-C121-No....	True	False	False	None
6	SHH-x1004	C 1 2 1	1.41	0.22294	0.26544	SG-C121-No....	True	False	False	None
7	SHH-x1005	C 1 2 1	1.33	0.16449	0.18668	SG-C121-No....	True	False	False	None
8	SHH-x1006	C 1 2 1	2.06	0.18989	0.23319	SG-C121-No....	True	False	False	None
9	SHH-x1009	C 1 2 1	1.67	0.16401	0.20044	SG-C121-No....	True	False	False	None
10	SHH-x101	C 1 2 1	1.11	0.15974	0.17569	SG-C121-No....	True	False	False	None
11	SHH-x1010	C 1 2 1	1.43	0.15886	0.18128	SG-C121-No....	True	False	False	None
12	SHH-x1011	C 1 2 1	1.51	0.15756	0.17944	SG-C121-No....	True	False	False	None
13	SHH-x1012	C 1 2 1	1.62	0.16154	0.18610	SG-C121-No....	True	False	False	None
14	SHH-x1013	C 1 2 1	1.36	0.16392	0.18547	SG-C121-No....	True	True	False	None
15	SHH-x1014	C 1 2 1	1.40	0.17439	0.19343	SG-C121-No....	True	False	False	None
16	SHH-x1015	C 1 2 1	1.46	0.16374	0.18237	SG-C121-No....	True	False	False	None
17	SHH-x1016	C 1 2 1	1.71	0.20015	0.25545	SG-C121-No....	True	False	False	None
18	SHH-x1018	C 1 2 1	1.84	0.17237	0.21639	SG-C121-No....	True	True	False	None
19	SHH-x1019	C 1 2 1	1.82	0.19399	0.21879	SG-C121-No....	True	False	False	None
20	SHH-x102	C 1 2 1	1.17	0.15891	0.17738	SG-C121-No....	True	False	False	None
21	SHH-x1020	C 1 2 1	3.28	0.26495	0.33837	SG-C121-No....	True	False	False	None
22	SHH-x1021	C 1 2 1	1.28	0.17132	0.19037	SG-C121-No....	True	False	False	None
23	SHH-x1022	C 1 2 1	1.56	0.15776	0.18612	SG-C121-No....	True	False	False	None

data directory  
[15/lb13320-1/processing/analysis/initial\_model/\*] Select Input Template

pdb style  
dimple.pdb

mtz style  
dimple.mtz

output directory  
[ta/2015/lb13320-1/processing/analysis/panddas] Select PANDDAs Directory

submit  
qsub

number of processors  
11

order events by:  
cluster\_size

Use space group of reference file as filter:  
SHH.x7\_monomerSoak C 1 2 1

Expert Parameters (only change if you know what you are doing!):

min\_build\_datasets  
40

max\_new\_datasets  
200

grid\_spacing (default=0.6)  
Note: higher values speed up calculations, but maps might be less pretty!  
0.6

Update Tables From Datasource

Datasets  
Get New Results from Autoprocessing Run Status

Maps & Restraints  
Create CIF/PDB/PNG file of ALL compounds Run Status

Hit Identification  
pandda.inspect Run Status

Refinement  
Open COOT Run Status

idle 0%

# pandda.inspect COOT interface



The screenshot displays the COOT 0.8.8 EL (ccp4) interface. The main window shows a 3D molecular model with a blue mesh and green sticks. The PANDDA inspect panel on the right contains the following information:

**PANDDA inspect**

Quit Overall Inspection Event/Site Progress: Go to Dataset: [ ] Go

Summary Event 1 of 404 <<< Go to Prev Site <<< >>> Go to Next Site >>>

Update HTML Site 1 of 21 >>> Go to Next Unviewed >>> >>> Go to Next Modelled >>>

<<< Prev <<< (Don't Save Model) >>> Next >>> (Don't Save Model) >>> Next >>> (Save Model)

Dataset ID				NUDT21A-x0743		Merge Ligand With Model		Save Model	
Event Information:		Dataset Information:		Move New Ligand Here		Reload Last Saved Model		Reset to Unfitted Model	
Event #	3	Resolution	1.65	Open Next Ligand					
1 - BDC	0.28	Map Uncertainty	0.23						
Z-blob Peak	11.0	R-Free / R-Work	0.213 / 0.247						
Z-blob Size	796								

Record Event Information (this event only)

Event Comment: [ DSPL Hit ]

Mark Event as Interesting  Mark Event as Not Interesting  Ligand Placed  No Ligand Placed  Model: High Confidence  Model: Medium Confidence  Model: Low Confidence

Record Site Information (for all events with this site)

Name: [ Active-mRNA site ]

Comment: [ LYS105 dangling in ]

Miscellaneous buttons: Load input mtz file Load average map Load unfitted model (for comparison only) Create new ligand

**Display Manager**

Maps: All

Maps	Display	Scroll	Properties	Delete Map
2 NUDT21A-x0743-z_map.native.ccp4	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
3 NUDT21A-x0743-event_3_1-BDC_0.28_map.nativ	<input checked="" type="checkbox"/>	<input type="checkbox"/>		

Molecules: All

Molecules	Display	Active	Bonds (Colour by Atom)	Delete Model
0 FMOPL000589a.pdb	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
1 NUDT21A-x0743-pandda-model.pdb	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		

Close



# pandda.inspect COOT interface

z\_map.native.ccp4

(set to appear like a difference map, on by default)

Shows the extent of deviations from the ensemble of crystallographic datasets. Large positive or negative Z-scores ( $\pm 3$ ) indicate significant deviations from the ensemble, and may represent interesting features.

event\_X\_1-  
BDC\_Y\_map.ccp4

(the important one! On by default)

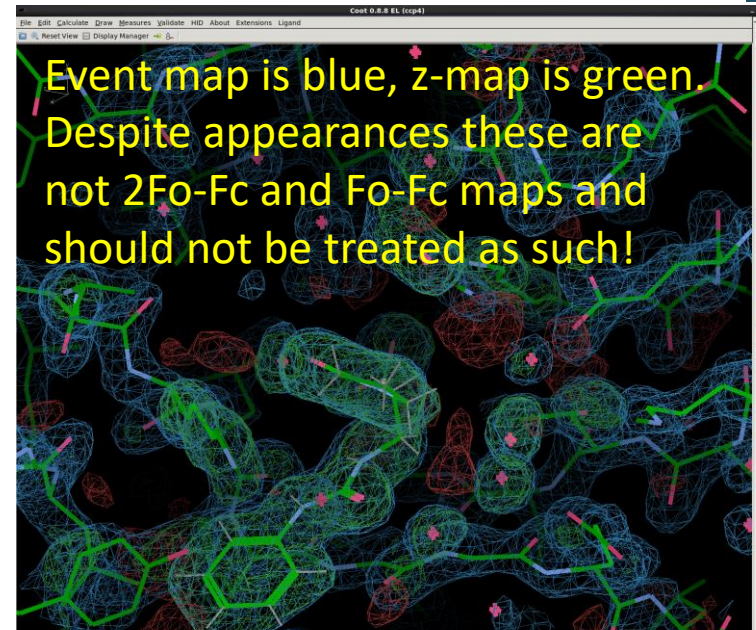
Partial-difference density obtained by subtracting a fraction of the mean map from the dataset map. This reveals the density for low-occupancy binding events. X indicates which event in this dataset is being inspected, and Y indicates the amount of mean map that has been subtracted (amount subtracted = 1-Y).

ligand files

Loaded automatically. PanDDA 2 will have attempted ligand fitting, but this file is present/hidden in case of multiple sites, re-fitting.

-pandda-model.pdb

The output of pandda2-analyse, with auto-fitted ligand in position (if an autobuild occurred)



Maps	
2	NUDT21A-x0743-z_map.native.ccp4
3	NUDT21A-x0743-event_3_1-BDC_0.28_map.nativ

Molecules	
0	FMOPL000589a.pdb
1	NUDT21A-x0743-pandda-model.pdb

# pandda.inspect COOT interface



- 1 Open html summary page of the data analysis
- 2 Indicates number of sites and events to review
- 3 Navigate through the events and sites, or go straight to a dataset of interest

- 1 Summary of PanDDA statistics
- 2 **Merge or add ligands to the model**
- 3 **Save your model** or roll back to previous models.

- 1 To **annotate the event**.
- 2 To **annotate the sites**. It will be used by XCE to categorise models in refinement.

**For your hits to be taken to the next step (If you do not follow these steps you will not be able to export your models!):**

- 'Mark Event as Interesting' and 'Ligand Placed' **must** be selected
- Save model (or 'Next' (Save model)). A pandda-model.pdb will be saved in processed\_datasets/\*/modelled\_structures/
- Update the event information as necessary
- Do not save useless/empty/dubious model

PANDDA inspect

Quit Overall Inspection Event/Site Progress: Go to Dataset: [ ] Go

Summary Event 1 of 404 <<< Go to Prev Site <<< >>> Go to Next Site >>>

Update HTML Site 1 of 21 >>> Go to Next Unviewed >>> >>> Go to Next Modelled >>>

<<< Prev <<< (Don't Save Model) >>> Next >>> (Don't Save Model) >>> Next >>> (Save Model)

Dataset ID **1** NUDT21A-x0743 Merge Ligand With Model **2** Save Model

Event Information:		Dataset Information:	
Event #	3	Resolution	1.65
1 - BDC	0.28	Map Uncertainty	0.23
Z-blob Peak	11.0	R-Free / R-Work	0.213 / 0.247
Z-blob Size	796		

Move New Ligand Here Reload Last Saved Model

Open Next Ligand Reset to Unfitted Model

**1** Record Event Information (this event only)

Event Comment: [ DSPL Hit ]

Mark Event as Interesting  Ligand Placed  Model: High Confidence

Mark Event as Not Interesting  No Ligand Placed  Model: Medium Confidence

Model: Low Confidence

**2** Record Site Information (for all events with this site)

Name: [ Active-mRNA site ]

Comment: [ LYS105 dangling in ]

Miscellaneous buttons Load input mtz file Load average map Load unfitted model (for comparison only) Create new ligand

# Modelling in pandda.inspect

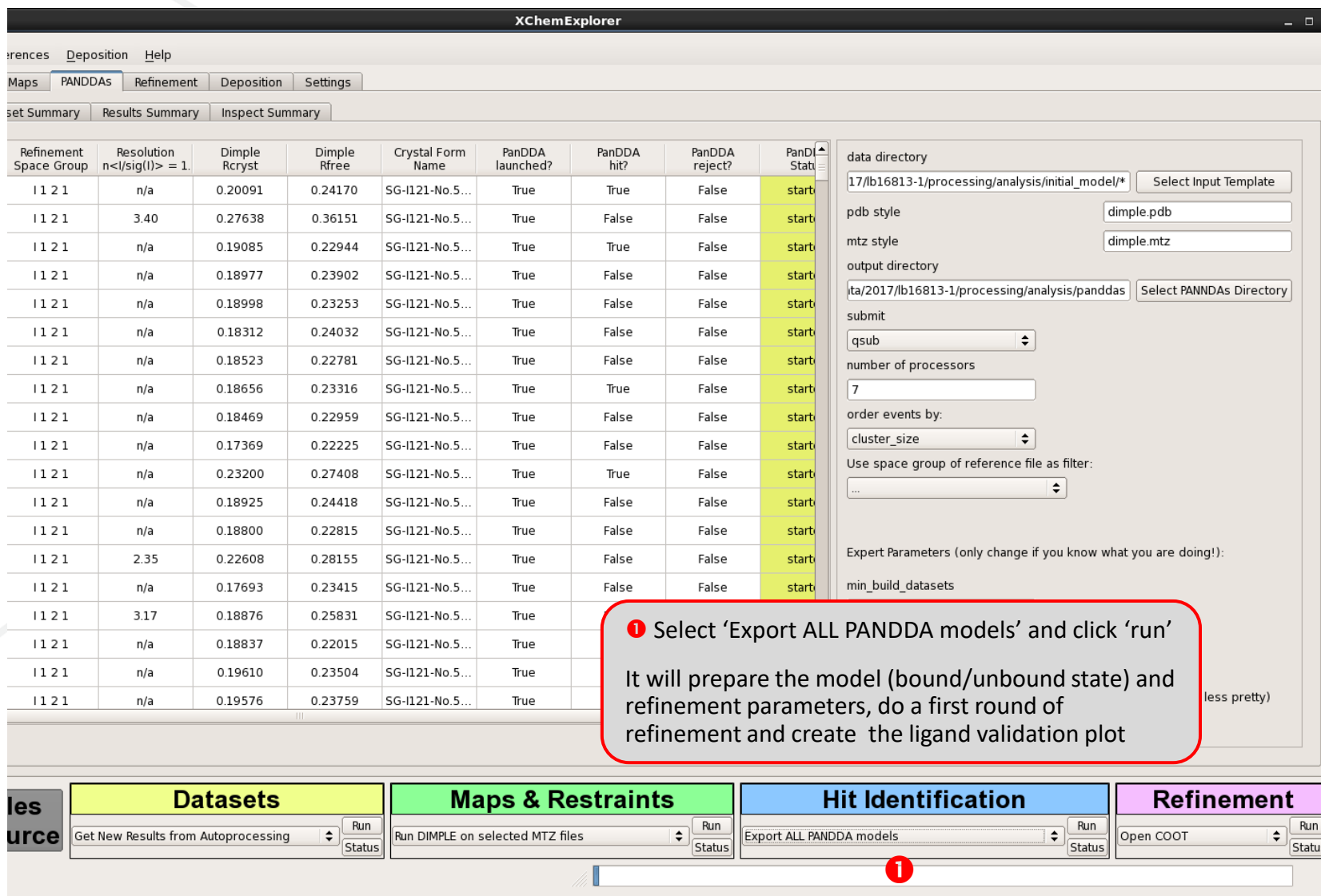


With PanDDA, you are not trying to build the entire model – just a model of the protein when something is bound to it: **i.e. the bound-state model.**

- Focus on the centred event map - **do not navigate away** from the initial view or search for blobs using Coot
- If you cannot clearly see the ligand pose in the PanDDA event map move on, there will be plenty more events to check!
- **Only change/delete atoms that are “important”** - with large peaks in the Z-map, clear shifts in location - other smaller changes can be built in refinement
- Think **‘would I give this model to a chemist for follow-up compound design?’** ‘Would I spend 3 months and £10K on follow-up chemistry’??

- |    |  |   |
|----|--|---|
| 1. | <b>Prune solvent molecules and alternate sidechain conformations</b> | Delete those atoms and alternate conformations that are not present in the event map.   |
| 2. | <b>Fix conformations and rotamers that have changed</b>              | For residues where a sidechain or water has changed, simply correct the model as normal. Every residue that is moved in the model will lead to an alternate conformation when the ensemble model is constructed, so it is normally only necessary to model large changes from the reference model.    |
| 3. | <b>Model the ligand (if present) and add new solvent molecules.</b>  | Add new solvent molecules to the protein model where required. The ligand should be modelled in a preliminary location of it was supplied to PanDDA2. You can move it using standard COOT tools, and use ‘Mark Event as Interesting’ and ‘Ligand Placed’ to add the structure to the list for export. |
| 4. | <b>Save the changes to the model.</b>                                | use the “Save Model” or “Next Event >>> (Save Model)” button to have the model before progressing   |

# PanDDAs Tab: pandda.export



The screenshot shows the XChem Explorer interface with the 'PanDDAs' tab selected. A table of refinement results is visible on the left, and a configuration panel is on the right. A red box highlights the 'Export ALL PANDDA models' option in the 'Hit Identification' section of the bottom toolbar.

Refinement Space Group	Resolution n<math>\langle \text{sig(I)} \rangle = 1</math>	Dimple Rcryst	Dimple Rfree	Crystal Form Name	PanDDA launched?	PanDDA hit?	PanDDA reject?	PanDDA Status
I 1 2 1	n/a	0.20091	0.24170	SG-I121-No.5...	True	True	False	start
I 1 2 1	3.40	0.27638	0.36151	SG-I121-No.5...	True	False	False	start
I 1 2 1	n/a	0.19085	0.22944	SG-I121-No.5...	True	True	False	start
I 1 2 1	n/a	0.18977	0.23902	SG-I121-No.5...	True	False	False	start
I 1 2 1	n/a	0.18998	0.23253	SG-I121-No.5...	True	False	False	start
I 1 2 1	n/a	0.18312	0.24032	SG-I121-No.5...	True	False	False	start
I 1 2 1	n/a	0.18523	0.22781	SG-I121-No.5...	True	False	False	start
I 1 2 1	n/a	0.18656	0.23316	SG-I121-No.5...	True	True	False	start
I 1 2 1	n/a	0.18469	0.22959	SG-I121-No.5...	True	False	False	start
I 1 2 1	n/a	0.17369	0.22225	SG-I121-No.5...	True	False	False	start
I 1 2 1	n/a	0.23200	0.27408	SG-I121-No.5...	True	True	False	start
I 1 2 1	n/a	0.18925	0.24418	SG-I121-No.5...	True	False	False	start
I 1 2 1	n/a	0.18800	0.22815	SG-I121-No.5...	True	False	False	start
I 1 2 1	2.35	0.22608	0.28155	SG-I121-No.5...	True	False	False	start
I 1 2 1	n/a	0.17693	0.23415	SG-I121-No.5...	True	False	False	start
I 1 2 1	3.17	0.18876	0.25831	SG-I121-No.5...	True			
I 1 2 1	n/a	0.18837	0.22015	SG-I121-No.5...	True			
I 1 2 1	n/a	0.19610	0.23504	SG-I121-No.5...	True			
I 1 2 1	n/a	0.19576	0.23759	SG-I121-No.5...	True			

data directory: 17/lb16813-1/processing/analysis/initial\_model/\*  
pdb style: dimple.pdb  
mtz style: dimple.mtz  
output directory: /ta/2017/lb16813-1/processing/analysis/panddas  
submit: qsub  
number of processors: 7  
order events by: cluster\_size  
Use space group of reference file as filter: ...  
Expert Parameters (only change if you know what you are doing!):  
min\_build\_datasets

less pretty

**1** Select 'Export ALL PANDDA models' and click 'run'  
It will prepare the model (bound/unbound state) and refinement parameters, do a first round of refinement and create the ligand validation plot

les  
ource

**Datasets**  
Get New Results from Autoprocessing Run Status

**Maps & Restraints**  
Run DIMPLE on selected MTZ files Run Status

**Hit Identification**  
Export ALL PANDDA models Run Status

**Refinement**  
Open COOT Run Status

# PanDDAs Tab: pandda.export



## pandda.export

- “Export **NEW/ALL/SELECTED** PanDDA models”:
  - Generates an ensemble model of bound and ground states and launches refinement
  - Uses REFMAC for refinement
  - Generates occupancy and restraints parameters for reftmac and phenix
  - Ligand stats are calculated
- “Refine **ALL/NEW** bound-state models with BUSTER”:
  - Launches refinement of **bound-state only**
  - Useful for high occupancy ligands with single protein conformations
  - Can launch without sanity checks (“no sanity check”) but not recommended
    - If refinement job fails then check the buster log files to see why and fix

# Refinement Tab



XChemExplorer

File Datasource Preferences **Refinement** Help Labels

Overview Datasets Maps PANDDAs Refinement Deposition Settings

Sample ID	Compound ID	Refinement Space Group	Refinement Resolution	Refinement Rcryst	Refinement Rfree	Refinement Outcome	buster-reports	Ligand CC	Refinement Status	
1	MID2A-x0041	Z57101343	P 21 21 21	1.570	0.2301	0.2463	4 - CompChem ready	<a href="#">Refine 13-report</a>	LIG-B-801: 0.795	finished
2	MID2A-x0109	Z190780124	P 21 21 21	1.540	0.2292	0.2498	3 - In Refinement	<a href="#">Refine 10-report</a>	LIG-B-801: 0.789	finished
3	MID2A-x0112	Z45656995	P 21 21 21	2.340	0.2257	0.2699	3 - In Refinement	<a href="#">Refine 9-report</a>	LIG-A-711: 0.742	finished
4	MID2A-x0135	Z1134990241	P 21 21 21	2.456	0.2568	0.2938	3 - In Refinement	<a href="#">Refine 8-report</a>	LIG-A-711: 0.824	finished
5	MID2A-x0139	Z1129283193	P 21 21 21	1.830	0.2271	0.2561	3 - In Refinement	<a href="#">Refine 7-report</a>	LIG-A-801: 0.692	finished
6	MID2A-x0144	Z57472297	P 21 21 21	2.066	0.2575	0.2858	3 - In Refinement	<a href="#">Refine 8-report</a>	LIG-A-711: 0.666	finished
7	MID2A-x0145	Z1407672867	P 21 21 21	2.089	0.2399	0.2822	3 - In Refinement	<a href="#">Refine 8-report</a>	LIG-A-711: 0.760	finished
8	MID2A-x0152	Z1101755952	P 21 21 21	1.911	0.2472	0.2774	3 - In Refinement	<a href="#">Refine 8-report</a>	LIG-A-711: 0.830	finished
9	MID2A-x0155	Z56792776	P 21 21 21	1.759	0.2399	0.2644	3 - In Refinement	<a href="#">Refine 8-report</a>	LIG-A-711: 0.755	finished
10	MID2A-x0169	Z1367324110	P 21 21 21	2.141	0.2406	0.2776	3 - In Refinement	<a href="#">Refine 4-report</a>	LIG-A-711: 0.605	finished
11	MID2A-x0183	Z135439900	P 21 21 21	2.090	0.2568	0.2975	3 - In Refinement	<a href="#">Refine 2-report</a>	LIG-A-801: 0.695	finished
12	MID2A-x0184	Z1955122823	P 21 21 21	1.970	0.2334	0.2596	3 - In Refinement	<a href="#">Refine 8-report</a>	LIG-A-711: 0.894	finished
13	MID2A-x0208	Z19755216	P 21 21 21	1.810	0.2518	0.2791	3 - In Refinement	<a href="#">Refine 8-report</a>	LIG-A-711: 0.879	finished
14	MID2A-x0301	Z729726784	P 21 21 21	1.549	0.2227	0.2444	4 - CompChem ready	<a href="#">Refine 9-report</a>	LIG-A-4000: 0.782	finished
15	MID2A-x0328	Z133716556	P 21 21 21	1.629	0.2234	0.2498	4 - CompChem ready	<a href="#">Refine 11-report</a>	LIG-A-801: 0.653	finished
16	MID2A-x0361	Z2856434762	P 21 21 21	1.670	0.2267	0.2474	4 - CompChem ready	<a href="#">Refine 2-report</a>	None	finished
17	MID2A-x0393	Z1545196403	P 21 21 21	1.600	0.2162	0.2345	4 - CompChem ready	<a href="#">Refine 7-report</a>	LIG-B-801: 0.795	finished
18	MID2A-x0398	Z26968795	P 21 21 21	1.820	0.2212	0.2514	4 - CompChem ready	<a href="#">Refine 7-report</a>	LIG-A-4000: 0.697	finished
19	MID2A-x0401	Z2242056442	P 21 21 21	1.879	0.2335	0.2621	4 - CompChem ready	<a href="#">Refine 3-report</a>	None	finished
20	MID2A-x0419	Z32014663	P 21 21 21	1.610	0.2175	0.2456	4 - CompChem ready	<a href="#">Refine 7-report</a>	LIG-A-4000: 0.739	finished
21	MID2A-x0425	Z1827602749	P 21 21 21	1.710	0.2227	0.2501	4 - CompChem ready	<a href="#">Refine 5-report</a>	LIG-A-801: 0.86...	finished
22	MID2A-x0434	Z1217960891	P 21 21 21	1.770	0.2532	0.2789	4 - CompChem ready	<a href="#">Refine 4-report</a>	LIG-A-801: 0.93...	finished
23	MID2A-x0452	Z228585534	P 21 21 21	1.600	0.2143	0.2398	4 - CompChem ready	<a href="#">Refine 6-report</a>	LIG-A-4000: 0.597	finished
24	MID2A-x0453	Z375990520	P 21 21 21	1.570	0.2171	0.2374	4 - CompChem ready	<a href="#">Refine 4-report</a>	LIG-B-801: 0.866	finished
25	MID2A-x0455	Z1270312110	P 21 21 21	1.789	0.2356	0.2660	4 - CompChem ready	<a href="#">Refine 4-report</a>	LIG-B-801: 0.776	finished
26	MID2A-x0456	Z383202616	P 21 21 21	1.490	0.2133	0.2257	4 - CompChem ready	<a href="#">Refine 5-report</a>	LIG-A-4000: 0.6...	finished
27	MID2A-x0457	Z32014663	P 21 21 21	1.589	0.2164	0.2378	4 - CompChem ready	<a href="#">Refine 4-report</a>	LIG-B-801: 0.602	finished
28	MID2A-x0478	Z300245038	P 21 21 21	1.850	0.2292	0.2574	3 - In Refinement	<a href="#">Refine 5-report</a>	LIG-A-4000: 0.6...	finished
29	MID2A-x0482	Z647156496	P 21 21 21	1.960	0.2280	0.2611	4 - CompChem ready	<a href="#">Refine 5-report</a>	LIG-A-801: 0.795	finished
30	MID2A-x0484	Z1432018343	P 21 21 21	1.787	0.2312	0.2660	4 - CompChem ready	<a href="#">Refine 4-report</a>	LIG-B-801: 0.928	finished
31	MID2A-x0508	Z235361315	P 21 21 21	1.740	0.2394	0.2621	4 - CompChem ready	<a href="#">Refine 4-report</a>	LIG-A-4000: 0.617	finished
32	MID2A-x0513	Z369936976	P 21 21 21	1.689	0.2273	0.2555	4 - CompChem ready	<a href="#">Refine 5-report</a>	LIG-B-801: 0.482	finished
33	MID2A-x0525	Z381474098	P 21 21 21	1.540	0.2070	0.2175	4 - CompChem ready	<a href="#">Refine 5-report</a>	LIG-A-4000: 0.720	finished
34	MID2A-x0526	Z198195770	P 21 21 21	1.640	0.2189	0.2423	3 - In Refinement	<a href="#">Refine 3-report</a>	LIG-A-...	finished
35	MID2A-x0528	Z56827661	P 21 21 21	1.720	0.2217	0.2426	3 - In Refinement	<a href="#">Refine 3-report</a>	LIG-B-...	finished
36	MID2A-x0531	Z1343633025	P 21 21 21	1.689	0.2245	0.2534	4 - CompChem ready	<a href="#">Refine 4-report</a>	LIG-A-...	finished
37	MID2A-x0535	Z65532537	P 21 21 21	1.880	0.2691	0.3102	4 - CompChem ready	<a href="#">Refine 5-report</a>	LIG-A-...	finished
38	MID2A-x0541	Z2856434865	P 21 21 21	1.769	0.2130	0.2446	4 - CompChem ready	<a href="#">Refine 5-report</a>	LIG-A-...	finished
39	MID2A-x0546	Z2856434829	P 21 21 21	1.540	0.2086	0.2327	4 - CompChem ready	<a href="#">Refine 5-report</a>	LIG-A-...	finished
40	MID2A-x0547	Z364328788	P 21 21 21	1.921	0.3193	0.3367	4 - CompChem ready	<a href="#">Refine 5-report</a>	LIG-B-...	finished
41	MID2A-x0549	Z26968795	P 21 21 21	1.510	0.2015	0.2241	4 - CompChem ready	<a href="#">Refine 5-report</a>	LIG-B-...	finished
42	MID2A-x0550	Z364321922	P 21 21 21	1.771	0.2161	0.2410	4 - CompChem ready	<a href="#">Refine 4-report</a>	LIG-B-801: 0.720	finished
43	MID2A-x0555	Z1449748885	P 21 21 21	1.570	0.2198	0.2481	4 - CompChem ready	<a href="#">Refine 4-report</a>	LIG-A-4000: 0.282	finished
44	MID2A-x0563	Z2856434918	P 21 21 21	1.620	0.2144	0.2391	4 - CompChem ready	<a href="#">Refine 4-report</a>	LIG-A-801: 0.920	finished
45	MID2A-x0564	Z1003207278	P 21 21 21	1.390	0.2090	0.2284	4 - CompChem ready	<a href="#">Refine 4-report</a>	LIG-A-801: 0.75...	finished

Update Tables From Datasource

Datasets Get New Results from Autoprocessing Run Status

Maps & Restraints Run DIMPLE on selected MTZ files Run Status

Hit Identification pandda.analyse Run Status

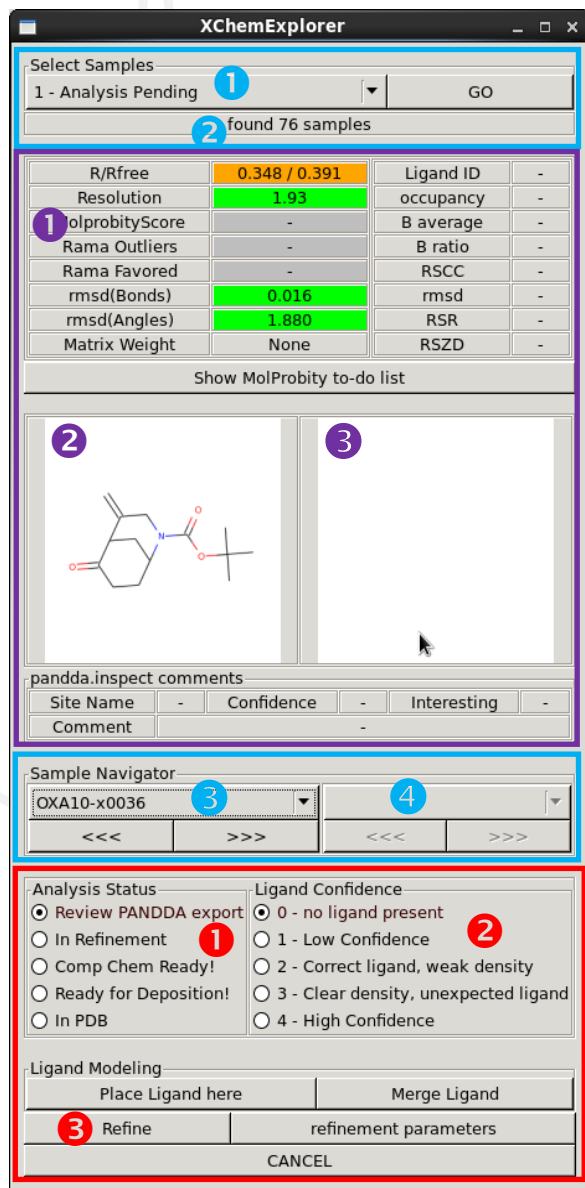
Refinement Open COOT - REFMAC refinement - Run Status

idle

1 Select 'Open coot' and click 'run'  
It will open coot and an the XCE refinement control panel



# Refinement Tab



Select Samples  
1 - Analysis Pending GO

found 76 samples

R/Rfree	0.348 / 0.391	Ligand ID	-
Resolution	1.93	occupancy	-
MolprobityScore	-	B average	-
Rama Outliers	-	B ratio	-
Rama Favored	-	RSCC	-
rmsd(Bonds)	0.016	rmsd	-
rmsd(Angles)	1.880	RSR	-
Matrix Weight	None	RSZD	-

Show MolProbity to-do list

2

3

pandda.inspect comments

Site Name	-	Confidence	-	Interesting	-
Comment	-				

Sample Navigator

OXA10-x0036

<<< >>> <<< >>>

Analysis Status

Review PANDDA export

In Refinement

Comp Chem Ready!

Ready for Deposition!

In PDB

Ligand Confidence

0 - no ligand present

1 - Low Confidence

2 - Correct ligand, weak density

3 - Clear density, unexpected ligand

4 - High Confidence

Ligand Modeling

Place Ligand here Merge Ligand

Refine refinement parameters

CANCEL

1 Select the category/status of samples you want to refine (at the beginning: **3 – in refinement**) and click ‘GO’

2 It will tell you how many samples were found for that category

3 To navigate through the samples in the selected category

4 To select the event of interest

*N.B* - XCE has already run on cycle of refinement straight after pandda.export

1 Summary of refinement statistics

2 & 3 are currently unavailable

1 Manually change the status of a model:

“**In Refinement**” – currently being refined

“**Comp Chem Ready!**” - Ligand and binding site refined, ready for interpretation, some atoms to refine elsewhere may remain.

“**Ready for Deposition!**” – drawn into any deposition actions

2 Manually select the ligand confidence for **this event**

3 Launch a refinement of the current model (plus other options)

‘**Comp chem ready**’ structure can be shared with your chemist to start follow-up work.

# PDB Group Deposition



- We can deposit XChem fragment structures to the RCSB in a single group
- Models and integrated data are deposited as .mmcif files
- Instructions are contained within the XCE interface – the process is still clunky so some manual file edits may be necessary, but your local contact should be able to help

# References



## XChem Explorer

Krojer, T., *et al.* **The XChem Explorer graphical workflow tool for routine or large-scale protein-ligand structure determination.** *Acta Cryst D*, **73**, 267-278 (2017). <https://doi.org/10.1107/S2059798316020234>

## PanDDA

Pearce, N., *et al.* **Partial-occupancy binders identified by the Pan-Dataset Density Analysis method offer new chemical opportunities and reveal cryptic binding sites.** *Structural Dynamics*, **4**, 032104 (2017). <https://doi.org/10.1063/1.4974176>

Pearce, N., *et al.* **A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density.** *Nat. Commun.*, **8**, 15123 (2017). <https://doi.org/10.1038/ncomms15123>

[https://github.com/ConorFWild/pandda\\_2\\_gemmi](https://github.com/ConorFWild/pandda_2_gemmi)

## XChem pipeline overview

Douangamath, A., *et al.* **Achieving Efficient Fragment Screening at XChem Facility at Diamond Light Source.** *JoVE journal* (2021). <https://www.jove.com/t/62414/achieving-efficient-fragment-screening-at-xchem-facility-at-diamond>

## XChem Bulletin Board

<https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=XCHEMBB>

