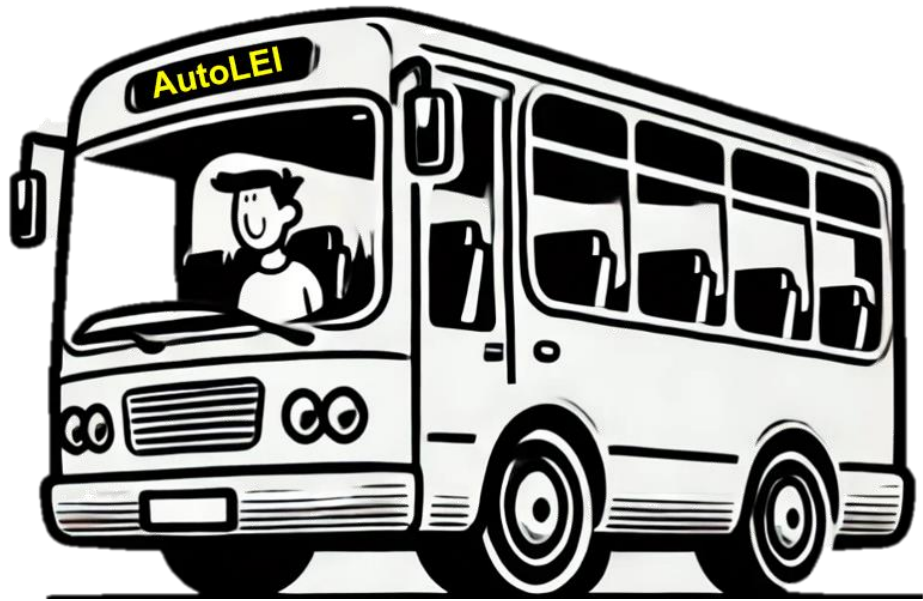


Station: **HKL**
0:00-24:00

Nästa: Station HKL



AutoLEI

Automated Real-time and Offline Batch
3D ED/MicroED Data Processing

Tutorial for AutoLEI

Lei Wang and Yinlin Chen,
Regarding AutoLEI version 1.0.2

Tutorial for AutoLEI	1
1. Installation.....	3
1.1 Requirements	4
1.2 Installation	4
2. Functions through command window	6
3. Functions through GUI	8
3.1 Input Tab.....	8
3.2 XDSRunner Tab	11
3.3 CellCorr Tab	13
3.4 XDSRefine Tab	14
3.5 DataMerge Tab	19
3.6 Cluster_Output Tab.....	20
3.7 Expert Tab	22
3.8 RealTime Tab	24
4. Tutorial Example 1: L-Tyrosine.....	28
4.1 General information.....	28
4.2 Goals	28
4.3 Load Working Directory.....	28
4.4 Batch Processing Datasets in P1 Mode.....	29
4.5 Batch Processing Datasets in Cell Mode	31
4.6 Refine Input Parameter	32
4.7 Finalise Data Reduction.....	33
5. Tutorial Example 2: CAU-36.....	36
5.1 General information.....	36

5.2 Goals	36
5.3 Load Working Directory.....	36
5.4 Batch Processing Datasets in P1 Mode.....	37
5.5 Batch Processing Datasets in Cell Mode	40
5.6 Refine Input Parameter	41
5.7 Merge Data	42
5.8 Intensity-Cluster	43
5.9 Generate and Update INS File	44
6. Tutorial Example 3: SU-100	47
6.1 General information.....	47
6.2 Goals	47
6.3 Load Working Directory.....	47
6.4 Batch Processing Datasets in P1 Mode.....	48
6.5 Batch Processing Datasets in Cell Mode	50
6.6 Refine Input Parameter	51
6.7 Merge Data	52
6.8 Intensity-Cluster	53
6.9 Generate and Update INS File	54
7. Tutorial Example 4: Lysozyme.....	58
7.1 General information.....	58
7.2 Goals	58
7.3 Load Working Directory.....	58
7.4 Batch Processing Datasets in P1 Mode.....	59
7.5 Batch Processing Datasets in Cell Mode	62
7.6 Merge Data	63

1. Installation

Obtain the original code from **Gitlab**, **PyPI** or **Zenodo** release folder and save it in a separate directory. Ensure that all necessary scripts are present in the directory. The complete list of scripts is provided below.

```
.
├── CHANGELOG.md
├── LICENSE
├── MANIFEST.in
├── README.md
├── autolei
│   ├── instrument_profile
│   │   └── [instrument_profiles]
│   ├── qt
│   │   └── style.qss
│   └── src
│       ├── __init__.py
│       ├── XDSINP
│       ├── analysis_hkl.py
│       ├── cif_io.py
│       ├── extinction.json
│       ├── file_io
│       │   └── folder_move_manager.py
│       ├── generate_pets.py
│       ├── image
│       │   ├── __init__.py
│       │   ├── beam_centre.py
│       │   ├── beam_stop.py
│       │   ├── generate_pets.py
│       │   ├── mrc.py
│       │   ├── nxs.py
│       │   └── tiff.py
│       ├── image_io.py
│       ├── symm_shelx
│       │   ├── __init__.py
│       │   ├── make_shelx_ins.py
│       │   ├── space_group_finder.py
│       │   ├── symmetry_operation.py
│       │   └── symmetry_ops_shelx.py
│       ├── symmetry_operation.py
│       ├── util.py
│       ├── visualisation
│       │   ├── __init__.py
│       │   ├── html_report.py
│       │   ├── recip_viewer.py
│       │   └── recip_viewer_GUI.py
│       ├── xds_analysis.py
│       ├── xds_cluster.py
│       ├── xds_input.py
│       ├── xds_refine.py
│       ├── xds_refiner
│       │   ├── __init__.py
│       │   ├── beam_centre.py
│       │   └── rotation_axis.py
│       ├── xds_report.py
│       ├── xds_runner.py
│       └── xds_shelx.py
├── __init__.py
├── _run.bat
├── autolei.py
├── autolei_core.py
├── realtime.py
├── setting.ini
├── tool.py
└── pyproject.toml
```

1.1 Requirements

AutoLEI is developed in Python and is compatible exclusively with UNIX-like environments, including **Linux** and **macOS**. For **Windows** users, it is necessary to install **WSL 2** or an X-service if using **WSL 1**.

Important: AutoLEI does not include the XDS suite. You must manually install and configure XDS-related software from the [official XDS website](#). It is highly recommended to install **xdsgui** as well. Additionally, ensure that **XPREP** (available for both Linux and Windows) is added to your system's PATH if you intend to use it during data processing.

1.2 Installation

AutoLEI requires several Python packages. All dependencies are available via pip. Besides, for operation system support of qt components, the necessary shared library should be installed as well. For users operating in a **pure Linux** environment, it is also recommended to install **LibreOffice** to fully utilise AutoLEI's features. To install pip, and LibreOffice, follow these instructions for your respective operating system:

For Debian-based system (Ubuntu, Debian):

```
sudo apt-get update
sudo apt-get install python3-pip qtbase5-dev
sudo apt-get install libreoffice
```

For Fedora / RedHat:

```
sudo dnf update
sudo dnf install qt5-qtbase python3-pip
sudo dnf install libreoffice
```

For CentOS / RHEL:

```
sudo yum install epel-release
sudo yum update
sudo yum install qt5-qtbase python3-pip
sudo yum install libreoffice
```

For Arch Linux / Manjaro

```
sudo pacman -Syu
sudo pacman -S qt5-base python-pip
sudo pacman -S libreoffice-still
```

1.2.1 Install dependencies via pip package

The whole package will be available on PyPI with the name **autolei**. The simplest way to install the package is through command

```
pip install autolei
```

You can install or update the entire package and its dependencies using pip, the package installer for Python. To do this, follow these steps:

- Download the distribution file and place it in your desired directory.

- Open the commander or a WSL terminal if on Windows and navigate to the directory containing the distribution file.

- Run the installation command:

```
pip install ./autolei-1.0.2.tar.gz
```

- After installation, you can use AutoXDSGUI from any directory by typing:

```
autolei
```

It can be uninstalled by

```
pip uninstall autolei
```

Figure 1 User interface of AutoLEI

2. Functions through command window

➤ **autolei_setting:**

This command opens AutoLEI's configuration file using the nano editor. The configuration is divided into several sections, each governing various aspects of the software:

- **General** – window_scaling / set_scaling:

Controls application scaling on high-DPI screens. Set window_scaling to False or set_scaling to 1.0 to disable scaling.

- **General** – multi-process / max_core:

Enables/disables multi-threading and sets the maximum number of cores for data processing.

- **General** – xprep_location:

Add the xprep location if it is not in environmental variable.

- **General** – is_MM:

If this setting is False, when the unit cell is larger than 300,000 Å³, the autolei's data analysis will be skipped on this dataset. This will lead the statistics inaccurate and will disable the web_report function. However, when XDS gives back some huge unit cell on small molecule structures, disable this function will enhance the speed.

- **General** – path_filter:

When True, AutoLEI will exclude directories starting with “!” and “.” from data reduction to protect invisible/system folders.

- **XDSInput** – max_core/max_job:

MAXIMUM_NUMBER_OF_JOBS and MAXIMUM_NUMBER_OF_PROCESSORS in XDS.INP. Setting both to 99 allows XDS to determine optimal values.

- **XDSInput** – use_relative_path:

Uses relative paths for image locations in XDS.INP. This may affect the display of indexed points in XDSGUI. Modify image paths via the Expert tab if necessary.

- **XDSRunner** –use_short_path:

When True, result files from batch runs use relative paths for dataset locations.

- **XDSRunner** – engine_hkl_analysis:

Switches between AutoLEI (“SU”) and XDS (“XDS”) for HKL analysis. Using AutoLEI provides additional resolution range information and suggested resolutions, while XDS may reduce statistical accuracy and disable web_report but may increase the processing speed. The option is also related to is_MM.

- **XDSRunner** – rerun_failed / rerun_strategy :

Activates the rerun_failed option to automatically retry failed datasets based on the specified rerun_strategy (“basic” or “advanced”).

- **Cluster** – cell_cluster_distance:

For the estimate_symmetry function on XDSRunner Page, it will try to estimate the lattice symmetry and corresponding Laue group of the structure. Then with the given unit cell, a unit cell clustering will be performed by a newly introduced mixed matrix method. The threshold of clustering will be given here, with its default and suggested value of 0.20.

- **Cluster** – cell_cluster_symmetry:

If True, splits clusters based on Bravais lattice and Laue group after unit cell clustering.

- **Inp_Refine:**

Parameters in this section will be appeared as the default value in the XDSRefine Page.

➤ **autolei_add_instrument:**

usage:

```
autolei_add_instrument [instrument_profile_file]
```

Adds an instrument profile to AutoLEI. The profile file must be in JSON format and include necessary details such as instrument type, detector specifications, and additional information.

```
{
  "instrument": "Titan Krios G3i",
  "detector": "Ceta-D",
  "NX": 1024,
  "NY": 1024,
  "QX": 0.056,
  "QY": 0.056,
  "overload": 65000,
  "rotation_axis": 6.0,
  "energy": 300,
  "additional_information": [
    "SIGNAL_PIXEL=7"
  ]
}
```

➤ **autolei_mrc2pets:**

usage:

```
autolei_mrc2pets root_dir [-o] [-m merge_number] [-s save_folder]
```

This function only supports FEI mrc. Options:

- root_dir: the working directory. This function will go through all subfolder contained non-stacked MRC and transferred them into TIFF. An automated PETS will be generated in the same folder of MRC files.
- -o or --overwrite: overwrites existing TIFF images if present.
- -m or --merge: merge multiple frames into one, default 1 (no merging).
- -s or --save_subfolder: specifies the subfolder which contains the tiff images, default “tiff”.

3. Functions through GUI

AutoLEI features two categories of tabs: **Core Tabs** and **Extended Tabs**. The **Core Tabs** include essential functionalities for the basic XDS workflow, while the **Extended Tabs** offer additional tools and features.

Core Tabs

The core tabs encompass the following components, which facilitate the primary XDS workflow:

1. **Input**
Handles data input and preparation.
2. **XDSRunner**
Executes the main XDS processing pipeline in P1 mode.
3. **CellCorr**
Performs unit cell corrections and adjustments.
4. **XDSRefine**
Refines XDS parameters for improved results.
5. **DataMerge**
Combines datasets for enhanced data quality.
6. **Cluster&Output**
Intensity-cluster the datasets and output refinement information.

Extended Tabs

The **Extended Tabs**, such as **Expert** and **Realtime**, provide supplementary functionalities to enhance user experience and expand the software's capabilities. These tabs may include advanced features or user-designed modules.

Customization and API Support

Starting with version 1.1.0, AutoLEI will introduce an open API, empowering users and developers to design and integrate custom tabs. This flexibility allows for tailored workflows and personalised functionality within the extended tabs.

Below, we explore the features and roles of the core tabs and highlight some official extended tabs available in AutoLEI.

3.1 Input Tab

The Input Tab serves as the starting point for setting up and configuring the working directory and instrument information. It allows users to:

- Load the working directory.
- Summarise instrument information (when an XDS.INP file is present).

- Input basic information about instruments and measurements (when an XDS.INP file is not available).

Features and Functionality

Hover over any input field, button, or drop-down menu to view tooltips explaining their specific functions. Most input fields in the Input Tab correspond to XDS input parameters. For detailed explanations, refer to https://xds.mr.mpg.de/html_doc/xds_parameters.html. The following sections provide details on input fields, buttons, and drop-down menus:

Figure 3.1 Input Tab

Input Fields

Fields (1)–(10) relate to instrument information and rarely need adjustment between sessions or measurements. Fields (11)–(15) are session- and material-dependent. For stable electron microscopes, these values can often remain consistent. However, it is recommended to verify and correct these values using image header data or tools in the XDSRunner Tab.

1. **Working directory:** Specify the Linux-format path to the working directory containing (or expected to contain) electron diffraction data. The program requires that the path be accessible and not protected.
2. **Pixel number (x-axis):** Number of pixels along the x-axis (fast direction).
3. **Pixel number (y-axis):** Number of pixels along the y-axis (slow direction).
4. **Pixel size (x-axis):** Pixel size along the x-axis (fast direction) in mm.
5. **Pixel size (y-axis):** Pixel size along the y-axis (slow direction) in mm.

6. **Pixel overload value:** Maximum pixel intensity; values exceeding this are ignored. Note: SMV files have a maximum intensity of 65535, so values beyond this limit are invalid.
7. **Wavelength:** Beam wavelength in Å.
8. **Rotation axis direction cosines:** Space-separated cosines for the rotation axis in the laboratory coordinate system (right-handed rotation). This is usually constant for a given instrument.
9. **Rotation axis angle:** Angle between the rotation axis projection and the x-axis, measured counterclockwise. Use only if field 8 is left empty.
10. **Additional information:** Specify beam stop details, untrusted detector areas, or other keywords. Note that AutoLEI does not validate these inputs, so users must ensure their correctness.
11. **Beam centre (x-axis):** Beam centre along the x-axis (fast direction). This can be corrected later using the “Beam Centre Finding” or “Beam Stop Finding” tools in the XDSRunner Tab.
12. **Beam centre (y-axis):** Beam centre along the y-axis (slow direction). This can also be corrected later using the tools in the XDSRunner Tab.
13. **Resolution range:** Space-separated values defining the lowest and highest resolutions for XDS. The default value is “30 0.8”.
14. **Detector distance:** Distance from the detector to the sample (camera length) in mm. Can be corrected later with “Correct Input with Metadata” in the XDSRunner Tab.
15. **Oscillation angle:** Oscillation angle of a single frame, measured in degrees. This value can also be corrected later using the “Correct Input with Metadata” in the XDSRunner Tab.

Drop-Down Menu

The drop-down menu allows users to:

- Select default instrument profiles implemented in AutoLEI as a starting point for generating the XDS.INP file then.
- Use the “Browse” option to select an existing XDS.INP file as a model.
- To add a new instrument to this menu, refer to Section 2 and use the **autolei_add_instrument** command.

Buttons:

- **Browse:** Opens a file dialog to select the working directory. Note that the directory loaded will be the one opened in the dialog, not the one highlighted.
- **Load Path:** Loads the specified working directory. Invalid characters or spaces in the path may affect AutoLEI’s functionality. A message box confirms successful loading.
- **Load:** Reads and loads the instrument profile selected in the drop-down menu into input fields (1)–(13).
- **Save Parameter:** Generates an Input_parameter.txt file, which is required to create the XDS.INP file then. Ensure all required fields are filled (except for fields 10). For fields 8 and 9 (rotation axis), at least one of them must be entered. If an XDS.INP file already exists, this step is unnecessary.

3.2 XDSRunner Tab

The XDSRunner Tab is the most complex among all core tabs. It provides tools for advanced data processing and correction. Its main functions include:

- Image conversion.
- Instamatic XDS.INP correction (for XDS.INP files from Instamatic).
- Generating and correcting input files for beam centre/stop and measurement parameters (pixel size, camera length, etc.).
- Running data reduction in *P1* mode.
- Viewing data reduction results.
- Estimating symmetry for preparation on cell run.

Features and Functionality

Hover over any radio button, checkbox, or button to view tooltips explaining their specific functions. The following details describe the functionality of radio buttons, checkboxes, and buttons in this tab:

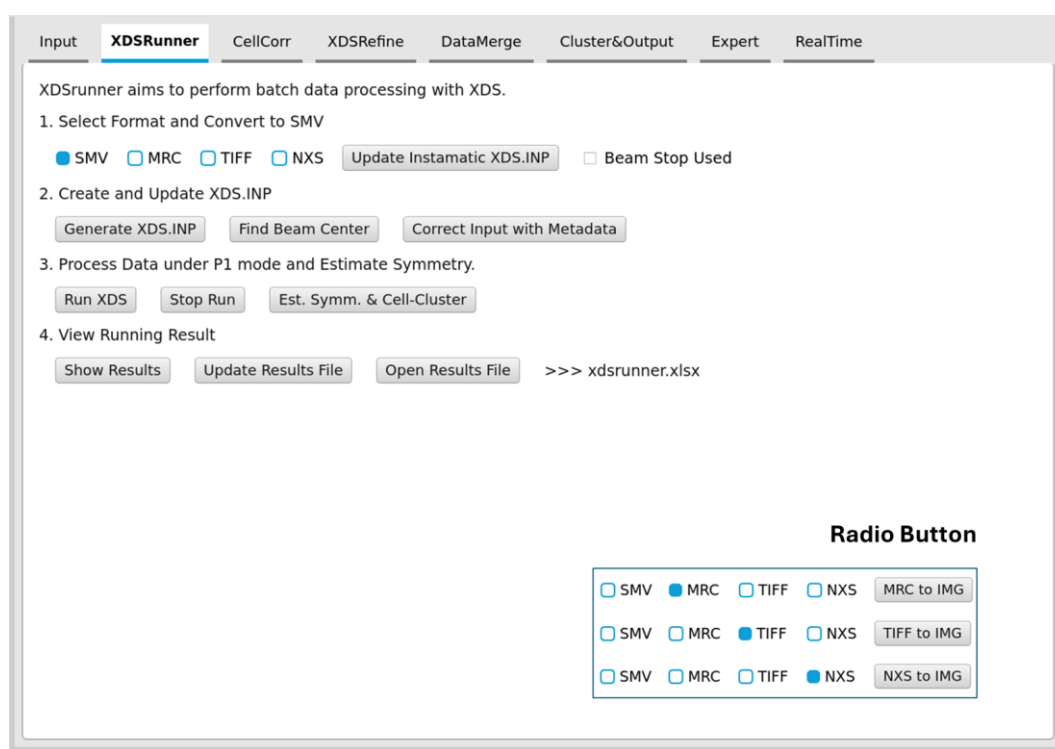


Figure 3.2 XDSRunner Tab

Radio Buttons:

Radio buttons allow users to select the raw data format of the diffraction image. AutoLEI primarily supports the SMV format, and other formats must be converted to SMV format for full functionality. In version 1, the following formats are supported, with more planned for future versions:

- **SMV**: The corresponding function button is “Update Instamatic XDS.INP.”

- **MRC:** The corresponding function button is “MRC to IMG.”
- **TIFF:** The corresponding function button is “TIFF to IMG.”
- **NXS:** The corresponding function button is “NXS to IMG.”

Checkbox:

Beam Stop Used: When checked the button function for “Find Beam Centre” changes to “Find Beam Stop”. Only check this box if a beam stop was used during the measurement.

Buttons:

1. **Update Instamatic XDS.INP:** Updates the XDS.INP file generated by Instamatic versions < 1.1, removing illegal characters and obsolete keywords for compatibility with the latest XDS versions.
2. **MRC to IMG:** Converts FEI .mrc files into SMV .img files. Requirements:
 - .mrc files must be non-stacked (separated) and contain more than 10 images.
 - Metadata must be present in the extended header.
 - Converted .img files will be saved in the same folder as the original .mrc files.
3. **TIFF to IMG:** Converts .tiff files into SMV .img files. Since TIFF files lack essential metadata, users will be prompted to manually input the required information. Details:
 - Required information includes wavelength, pixel size, starting angle, ending angle, and camera length.
 - If Input_parameter.txt exists, it will pre-load wavelength, pixel size, and camera length values.
 - Start and end angles can be auto-loaded if a metadata.txt file exists in the same folder with keywords like START_ANGLE=, END_ANGLE=, or OSCILLATION_RANGE= (at least two required).
 - Missing frame names in XDS will result in empty placeholder images being generated.
 - Converted .img files will be saved in the same folder as the original .tiff files.
4. **NXS to IMG:** Converts DLS .nxs files into SMV .img files. Requirements:
 - Associated .h5 files must be present in the same folder as the .nxs files.
 - Metadata must be included in the file headers.
 - Converted .img files will be saved in the same folder as the original .nxs files.
5. **Generate XDS.INP:** Generates XDS.INP in the xds subfolder based on the saved Input_parameter.txt. SMV .img files that already have XDS.INP will not be generated again.
6. **Find Beam Centre:** Finds the most intensive region in the image and marks its centre as the beam origin point.
7. **Find Beam Stop:** Activated if the “Beam Stop Used” checkbox is selected. This function uses diffuse scattering of the direct beam to locate the beam centre and the shadow of the beam stop. It also marks the beam stop location as an untrusted area.
8. **Correct Input with Metadata:** Reads header information from the SMV .img file and corrects pixel size, starting angle, oscillation range (single frame), wavelength, and camera length. If certain metadata is missing, the function will skip those corrections.

9. **Run XDS:** Runs XDS batch processing on all XDS.INP files in the working directory and its subfolders. Two additional files, XDS.LP and STATISTICS.json, are generated upon successful processing.
 - XDS.LP logs the processing details showing on screen.
 - STATISTICS.json collects necessary data reduction information.
 The auto-re-run function, which can be open/close in the setting, can automatically re-run the failed datasets with set strategies in AutoLEI.
10. **Stop Run:** Terminates the XDS runner after completing the current XDS file.
11. **Est. Symm. & Cell-Cluster:** After running XDS, autolei calculates lattice symmetry and evaluates R_{meas} and $CC_{1/2}$ for the corresponding Laue group. This function will cluster the data based on the lattice with the highest symmetry with an innovative method. Results are displayed in the command window and saved in lattice_cluster.txt.
12. **Show Results:** Show the summarised results of the batch processing and display them below the button.
13. **Update Results File:** Summarises successfully reduced results into “xdsrunner.xlsx” by extracting data from STATISTICS.json. If the file is not found, the function will analyse the results, which may take a few seconds. The results file will be automatically updated after running XDS batch processing.
14. **Open Result File:** Opens xdsrunner.xlsx using Excel (on WSL) or LibreOffice (on Linux). The Excel file contains additional detailed information.

3.3 CellCorr Tab

After the first round of XDS processing, the resulting files are generated. However, the unit cell and space group provided by XDS may not always be accurate. The CellCorr Tab allows users to:

- Re-run data reduction with specified unit cell parameters and space group.

Features and Functionality

Hover over any input fields, or button to view tooltips explaining their specific functions. The following details describe the functionality of input fields, and buttons in this tab:

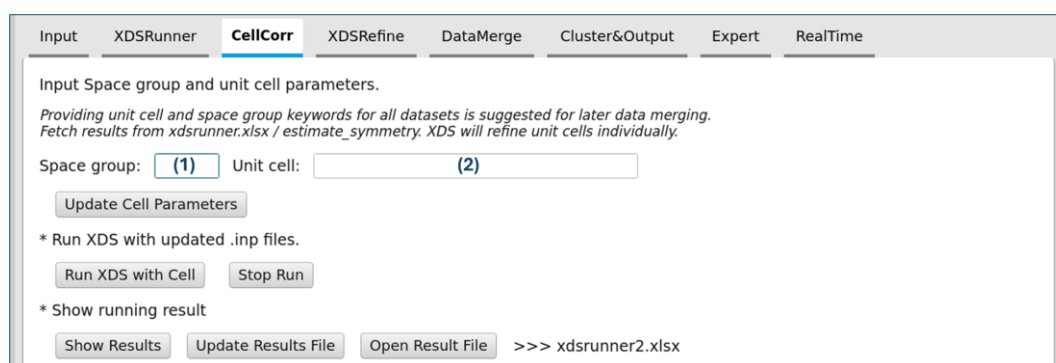


Figure 3.3 CellCorr Tab

Input Fields

1. **Space Group:** Enter the space group number (1 to 230). XDS will automatically recognise different lattice settings for the same number and convert them to the standard setting.
2. **Unit Cell:** the cell parameters in order of a, b, c (in Å) and α , β , γ (in °). Valid input formats include:
 - 3.0, 4.0, 5.0, 90.0, 110, 90
 - 3.0 4.0 5.0 90.0 110 90
 - 3,0 4,0 5,0 90,0 110 90

Buttons:

1. **Update Cell Parameters:** Modifies all XDS.INP files in the working directory with the specified unit cell and space group parameters. A backup file named BACKUP-P1 is created. Backup files can be restored using functions in the Expert tab.
2. **Run XDS with Cell:** Runs XDS batch processing on all XDS.INP files in the working directory and its subfolders with the modification. Two additional files, XDS.LP and STATISTICS.json, are generated upon successful processing.
 - XDS.LP logs the processing details showing on screen.
 - STATISTICS.json collects necessary data reduction information.

The auto-re-run function (enabled/disabled in settings) will automatically retry failed datasets using predefined strategies in AutoLEI.
3. **Stop Run:** Terminates the XDS runner after completing the current XDS file.
4. **Show Results:** Show the summarised results of the batch processing and display them below the button.
5. **Update Results File:** Summarises successfully reduced results into **xdsrunner2.xlsx** by extracting data from STATISTICS.json. If the file is not found, the function will analyse the results, which may take a few seconds. The results file will be automatically updated after running XDS batch processing.
6. **Open Result File:** Opens **xdsrunner2.xlsx** using Excel (on WSL) or LibreOffice (on Linux). The Excel file contains additional detailed information.

3.4 XDSRefine Tab

With a given unit cell, results are closer to structure solution. However, due to potential instability of the goniometer and beam, results can often be improved by refining input parameters. For datasets with sufficient completeness and quality, users can also generate output files such as hkl, p4p, and cif_od. The XDSRefine Tab provides functionality to:

1. Refine input parameters in XDS.INP, including:
 - Rotation axis
 - Beam centre (if using beam stop)
 - Divergence and mosaicity
 - Scale outliers
 - Index error

- Resolution
- 2. Batch update or mute input parameters
- 3. Calibrate the camera length using a specified ratio
- 4. Re-run data reduction with specified unit cell parameters and space group
- 5. Utilise tools specific to single datasets or failed datasets:
 - Run XDSGUI
 - View reconstructed reciprocal space
 - Open the data folder
 - Generate SHELX hkl, p4p, and cif_od files
 - View a web report for data reduction

Features and Functionality

Hover over any input field, drop-down menu, checkbox, or button to view tooltips explaining their specific functions. Below are details on the features available in this tab:

Drop-down menu:

1. **Data Selection [I]:** Options to select datasets for re-running data reduction:
 - **All:** Includes all datasets.
 - **Selected:** Filters data by specific criteria.
 - **Ranged:** Limits data to a specified numbering range.
 - **Single:** Focuses on a single dataset.
 - **Failed:** Targets datasets that failed.

Additional functions are available for **Single** and **Failed** modes.
2. **Data Filter [II, III]:** Appears when “Selected” is chosen in Data Selection.
 - First menu: Criteria options (Index%, I/Sigma, CC1/2, or Resolution).
 - Second menu: Comparison operators (“<” or “>”).
3. **Single Data Selector [IV]:** For “Single” mode, displays available datasets by name and number.
4. **Failed Data Selector [V]:** For “Failed” mode, displays failed datasets by name and number.

Input Fields

1. **Keyword Manager – Add – Keyword:** Supports autocomplete for entering keywords. Keywords unavailable in this field can be modified using XDSGUI instead.
2. **Keyword Manager – Add – Value:** Specifies the value for a corresponding keyword.
3. **Keyword Manager – Delete – Keyword:** Similar to the add tab. Instead of simply deleting, AutoLEI comments out the parameter using “!”.

Input XDSRunner CellCorr **XDSRefine** DataMerge Cluster&Output Expert RealTime

Refine Input Parameters in XDS.INP. Get data reduction result from single dataset.
Use AutoLEI to refine XDS.INP files in the target folder.

Refine on data as **All** **[I]** **(8)**

☐ Rotation Axis ☐ Divergence & Mosaicity ☒ Remove Scale Outlier > **2.0** IQR ☐ Beam Centre

☐ Refine Index Ratio on datasets with index% < **85.0** **(9)** ☐ Change Resolution to **30 0.8** **(10)**

Run XDS with Cell Stop Run | Change Input Parameters |

Show Results Update Results File Open Results File |

Keyword Manager
Add, Replace or Delete Keywords in XDS.INPs.
The Keyword Entry has AutoComplete Function.

Add Delete Calibrate

Add Row Delete Row Reset

(1) **(2)** **(3)** **(4)** **(5)**

Force Replace Save Cancel

Refine on data as **Selected** with **--** **[I]** **[II]** **[III]** **(6)**

Refine on data as **Ranged** **(7)** >> Example: 1, 2-4

Refine on data as **Single** on **--** **[IV]**

Refine on data as **Failed** on **--** **[V]**

Run XDS with Cell Stop Run | Change Input Parameters | Run XDSGUI View Reciprocal Space

Show Results Update Results File Open Results File | Bus to SHELX Web Report Open Folder

Figure 3.4 XDSRefine Tab

4. **Keyword Manager – Calibrate – Universe Ratio:** Specifies a calibration ratio. Ratios >1 indicate the measured unit cell is larger than the ideal one. Camera length is calculated as input length ÷ ratio.
5. **Keyword Manager – Calibrate – Distance Ratio:** Converts displayed camera length to the normal reported length on the instrument.
6. **Scale Outlier Ratio:** Threshold for filtering outliers. Default: 2.0 (modifiable in settings).
7. **Index% Threshold:** Prevents index error refinement for datasets exceeding this percentage. Default: 85.0 (configurable).
8. **Resolution Range:** Defines the resolution range (lowest to highest) for XDS. Default: 30 0.8.

Checkboxes:

- **Rotation Axis:** If checked, AutoLEI will optimise the rotation axis ($\pm 10^\circ$) based on the peaks identified in SPOT.XDS. The optimization algorithm is adapted from edtools. To refine the rotation axis, XDS must at least complete the COLSPOT step.

- **Divergence & Mosaicity:** If checked, AutoLEI will copy the divergence and mosaicity values from INTEGRATE.LP. To refine these parameters, XDS must at least complete the INTEGRATE step.
- **Scale Outlier Frame:** If checked, AutoLEI will remove frames with scale factors outside the specified range. These frames often result from issues such as crystal displacement, movement out of the selected area, or image mode (e.g., in instamatic). To remove scale outlier frames, XDS must at least complete the INTEGRATE step.
- **Beam Centre:** If checked, AutoLEI will optimise the beam centre using Friedel's pairs in SPOT.XDS, without relying on indexed results. This feature is particularly useful when a beam stop is used. To refine the beam centre, XDS must at least complete the COLSPOT step.
- **Index Ratio:** If checked, AutoLEI will optimise the indexation result using the INDEX_ERROR keyword. This can be helpful for unstable electron microscopy goniometers. To refine the index ratio, XDS must at least complete the COLSPOT step.
- **Resolution Range:** If checked, AutoLEI will rewrite the resolution range keyword in the corresponding XDS.INP file.
- **Keyword Manager – Force Replace:** If checked, AutoLEI will overwrite all existing values for multi-value keywords (e.g., JOB, EXCLUDE_RESOLUTION_RANGE) in XDS.INP and replace them with new values. If unchecked, new values will be added after the existing ones for multi-value keywords.

Buttons:

1. **Run XDS with Cell:** AutoLEI first attempts to refine the input parameters in the provided XDS.INP files. It then executes XDS batch processing using the successfully refined parameters. Upon successful processing, two additional files are generated:
 - XDS.LP: Logs the processing details displayed on the screen.
 - STATISTICS.json: Collects necessary data reduction information.
 Additionally, the Auto-Re-run function (enabled/disabled in settings) automatically retries failed datasets using predefined strategies in AutoLEI.
2. **Stop Run:** Terminates the XDS runner after completing the current XDS file.
3. **Change Input Parameters:** If one or more XDS.INP files are selected, this button opens the Keyword Manager. Within the Keyword Manager, you can:
 - Add new keyword/value pairs.
 - Modify existing keywords.
 - Batch change the camera length based on a standard sample.
 The pop-out window contains three tabs: Add, Delete, and Calibrate. Note that Calibration mode cannot be used simultaneously with Keyword Editing mode.
4. **Keyword Manager – Add/Delete – Add Row:** Adds a new row below the existing rows, with a maximum of eight rows allowed.
5. **Keyword Manager – Add/Delete – Delete Row:** Deletes the last empty row.

6. **Keyword Manager – Add/Delete – Reset:** Resets the tab to its initial state.
7. **Keyword Manager – Calibrate – Load:** Extracts the camera length from the selected XDS.INP files and displays the unique values below the segment sign. The displayed values are influenced by the distance factor.
8. **Keyword Manager – Calibrate – Clean:** Resets the Calibration tab to its initial state.
9. **Keyword Manager – Save:**
 - Editing Mode: Summarises and validates the edited keyword values.
 - Calibration Mode: Summarises and validates the camera length calibration.

After confirmation in the message box, the changes are applied to the selected XDS.INP files.
10. **Keyword Manager – Cancel:** Discards any changes and returns to AutoLEI.
11. **Show Results:** Displays the summarised results of the batch processing below the button.
12. **Update Results File:** Summarises successfully reduced results into xdsrunner2.xlsx by extracting data from STATISTICS.json. If the file is not found, the function will analyse the results, which may take a few seconds. The results file is automatically updated after running XDS batch processing.
13. **Open Results File:** Opens xdsrunner2.xlsx using Excel (on WSL) or LibreOffice (on Linux). This Excel file contains additional detailed information and will disappear in Single and Failed modes.
14. **View Reciprocal Space:** Opens a new window called **Lattice Visualization**. In this viewer:
 - Failed Indexation: All points are shown in light blue.
 - Successful Indexation: Reflection points from the most populated indexed result are shown in black, while remaining unindexed points are shown in red.

The centre displays basis sets from the origin points, with vector lengths doubled relative to their corresponding vectors in reciprocal space. The 3D image can be freely rotated and scaled. Below the image, multiple buttons are available. Due to scaling issues, some buttons may be hidden; resizing the window will reveal hidden buttons.
15. **View Reciprocal Space – All / Indexed / Unindexed:** Filters the points displayed to show all points, only indexed points, or only unindexed points.
16. **View Reciprocal Space – Show / Hide Intensity:** Displays points with sizes proportional to the square root of their intensity.
17. **View Reciprocal Space – HK0, H0L, 0KL:** Shows slicing of the reciprocal space along the specified planes (HK0, H0L, 0KL).
18. **View Reciprocal Space – View [001] / [010] / [100]:** Changes the view direction to c^* , b^* , and a^* , respectively.
19. **Bus to SHELX:** For a selected single dataset, generates the following files in the same folder as XDS.INP:
 - SHELX hkl
 - XPREP p4p
 - cif_od: Contains basic information from data reduction, which can be read by OLEX2 to merge into the final CIF.

20. **Web Report:** Creates and opens an HTML report using SU as the analysis engine. The report includes data collection and reduction information presented in both textual and graphical formats.
21. **Open Folder:** Opens the folder containing the currently selected dataset.

3.5 DataMerge Tab

When the completeness is not high enough for single datasets, it is still possible to solve the structure by merging different datasets together. This is especially important for beam-sensitive materials, such as proteins. Merging also improves statistics and reduces the influence of the dynamical effect. The DataMerge Tab provides functionality to:

- Filter data by statistics
- Merge data
- View statistics on merged data
- Generate SHELX hkl and p4p files.

Features and Functionality

Hover over any input field, drop-down menu, or button to view tooltips explaining their specific functions. Below are details on the features available in this tab:

Drop-down menu:

The data quality filter can be selected between ISa, $CC_{1/2}$, R_{meas} , and Resolution. The ISa here refers to the model ISa from XDS and not the actual I/Sigma of the datasets.

Input Fields:

The input field is used to enter the threshold value for filtering the data. The default values for each filter are ISa (5), $CC_{1/2}$ (95), R_{meas} (50), and Resolution (1.0), which are reasonable for small molecules. For macromolecules and beam-sensitive materials, the threshold can be decreased.

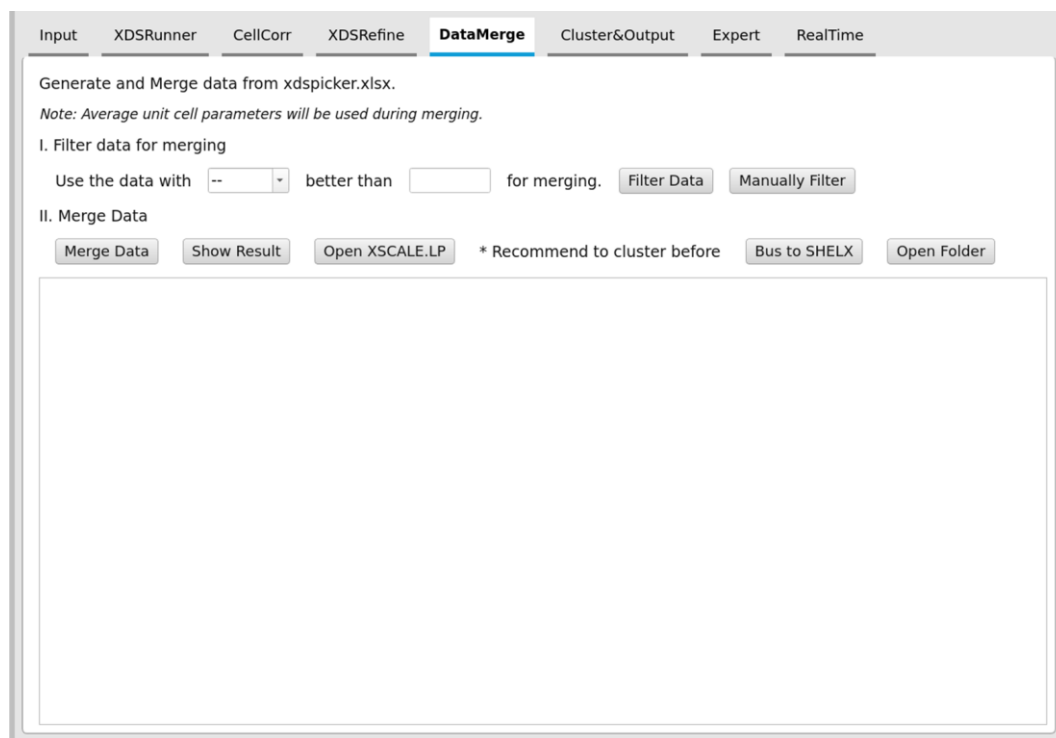


Figure 3.5 DataMerge Tab

Buttons:

1. **Filter Data:** Applies the selected filter to the data. The filtered data will be saved into “xdspicker.xlsx”.
2. **Manually Filter:** Opens “xdspicker.xlsx” in Excel (if WSL) or LibreOffice (if Linux). Entries can be manually removed from the form as needed. If “xdspicker.xlsx” does not exist, it will be generated with the same content as “xdsrunner2.xlsx.”
3. **Merge Data:** Merges the datasets in “xdspicker.xlsx” using XSCALE. For single datasets, the highest resolution is limited to the estimated resolution in the form. The merged data is saved in the “merge” subfolder under the working directory. AutoLEI will also generate “CLUSTER.JSON” in the merged folder.
4. **Show Result:** Displays a statistical table of resolution shells below the button.
5. **Open XSCALE.LP:** Opens “XSCALE.LP” in the internal text editor.
6. **Bus to SHELX:** For the merged datasets, generates the following files in the same folder as XSCALE.LP:
 - SHELX hkl
 - XPREP p4p
7. **Open Folder:**

3.6 Cluster_Output Tab

Merged data theoretically have better statistics than single datasets. However, due to dynamical effects, subtle structural changes, or suboptimal data reduction, intensity clustering of correlated data can be beneficial. The Cluster_Output Tab provides functionality to:

- Intensity-cluster datasets with a given distance

- Compare clusters with different distance cut-off
- Run XPREP on a selected cluster
- Supply necessary information on data collection and generated a cif_od file.

Figure 3.6 Cluster_Output Tab

Features and Functionality

Hover over any input field, checkbox, drop-down menu, or button to view tooltips explaining their specific functions. Below are details on the features available in this tab:

Input fields:

1. **Distance:** Specifies the cut-off distance for clustering. The distance between datasets is calculated from the correlation coefficient, following the equation: $d = \sqrt{1 - CC^2}$. CC of 0.90 equals to distance of 0.435.
2. **TEM Instrument Name:** Specifies the name of the electron microscope.
3. **Detector Name:** Specifies the name of the detector.
4. **Temperature:** Indicates the measurement temperature. Reference values are:
 - Cryo-column: 80 K
 - Cryo-holder: 100 K
 - Room temperature: 298 K
5. **Short Name:** A concise name for the sample. It should be a single word and free of illegal characters for both Windows and Linux. The short name appears as the file and data name in the cif file.

6. **Long Name:** A descriptive name for the sample.

Drop-down menu:

1. **Select Cluster (I):** Allows selection of clusters for processing.
2. **Select Instrument Profile (II):** Autofill instrument and detector details based on the selected profile.

Checkboxes:

1. **Overwrite previous result:** If checked, newly generated clusters will overwrite previous results. Otherwise, the new cluster will be saved with a new identifier.
2. **Cryoholder:** If checked, the temperature is set to 100 K. If unchecked, the temperature defaults to 298 K.

Buttons:

1. **Set Distance from Dendrogram:** Opens a window displaying the dendrogram calculated from “XSCALE.LP” in the merge subfolder. Clicking on the dendrogram changes the distance, and the clustering result is updated with assorted colours. Close the window to apply the updated distance.
2. **Make Cluster based on Distance:** Generates clusters based on the specified distance. If the overwrite checkbox is unchecked, new clusters are assigned a unique identifier and saved in a subfolder named “clusterX” in the merge folder. Otherwise, previous clusters are overwritten.
3. **Refresh and Show Summary:** Refreshes the cluster drop-down menu and displays cluster statistics in the command window.
4. **Open Dendrogram:** Open the dendrogram of the current selected cluster.
5. **Open XSCALE.LP:** Displays “XSCALE.LP” of the current selected cluster in the internal text editor.
6. **Run XPREP:** Launches XPREP (if available in Linux or WSL, or if its location is specified in the settings file) to determine symmetry and generate an .INS file for the current cluster.
7. **Open Report:** Creates and opens an HTML report for the current cluster using SU as the analysis engine. The report includes data collection and reduction information in both textual and graphical formats.
8. **Update INS and Metadata:** Updates the wavelength, title, and temperature in the INS file for the current cluster. Based on input and data reduction parameters, generates a cif_od file readable by Olex2.
9. **Open Folder:** Opens the folder for the current cluster. If no cluster is selected, the working directory is opened instead.

3.7 Expert Tab

The Expert Tab is an official extended feature of AutoLEI, providing several utility functions for advanced file management and processing. These functionalities include:

- Creating REDp files from FEI .mrc files.
- Rolling back XDS.INP files to a specific stage.
- Modifying the image path format in XDS.INP.

- Generating PETS input files and TIFF images from FEI .mrc files.

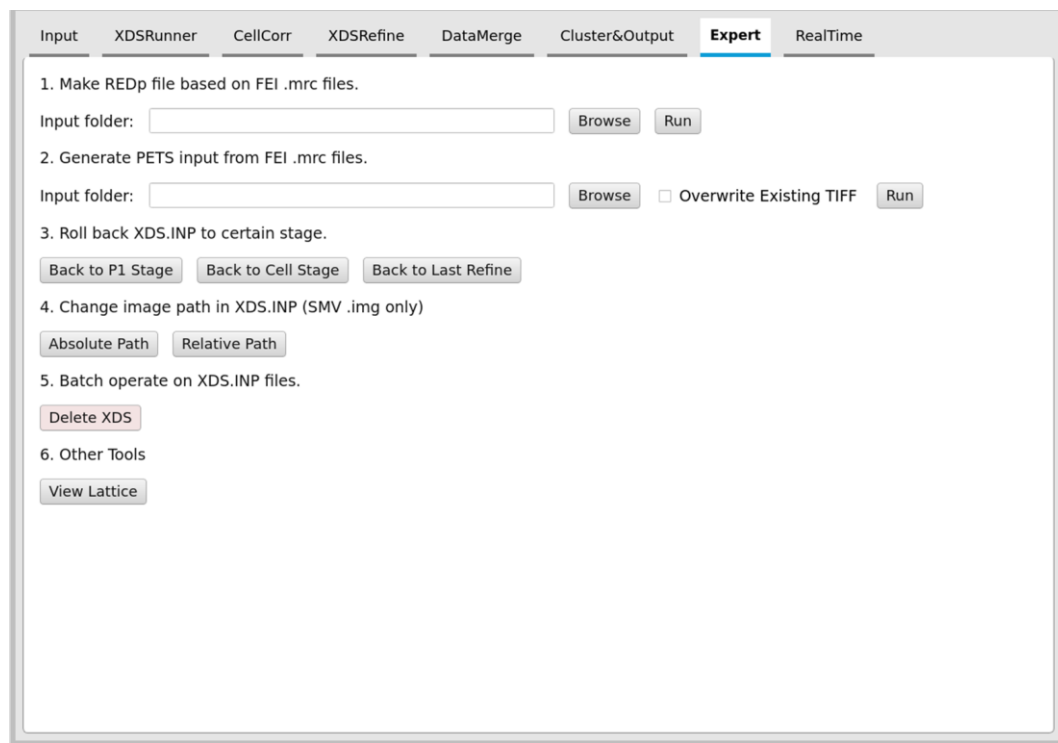


Figure 3.7 Expert Tab

Functionalities

1. Make REDp File on FEI .mrc Files

- Input the folder path in Linux format or use the Browse button to select the directory for processing.
- Click Run to generate REDp .ed3d input files in a subfolder named “redp” and convert FEI .mrc files to standard .mrc format.
- This process applies to all subfolders containing more than 10 FEI .mrc images, excluding folders generated by Instamatic.

2. Roll Back XDS.INP to a Certain Stage

- AutoLEI creates backup files when modifying XDS.INP. These backups are stored in the same folder as the original and named with the prefix “BACKUP-” (e.g., BACKUP-P1, BACKUP-CELL, BACKUP-REFINE).
- Use the corresponding button to roll back all XDS.INP files in the working directory to a selected backup stage.
- **Delete XDS:** Deletes XDS.INP and the created xds folder. Note: If the data folder has a different structure, some XDS-related files may not be deleted by this function.

3. Change the Image Path Format in XDS.INP

- By default, AutoLEI generates XDS.INP files with relative image paths. If absolute paths are required, the Absolute Path button can convert all XDS.INP files in the working directory to use absolute paths.
- This function also checks whether the specified images exist and attempts to locate missing images.

4. Generate PETS Input Files and TIFFs from FEI .mrc Files

- Input the folder path in Linux format or use the Browse button to select the directory for processing.
- Click Run to generate PETS .pts input files with autorun functionality in the same folder as the FEI .mrc files.
- Additionally, FEI .mrc files are converted to TIFF format and saved in a subfolder named “tiff.”
- This process applies to all subfolders containing more than 10 FEI .mrc images, excluding folders generated by Instamatic.

3.8 RealTime Tab

The RealTime Tab is an extended feature of AutoLEI, enabling real-time processing during data collection. Currently, this feature supports datasets from **EPU-D** and **Instamatic**.

- EPU-D Data: For data in non-stacked FEI .mrc format, processing includes:
 - Image conversion
 - XDS.INP creation
 - Beam centre/stop finding
 - Input parameter correction (optional)
 - Cell parameter/space group updates (optional)
- Instamatic Data: For data with pre-existing XDS.INP files, processing includes:
 - Updating XDS.INP
 - Input parameter correction (optional)
 - Cell parameter/space group updates (optional)

Its main functionalities include:

- **Cell Mode:** If the unit cell, space group, and resolution limit are provided, RealTime processes data in this mode. All collected data are processed using the given cell parameters and space group, and new data will be merged with existing good data if they meet the criteria. Completeness and $CC_{1/2}$ of merged data are displayed as plots.
- **Screen Mode:** If the unit cell, space group, or resolution limit is missing, RealTime operates in P1 mode. Data are processed and reduced without further refinement, and resolution, $CC_{1/2}$, and I/Sigma plots are shown.

Features and Functionality

Hover over any input field, drop-down menu, checkbox, display area or button to view tooltips explaining their specific functions. Below is a detailed breakdown of the components in this tab.

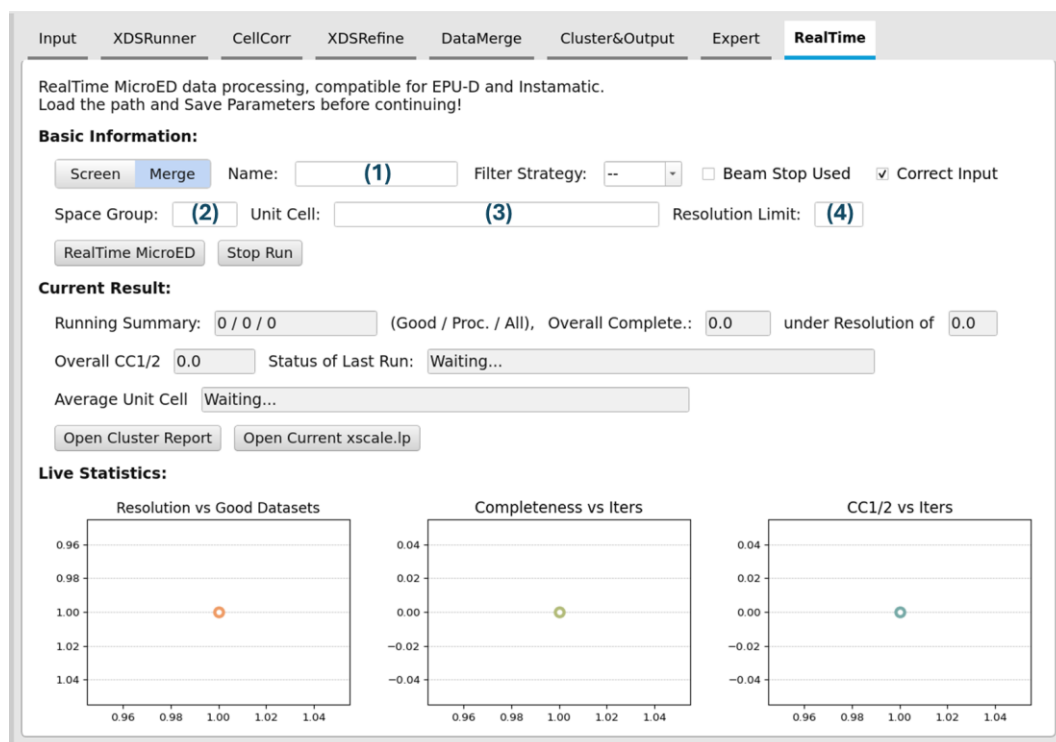


Figure 3.8 RealTime Tab

Input fields:

1. **Name:** A concise name for the sample. It should be a single word and free of illegal characters for both Windows and Linux. The short name appears as the folder name. If leave empty, “online” will be the default name.
2. **Space Group:** Enter the space group number (1 to 230) or any space group name (like C2/c, P212121, Fm-3m). XDS will automatically recognise different lattice settings for the same number and convert them to the standard setting.
3. **Unit Cell:** the cell parameters in order of a, b, c (in Å) and α , β , γ (in °). Valid input formats include:
 - 3.0, 4.0, 5.0, 90.0, 110, 90
 - 3.0 4.0 5.0 90.0 110 90
 - 3,0 4,0 5,0 90,0 110 90
4. **Resolution Limit:** Sets the resolution used for completeness and CC1/2 calculations in **Cell Mode**.

Drop-down menu:

Choose and configure the run and filter strategy for RealTime processing. Selecting the “Edit” option opens a new window where keywords are explained:

- **CYCLE_TIME:** Time interval (in seconds) for detecting changes in the folder. Default: 5 seconds.
- **WAITING_TIME:** Delay (in seconds) to ensure data collection completion. Default: 20 seconds.
- **RUN_FILTER:** Criteria for assessing data quality. Supported keywords include:
 - **CC1/2:** Default is $CC_{1/2} > 85$.

- **ISA:** Signal-to-noise ratio (default: $50 > \text{ISA} > 2$).
- **CC1/2, REXP, RMEAS, VOL_DEV, RESOLUTION:** units in %.
- **MERGE_FILTER:** Criteria for clustering (not implemented in version 1.0).

Checkboxes:

- **Beam Stop Used:** if checked, the beam centre finding will be replaced by the beam centre finding.
- **Correct Input:** if checked, the XDS.INP will be corrected by image header.

Display Area:

- **Running Summary:** Displays the status of datasets categorised as good, processable, or transferred. A valid dataset must contain an XDS.INP file.
- **Status of Last Run:** Indicates whether the last processed dataset was good or bad.
- **Overall Completeness (Cell Mode):** Shows the completeness of the latest merged cluster.
- **Resolution (Cell Mode):** Displays the resolution used for calculating completeness and $CC_{1/2}$.
- **Overall CC1/2 (Cell Mode):** Shows $CC_{1/2}$ of the latest merged cluster.
- **Average Unit Cell (Cell Mode):** Displays the unit cell parameters of the latest merged cluster.

Buttons:

1. **Realtime MicroED:** All data from AutoLEI Realtime processing will be stored in realtime.json under the working directory. AutoLEI can automatically read the previous realtime.json to continue the data processing.

In the initial stage, AutoLEI first attempts to analysis all already existing data in the direct subfolders and use the provided criteria to judge if it is good or not. If it is in the cell mode, AutoLEI will then go through the merged folder (with the same name as input), using the given resolution to calculate the completeness and the $CC_{1/2}$. After the initial stage, the display area will be updated. The setting-up process will take several minutes if lots of datasets (>40) are already present in the folder.

Then AutoLEI will continuously monitor the working directory every CYCLE_TIME and mark the subfolder with data collecting in. When the number of the image is equal to the Instamatic's log, or do not change after given WAITING_TIME, AutoLEI will go through processing functions. It then executes XDS batch processing using the successfully refined parameters. Upon successful processing, two additional files are generated:

- XDS.LP: Logs the processing details displayed on the screen.
- STATISTICS.json: Collects necessary data reduction information.

Additionally, the Auto-Re-run function (enabled/disabled in settings) automatically retries failed datasets using predefined strategies in AutoLEI.

2. **Stop Run:** Terminates the XDS runner after completing the current XDS file.

3. **Open Cluster Report:** Creates and opens an HTML report for the latest cluster using SU as the analysis engine. The report includes data collection and reduction information in both textual and graphical formats.
4. **Open Current XSCALE.LP:** Displays “XSCALE.LP” of the latest cluster in the internal text editor.

Buttons:

1. Mode Shift Button:

- Screen/Merge mode: By clicking corresponding button, AutoLEI will shift between **Screen** mode and **Merge** mode. In the Screen mode, only name is allowed to be input.

2. Realtime MicroED:

- Processes all data collected by AutoLEI RealTime, storing information in realtime.json within the working directory.
- AutoLEI can read the existing realtime.json file to continue processing previously collected data.

Initial Stage:

- AutoLEI analyses all existing data in direct subfolders to evaluate quality based on the provided criteria.
- In **Cell Mode**, it processes the merged folder (named after the input) and calculates completeness and $CC_{1/2}$ using the specified resolution.
- The display area is updated after this stage, which may take several minutes if there are many datasets (>40).

Real-Time Monitoring:

- AutoLEI continuously monitors the working directory at intervals specified by CYCLE_TIME and tracks subfolders where data is being collected.
- When the number of images matches the Instamatic log or remains unchanged after WAITING_TIME, AutoLEI processes the data.
- XDS batch processing is executed using the refined parameters.

Outputs for single datasets:

- **XDS.LP:** Logs the processing details displayed on the screen.
- **STATISTICS.json:** Summarises data reduction results.

Auto-Re-run Function:

- If enabled in settings, AutoLEI automatically retries failed datasets using predefined strategies.

3. **Stop Run:** Halts the RealTime process after the current XDS file is completed.
4. **Open Cluster Report:**
 - Generates and opens an HTML report for the latest cluster using SU as the analysis engine.
 - The report provides data collection and reduction information, displayed in textual and graphical formats.
5. **Open Current XSCALE.LP:** Opens the “XSCALE.LP” file of the latest cluster in AutoLEI’s internal text editor.

4. Tutorial Example 1: L-Tyrosine

4.1 General information

- **Raw Data:** The L-Tyrosine datasets can be downloaded from <https://doi.org/10.5281/zenodo.13835773>
- **Data Collection:** 12 datasets were collected using an ASI TimePix hybrid detector installed on a JEOL JEM-2100 (200 kV) microscope equipped with a LaB₆ filament. Data was collected at cryo temperature with a Gatan 914 cryo-holder.
- **Collection Method:** Data collection utilised Instamatic, with every 20th frame defocused to track the crystal.
- **References:**
 - Instamatic details: [<https://zenodo.org/records/11091622>]
 - Sample and data collection: [AutoLEI paper]
 - Final structure deposited in CCDC: [AutoLEI CCDC]
- **Sample Details:** L-Tyrosine (C₉H₁₁NO₃), space group: *P*2₁2₁2₁, unit cell dimensions: 6.01 Å, 7.03 Å, 21.70 Å, 90°, 90°, 90°.
- **Software Versions:** Tutorial conducted using AutoLEI version 1.0 and XDS version 2025/04/30.

4.2 Goals

1. Understand the workflow of AutoLEI.
2. Perform batch processing of electron diffraction data using provided XDS.INPs.
3. Finalise data processing for a single dataset.

4.3 Load Working Directory

1. Start AutoLEI:
 - If using a pip-installed version, activate the Python environment in Linux and type **autolei**.
 - Alternatively, navigate to the folder containing **autolei.py** and run it.
 - After version and scaling information appear in the command window, the main AutoLEI window will open.
 - **Note:** If a message appears stating “XDS not existing! Do you wish to continue?”, ensure XDS is accessible to AutoLEI (e.g., added to the system path).
2. Select Input Path:
 - Click **Browse** to select the folder containing the 12 datasets.
 - Click **Load**. Once loaded, the title updates to display the selected path and the number of datasets. Input parameters will also reflect communal settings.
 - Proceed to the XDSRunner tab as XDS.INPs are pre-generated.

AutoLEI 1.0.2 - /mnt/c/AutoLEI_demo/Tyrosine - 12 datasets

Input XDSRunner CellCorr XDSRefine DataMerge Cluster&Output Expert RealTime

Browse and load the work path where the program will load the measurement settings.
For XDS input generation, supply basic parameters and click the 'Save Parameter' button.

Input path: /mnt/c/AutoLEI_demo/Tyrosine Browse Load Path Instrument: Custom Load

I. Instrument Parameters

1. Detector parameters: NX= 516 NY= 516 QX= 0.0550 QY= 0.0550

2. Overloading: OVERLOAD= 130000 3. Wavelength: WAVELENGTH= 0.0251 Å

4. Rotation axis: ROTATION_AXIS= -0.6204 0.7843 0.0000 or Angle in Degree= 128.35

5. Additional information
(Please copy from XDS):

```
UNTRUSTED_RECTANGLE= 255 262 0 517
UNTRUSTED_RECTANGLE= 0 517 255 262
```

II. Measurement Parameters

6. Direct beam position: ORGX= ORGY=

7. Resolution range: INCLUDE_RESOLUTION_RANGE= 20 0.8 >>> Use space to segment

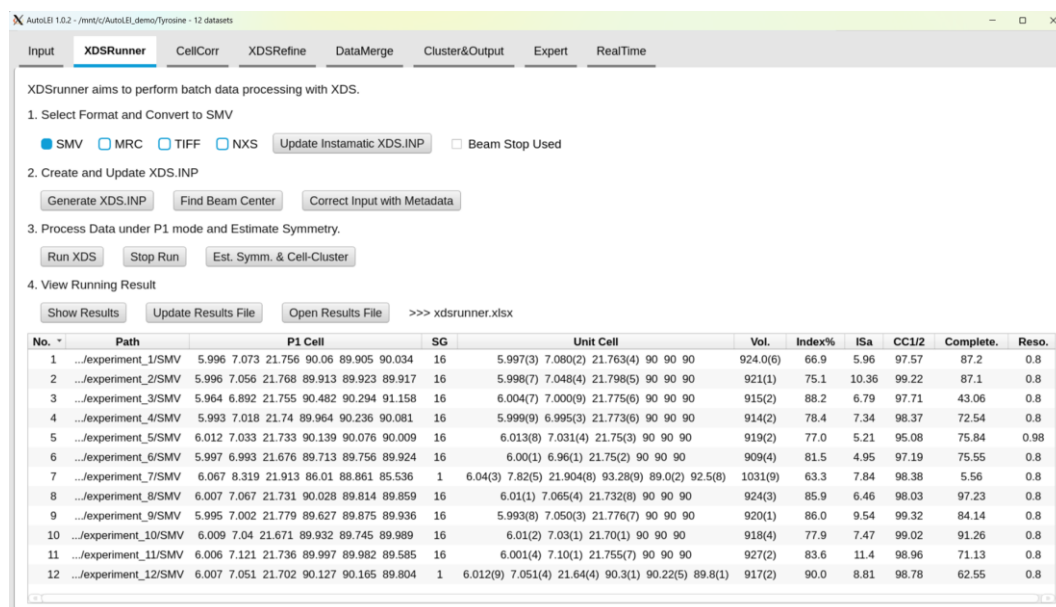
8. Camera Length: DETECTOR_DISTANCE= +439.48 9. Rotation step: OSCILLATION_RANGE=

Save Parameters

4.4 Batch Processing Datasets in P1 Mode

- Correct the XDS.INP:
 - Since the data is in SMV .img format, no conversion is necessary.
 - Select SMV format and click **Update Instamatic XDS.INP** to batch edit all XDS.INP files, removing illegal characters and obsolete keywords.
- Run XDS and View the result
 - Click **Run XDS** to perform batch processing electron diffraction data. The processing will be finished with several minutes. Processing time varies with the operating system and hardware performance (e.g., WSL1 is faster than WSL2).
 - The process completes when the command window shows “All information extracted. Note that resolution is ideal,” or the animation in AutoLEI stops. Monitor progress in the command window’s bottom line.
 - After batch processing, AutoLEI will analysis the processing results automatically and generate xdsrunner.xlsx with all result. Click **Show Results** to show the statistics below the button.
 - The table will contain information for all datasets which successful processed:
 - [A] Serial Number
 - [B] Data Path
 - [C] P1 Unit Cell
 - [D] Space Group
 - [E] Symmetry-Adopted Unit Cell

- [F] Index% (from SPOT.XDS)
 [G] Model ISa (from CORRECT.LP)
 [H] CC1/2
 [I] Completeness
 [J] Resolution



3. Estimate Symmetry:

- Click **Estimate Symmetry & Cell-Cluster**. Results are displayed in the command window and saved as lattice_cluster.txt in the working directory.
- The cell cluster report provides detailed analysis on:
 - Potential Symmetry: Indicates to which symmetry the datasets align.
 - Cluster Analysis: Reveals that 11 of 12 datasets form Cluster 1 (similar unit cell and high symmetry), while the 7th dataset is classified as Cluster 2 due to slightly differing unit cell parameters and low symmetry.
- Based on Cluster 1, the primary unit cell is identified as having an orthorhombic lattice and Laue group *mmm* with:
 - R_{meas} ratio increased slightly: 1.175
 - CC ratio enhancement: 0.918
- For further processing, use the refined unit cell parameters: 6.0, 7.027, 21.718 (in Å) and angles 90°, 90°, 90°. Adopt space group 16 (the base group for *mmm*).

```

***** Cluster-1: *****
** Data Path: 11 datasets, (91.7%)
/mnt/c/AutoLEI_demo/Tyrosine/experiment_1/SMV
/mnt/c/AutoLEI_demo/Tyrosine/experiment_2/SMV
/mnt/c/AutoLEI_demo/Tyrosine/experiment_3/SMV
/mnt/c/AutoLEI_demo/Tyrosine/experiment_4/SMV
/mnt/c/AutoLEI_demo/Tyrosine/experiment_5/SMV
/mnt/c/AutoLEI_demo/Tyrosine/experiment_6/SMV
/mnt/c/AutoLEI_demo/Tyrosine/experiment_8/SMV
/mnt/c/AutoLEI_demo/Tyrosine/experiment_9/SMV
/mnt/c/AutoLEI_demo/Tyrosine/experiment_10/SMV
/mnt/c/AutoLEI_demo/Tyrosine/experiment_11/SMV
/mnt/c/AutoLEI_demo/Tyrosine/experiment_12/SMV

** Lattice Choice:
-- Bravais Lattice: oP
Averaged Cell: (6.0, 7.036, 21.736, 90.0, 90.0, 90.0), Lattice Difference: 0.0364, FOM: 13.4
SEM of Cell: (0.0, 0.019, 0.015, 0.0, 0.0, 0.0)
Suggested Space Group: P222 (No. 16) : 1.431(R), 1.026(C)

-- Bravais Lattice: mP (setting 1)
Averaged Cell: (6.0, 7.036, 21.736, 90.0, 90.164, 90.0), Lattice Difference: 0.0345, FOM: 8.3
SEM of Cell: (0.0, 0.019, 0.015, 0.0, 0.027, 0.0)
Suggested Space Group: P121 (No. 3) : 1.204(R), 0.987(C)

-- Bravais Lattice: mP (setting 2)
Averaged Cell: (7.036, 6.0, 21.736, 90.0, 90.155, 90.0), Lattice Difference: 0.0364, FOM: 7.7
SEM of Cell: (0.019, 0.0, 0.015, 0.0, 0.049, 0.0)
Suggested Space Group: P121 (No. 3) : 1.211(R), 0.998(C)

-- Bravais Lattice: mP (setting 3)
Averaged Cell: (6.0, 21.736, 7.036, 90.0, 90.1, 90.0), Lattice Difference: 0.0145, FOM: 11.3
SEM of Cell: (0.0, 0.015, 0.019, 0.0, 0.114, 0.0)
Suggested Space Group: P121 (No. 3) : 1.32(R), 1.062(C)

-- Bravais Lattice: aP
Averaged Cell: (6.0, 7.036, 21.736, 89.845, 89.891, 89.791), Lattice Difference: 0.0, FOM: 1.0
SEM of Cell: (0.0, 0.019, 0.015, 0.049, 0.045, 0.1)
Suggested Space Group: P1 (No. 1) : 1.0(R), 1.0(C)

***** Cluster-2: *****
** Data Path: 1 datasets, (8.3%)
/mnt/c/AutoLEI_demo/Tyrosine/experiment_7/SMV

** Lattice Choice:
-- Bravais Lattice: aP
Averaged Cell: (6.1, 8.3, 21.9, 86.0, 88.9, 85.5), Lattice Difference: 0.0, FOM: 0.0
SEM of Cell: (0.0, 0.0, 0.0, 0.0, 0.0, 0.0)
Suggested Space Group: P1 (No. 1) : 1.0(R), 1.0(C)

```

4.5 Batch Processing Datasets in Cell Mode

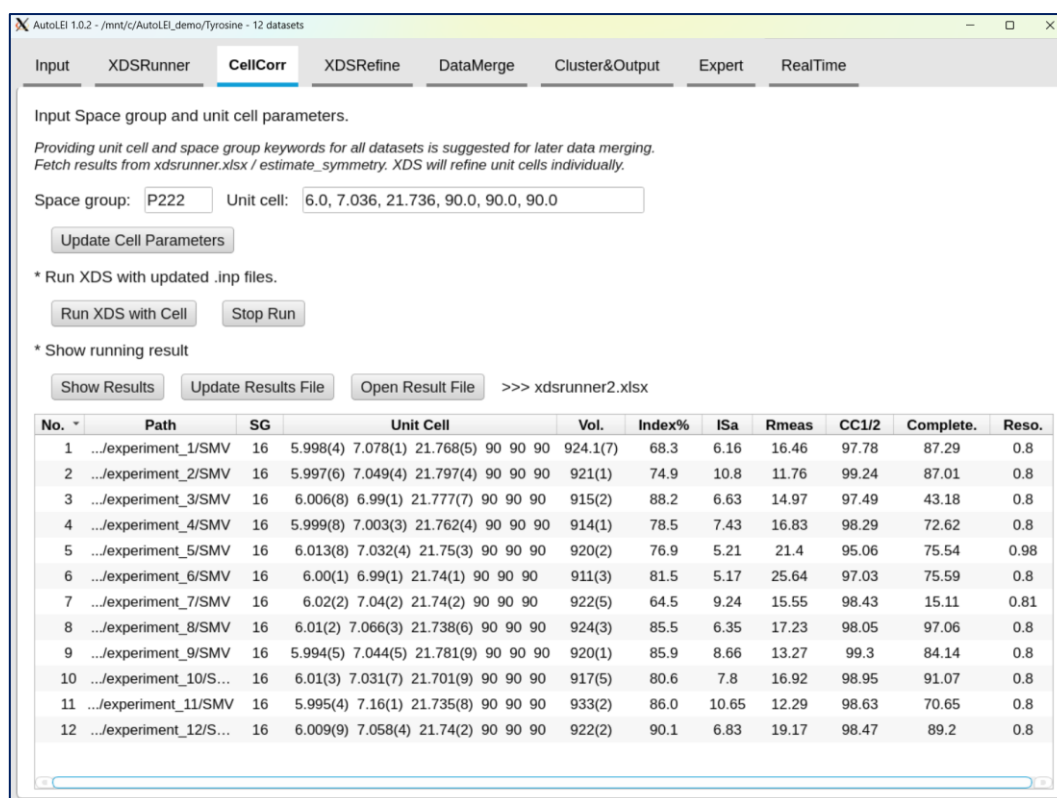
Shift to CellCorr Tab to continue.

1. Supply Unit Cell and Space Group information:

- Enter the space group (16) and unit cell dimensions (6.0, 7.027, 21.718, 90, 90, 90) in the respective fields.
- Click **Update Cell Parameters**, a confirmation message will appear.

2. Run XDS and View the result:

- Click **Run XDS with Cell** to perform batch processing electron diffraction data. The processing will be finished with several minutes.
- After batch processing, AutoLEI will analysis the processing results automatically and generate xdsrunner2.xlsx with all result. Click **Show Results** to show the statistics below the button.
- The table will contain information for all datasets which successful processed:
[A] Serial number, [B] Path of the data, [C] Space group, [D] Symmetry-adopted Unit cell, [E] Index% (calculated from SPOT.XDS), [F] Model ISa from XDS (extracted from CORRECT.LP), [G] R_{meas} , [H] $CC_{1/2}$, [I] Completeness, [J] Resolution.

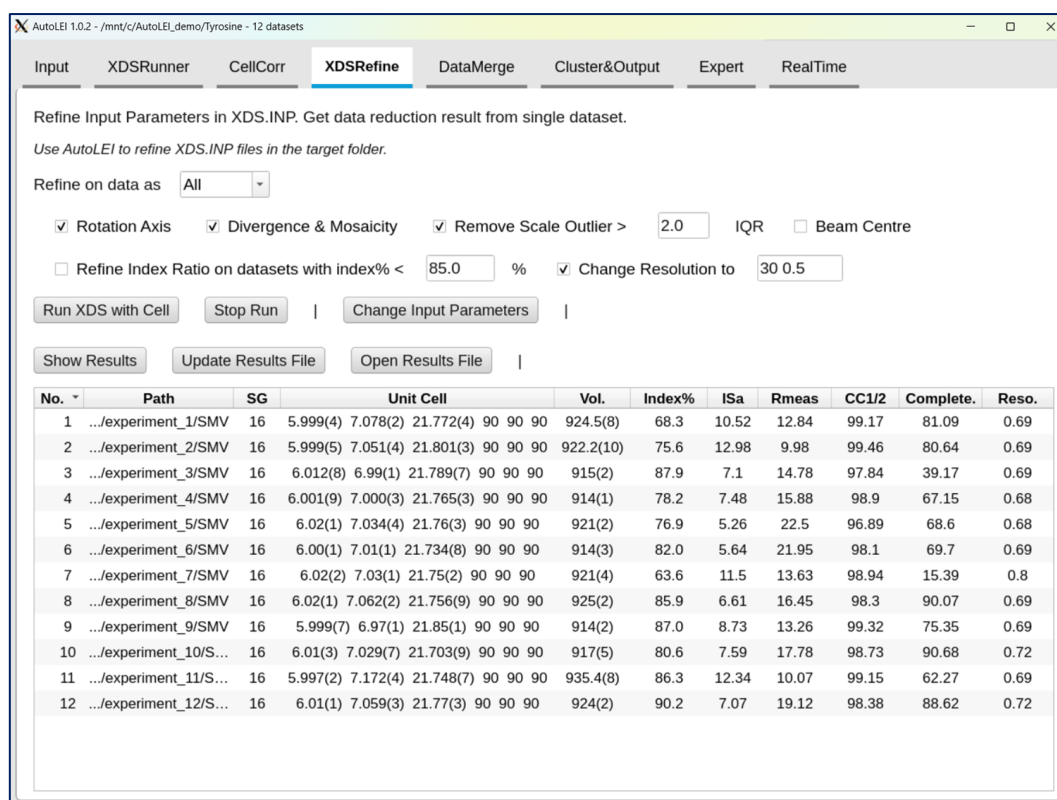


4.6 Refine Input Parameter

The summary table reveals two noteworthy observations. First, nearly all datasets exhibit high completeness, enabling structure determination using a single dataset. Second, the reported resolution values are commonly at 0.8, which is the upper limit of the resolution range. Before proceeding to structure solution, it is essential to refine the input parameters to achieve better results.

- Refine input parameters:
 - Check **Rotation Axis, Divergence & Mosaicity, Remove Scale Outlier** and **Change Resolution**
 - Change the value of resolution range to “**30 0.5**”
- Run XDS and View the result:
 - Click **Run XDS with Cell** to perform batch processing electron diffraction data. In the command window, the refinement process will be shown. After input parameter refinement, data reduction process will be started. The processing will be finished with several minutes.
 - After batch processing, AutoLEI will analysis the processing results automatically and update xdsrunner2.xlsx with all result. Click **Show Results** to show the statistics below the button.
 - The table will contain information for all datasets which successful processed:

[A] Serial number, [B] Path of the data, [C] Space group, [D] Symmetry-adopted Unit cell, [E] Index% (calculated from SPOT.XDS), [F] Model ISa from XDS (extracted from CORRECT.LP), [G] R_{meas} , [H] $CC_{1/2}$, [I] Completeness, [J] Resolution.
 - All datasets, excluding experiment 7, will now have improved results. Then, Proceed to generate the SHELX HKL files.



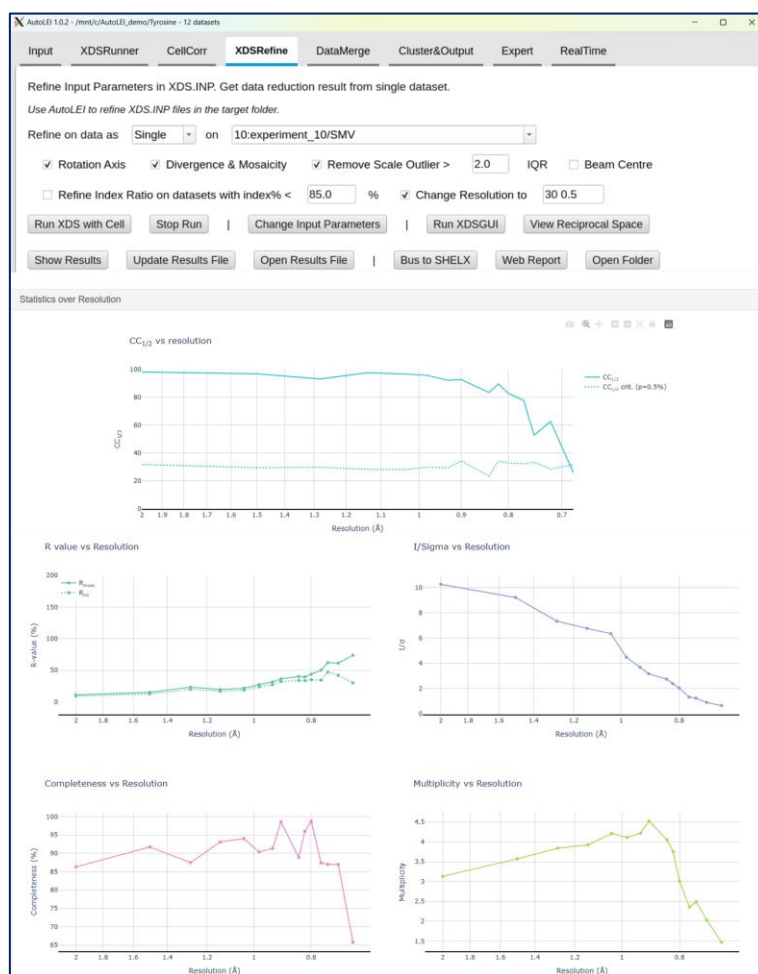
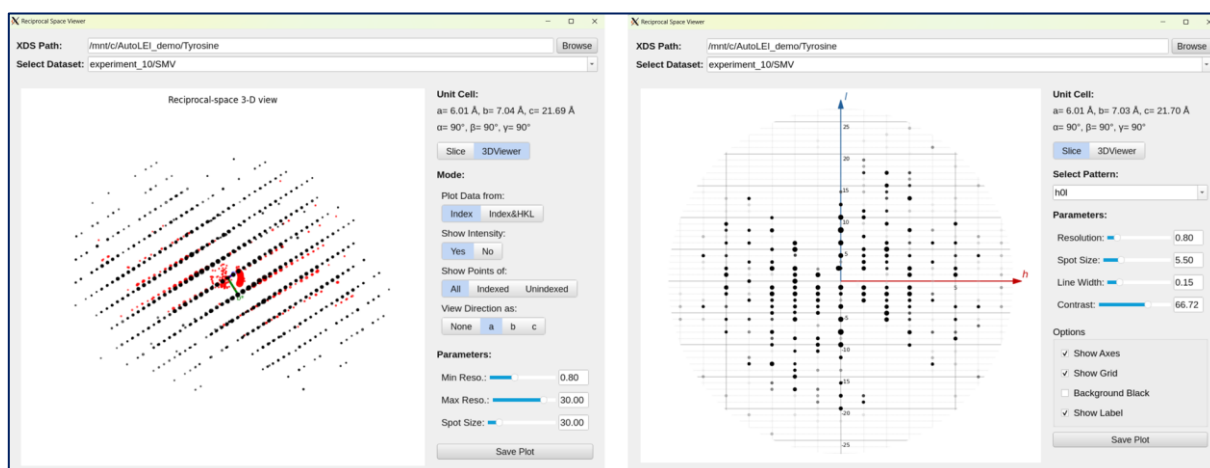
4.7 Finalise Data Reduction

1. Generate File for Structure Solvation:

- Change Data range drop-down menu from **All** to **Single**, additional buttons will appear.
- You can also right-click the dataset to do below options.
- Select **10:experiment_10/SMV** as example.
- View Reciprocal Space:
 - Click **View Reciprocal Space** to open the lattice visualization.
 - You can switch between 3D viewer and Slice viewer, where the former one is the default option when opening the app.
 - In the slice viewer, choose **h0l** to generate a slice of **h0l**. The slice clearly shows that the **00l** reflections follow the rule **00l: l = 2n**. On other axes, the lack of reflections makes it difficult to determine symmetry absence on other planes.
 - Close the window after view.
- Generate Web Report:
 - Click **Web Report** to generate and open the web report for the data reduction result of experiment 10.
 - The web report includes three tables:
 1. Collection Information
 2. Overall Statistics on Data Reduction Results

3. Sub-Resolution-Shell Statistics

- Below the tables, there are two parts of plots, Statistics over Resolution and Statistics over Frames:
 - CC1/2 vs. Resolution: CC1/2 drops dramatically from 0.7.
 - Scale vs. Frames: The scale appears healthy. These indicators confirm the reliability of the data reduction results.
- Click **Bus to SHELX** to generate SHELX hkl, XPREP p4p and metadata cif_od file.
- Click **Open Folder** to open the **experiment_10/SMV** folder, where the data reduction is finished.



Contents of 1.cif_od:

```
data_1
_cell_measurement_reflns_used      2307
_cell_measurement_theta_min        0.0331
_cell_measurement_theta_max        1.0529
_exptl_crystal_description         nano-crystal
_diffraction_ambient_environment    vacuum
_diffraction_measurement_device_type TEM
_diffraction_radiation_type         electron
_diffraction_radiation_source       electron
_diffraction_radiation_probe        electron
_diffraction_radiation_wavelength   0.0251
_diffraction_source                 'electron microscope'
_diffraction_source_voltage          200
_diffraction_measurement_device      TEM
_diffraction_measurement_method      'Continuous Rotation Electron Diffraction'
_diffraction_detector_area_resol_mean 18.1818
_diffraction_measurement_details
;
List of Samples for Merging (Abbr. = Camera Length, Resolution):
#   Type   Start      End        Step    t~exp~   #Frames   CL        Reso.   ISa     CC1/2
-----
1   /a     -56.38    61.9072    0.2328   0.5      509       439.48   0.72    4.31    98.73
;
_computing_cell_refinement          'XDS (Kabsch et al., 2010)'
_computing_data_reduction            'XDS (Kabsch et al., 2010)'
_computing_data_collection           'Instamatic (Wang et al., 2019)'
```

5. Tutorial Example 2: CAU-36

5.1 General information

- **Raw Data:** The CAU-36 datasets can be downloaded from <https://doi.org/10.5281/zenodo.10891006>, of CAU-36only.zip. In this tutorial, part 4 will be used.
- **Data Collection:** 8 datasets were collected using an ASI TimePix hybrid detector installed on a JEOL JEM-2100 (200 kV) microscope equipped with a LaB₆ filament. Data was collected at cryo temperature with a Gatan 914 cryo-holder.
- **Collection Method:** Data collection utilised Instamatic. And generated XDS.INPs are purposely deleted.
- **References:**
 - Instamatic details: [<https://zenodo.org/records/11091622>]
 - Sample and data collection: [<https://doi.org/10.1002/chem.201804133>]
 - Final structure deposited in CCDC:
[<https://www.ccdc.cam.ac.uk/structures/Search?refcode=HIQDUJ>]
- **Sample Details:** CAU-36 [Co₂(NiC₄₄H₂₄N₄P₄O₁₂)·(C₆H₁₂N₂)_n], space group: $P\bar{4}c2$, unit cell dimensions: 21.980(5)Å, 21.980(5)Å, 8.9600(17)Å, 90°, 90°, 90°
- **Software Versions:** Tutorial conducted using AutoLEI version 1.0 and XDS version 2025/04/30.

5.2 Goals

1. Understand the workflow of AutoLEI.
2. Generate and Correct XDS.INPs.
3. Perform batch processing of electron diffraction data using generated XDS.INPs.
4. Finalise data processing by merging dataset.

5.3 Load Working Directory

1. Start AutoLEI:
 - If using a pip-installed version, activate the Python environment in Linux and type **autolei**.
 - Alternatively, navigate to the folder containing **autolei.py** and run it.
 - After version and scaling information appear in the command window, the main AutoLEI window will open.
 - **Note:** If a message appears stating “XDS not existing! Do you wish to continue?”, ensure XDS is accessible to AutoLEI (e.g., added to the system path).
2. Select Input Path:
 - Click **Browse** to select the folder containing the 8 datasets.
 - Click **Load**. Once loaded, the title updates to display the selected path and the number of datasets. Input parameters will also reflect communal settings.

3. Select Instrument Profile:

- Use the drop-down menu on the top-right corner to select **TimePix-512-2100_SU** as the instrument profile.
- Click **Load** to load the instrument profile into the current window.
- Most input fields are fulfilled. In the additional information part, the untrusted area excludes the gap between the chips on the TimePix Detector.

4. Input Parameters:

- For the eight datasets, the oscillation angle is around 0.23 and the camera length is 531.85 mm, except datasets 8 with 439.13 mm. Thus, to generate XDS.INPs with correct parameters, correct with image header data is necessary, which will be performed then. In this step, we just input some arbitrary number.
- Enter 1 in DETECTOR_DISTANCE box.
- Enter 0.2 in OSCILLATION_RANGE box.

5. Save the parameters through clicking **Save Parameters** button. A window will appear to confirm.

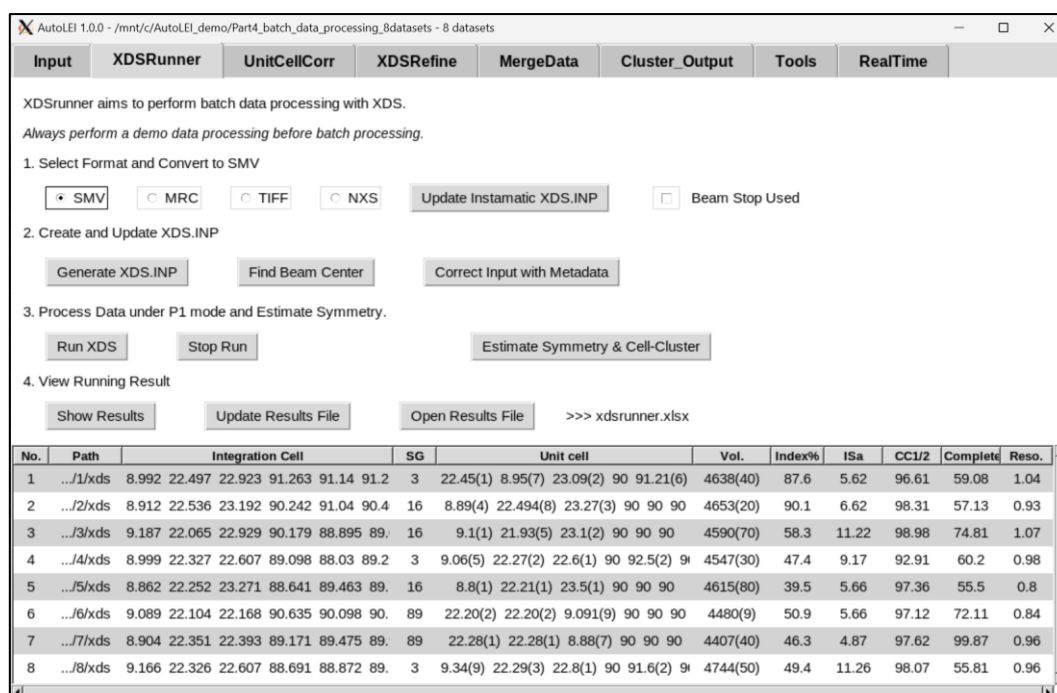
The screenshot shows the 'Input' tab of the AutoLEI 1.0.0 software. The window title is 'AutoLEI 1.0.0 - /mnt/c/AutoLEI_demo/Part4_batch_data_processing_8datasets - 8 datasets'. The top menu bar includes 'Input', 'XDSRunner', 'UnitCellCorr', 'XDSRefine', 'MergeData', 'Cluster_Output', 'Tools', and 'RealTime'. Below the menu, there is a text box with instructions: 'Browse and load the work path where the program will load the measurement settings. For XDS input generation, supply basic parameters and click the 'Save Parameter' button.' Below this, there is an 'Input path:' field with the value '/mnt/c/AutoLEI_demo/Part4_batch_d' and a 'Browse' button. To the right, there is an 'Instrument File:' dropdown menu with 'TimePix-512_2100_SU' selected and a 'Load' button. The main area is divided into two sections: 'I. Instrument Parameters' and 'II. Measurement Parameters'. Section I includes fields for '1. Detector parameters' (NX=512, NY=512, QX=0.055, QY=0.055), '2. Overloading' (OVERLOAD=65000), '3. Wavelength' (WAVELENGTH=0.0251 Å), '4. Rotation axis' (ROTATION_AXIS=-0.7550 0.6557 0) with a note 'Use space to segment OR Angle in Degree = -139.025', and '5. Additional information (Please copy from XDS):' with a text box containing 'UNTRUSTED_QUADRILATERAL= 255 1 258 1 258 512 255 512' and 'UNTRUSTED_QUADRILATERAL= 512 255 512 258 1 258 1 255'. Section II includes fields for '6. Direct beam position' (ORGX=256, ORGY=256), '7. Resolution range' (INCLUDE_RESOLUTION_RANGE=30 0.8) with a note 'Use space to segment', '8. Camera Length' (DETECTOR_DISTANCE=600), and '9. Rotation step' (OSCILLATION_RANGE=0.2). At the bottom left, there is a 'Save Parameters' button.

5.4 Batch Processing Datasets in P1 Mode

1. Generate XDS.INP:

- Since the image is already SMV .img, there will be no need to convert the format.
- Click **Generate XDS.INP** to generate the input file for XDS. XDS.INP will be generated in all 8 data folders in the subfolder "xds". (You can monitor the updates in the bash window.)

2. Beam Centre Finding:
 - As an arbitrary beam centre position is provided in the input section, a more precise estimation of the beam centre location for each dataset is needed.
 - Click **Find Beam Centre**, the beam centre will be written to the XDS.INP and direct_beam_centers.txt under the image folder.
3. Correct Input with Metadata:
 - Because of approximate input parameters given in the Input and the possibility of parameters changed during the whole experiment, the XDS.INP should be corrected to adapt every single experiment. The header information read from the image can be used to correct parameters on image size, pixel size, detector distance, and oscillation range if not properly input.
 - Click **Correct Input with Metadata** to make these corrections.
4. Run XDS and View the result
 - Click **Run XDS** to perform batch processing electron diffraction data. The processing will be finished with several minutes. Processing time varies with the operating system and hardware performance (e.g., WSL1 is faster than WSL2).
 - The process completes when the command window shows “All information extracted. Note that resolution is ideal,” or the animation in AutoLEI stops. Monitor progress in the command window’s bottom line.
 - After batch processing, AutoLEI will analysis the processing results automatically and generate xdsrunner.xlsx with all result. Click **Show Results** to show the statistics below the button.
 - The table will contain information for all datasets which successful processed:
 - [A] Serial Number
 - [B] Data Path
 - [C] P1 Unit Cell
 - [D] Space Group
 - [E] Symmetry-Adopted Unit Cell
 - [F] Index% (from SPOT.XDS)
 - [G] Model ISa (from CORRECT.LP)
 - [H] CC1/2
 - [I] Completeness
 - [J] Resolution



5. Estimate Symmetry:

- Click **Estimate Symmetry & Cell-Cluster**. Results are displayed in the command window and saved as lattice_cluster.txt in the working directory.
- The cell cluster report provides detailed analysis on:
 - Potential Symmetry: Indicates to which symmetry the datasets align.
 - Cluster Analysis: All datasets are within same cluster.
- The primary unit cell is identified as having an *tP* lattice (though the difference is a bit large), and Laue group *4/mmm* has:
 - R_{meas} ratio increased slightly: 1.471
 - CC ratio increased: 1.412
- It is trustable that the Laue Group is *4/mmm*, so the lattice symmetry can be decided as *tP*. For further processing, use the refined unit cell parameters: 22.537, 22.537, 9.025 (in Å) and angles 90.0, 90.0, 90.0. Adopt space group 89 (the base group for *4/mmm*).


```

***** Cluster-1: *****
** Data Path: 8 datasets, (100.0%)
/mnt/c/AutoLEI_demo/Part4_batch_data_processing_8datasets/1/xds
/mnt/c/AutoLEI_demo/Part4_batch_data_processing_8datasets/2/xds
/mnt/c/AutoLEI_demo/Part4_batch_data_processing_8datasets/3/xds
/mnt/c/AutoLEI_demo/Part4_batch_data_processing_8datasets/4/xds
/mnt/c/AutoLEI_demo/Part4_batch_data_processing_8datasets/5/xds
/mnt/c/AutoLEI_demo/Part4_batch_data_processing_8datasets/6/xds
/mnt/c/AutoLEI_demo/Part4_batch_data_processing_8datasets/7/xds
/mnt/c/AutoLEI_demo/Part4_batch_data_processing_8datasets/8/xds

** Lattice Choice:
-- Bravais Lattice: tP
Averaged Cell: (22.537, 22.537, 9.025, 90.0, 90.0, 90.0), Lattice Difference: 1.21, FOM: 209.9
SEM of Cell: (0.054, 0.054, 0.03, 0.0, 0.0, 0.0)
Suggested Space Group: P422 (No. 89) : 1.311(R), 0.963(C) / P4 (No. 75) : 1.252(R), 1.095(C)

-- Bravais Lattice: oC
Averaged Cell: (31.583, 32.067, 9.017, 90.0, 90.0, 90.0), Lattice Difference: 0.1367, FOM: 120.1
SEM of Cell: (0.155, 0.135, 0.044, 0.0, 0.0, 0.0)
Suggested Space Group: C222 (No. 21) : 1.224(R), 0.983(C)

-- Bravais Lattice: oP
Averaged Cell: (9.025, 22.312, 22.762, 90.0, 90.0, 90.0), Lattice Difference: 0.2875, FOM: 83.0
SEM of Cell: (0.042, 0.051, 0.126, 0.0, 0.0, 0.0)
Suggested Space Group: P222 (No. 16) : 1.172(R), 0.906(C)

-- Bravais Lattice: mC (setting 1)
Averaged Cell: (31.88, 31.91, 9.03, 90.0, 90.91, 90.0), Lattice Difference: 0.138, FOM: 154.0
SEM of Cell: (0.127, 0.14, 0.04, 0.0, 0.195, 0.0)
Suggested Space Group: C121 (No. 5) : 1.14(R), 1.134(C)

-- Bravais Lattice: mP (setting 1)
Averaged Cell: (22.312, 9.025, 22.762, 90.0, 90.837, 90.0), Lattice Difference: 0.1175, FOM: 36.9
SEM of Cell: (0.051, 0.042, 0.126, 0.0, 0.159, 0.0)
Suggested Space Group: P121 (No. 3) : 1.095(R), 1.026(C)

-- Bravais Lattice: mP (setting 2)
Averaged Cell: (9.018, 22.582, 22.591, 90.0, 90.791, 90.0), Lattice Difference: 0.12, FOM: 60.0
SEM of Cell: (0.038, 0.12, 0.115, 0.0, 0.176, 0.0)
Suggested Space Group: P121 (No. 3) : 1.11(R), 0.962(C)

-- Bravais Lattice: aP
Averaged Cell: (9.025, 22.312, 22.762, 89.938, 89.962, 90.2), Lattice Difference: 0.0, FOM: 0.0
SEM of Cell: (0.042, 0.051, 0.126, 0.335, 0.372, 0.238)
Suggested Space Group: P1 (No. 1) : 1.0(R), 1.0(C)

```

5.5 Batch Processing Datasets in Cell Mode

Shift to CellCorr Tab to continue.

1. Supply Unit Cell and Space Group information:

- Enter the space group (89) and unit cell dimensions (22.537, 22.537, 9.025, 90.0, 90.0, 90.0) in the respective fields.
- Click **Update Cell Parameters**, a confirmation message will appear.

2. Run XDS and View the result:

- Click **Run XDS with Cell** to perform batch processing electron diffraction data. The processing will be finished with several minutes.
- After batch processing, AutoLEI will analysis the processing results automatically and generate xdsrunner2.xlsx with all result. Click **Show Results** to show the statistics below the button.
- The table will contain information for all datasets which successful processed:
[A] Serial number, [B] Path of the data, [C] Space group, [D] Symmetry-adopted Unit cell, [E] Index% (calculated from SPOT.XDS), [F] Model ISa from XDS (extracted from CORRECT.LP), [G] R_{meas} , [H] $CC_{1/2}$, [I] Completeness, [J] Resolution.

AutoLEI 1.0.0 - /mnt/c/AutoLEI_demo/Part4_batch_data_processing_8datasets - 8 datasets

Input XDSRunner UnitCellCorr XDSRefine MergeData Cluster_Output Tools RealTime

Input Space group and unit cell parameters.

Providing unit cell and space group keywords for all datasets is helpful for later data merging. XDS will refine unit cells individually.
Blind unit cell searching should be performed in XDSrunner. Check results in xdsrunner.xlsx / estimate symmetry

Space group: 89 Unit cell: 22.537, 22.537, 9.025, 90.0, 90.0, 90.0

Update Cell Parameters

* Run XDS with updated .inp files.

Run XDS with Cell Stop Run

* Show running result

Show Results Update Results File Open Result File >>> xdsrunner2.xlsx

No.	Path	SG	Unit cell	Vol.	Index%	ISa	Rmeas	CC1/2	Completeness	Reso.
1	.../1/xds	89	22.5(1) 22.5(1) 9.01(4) 90 90 90	4580(50)	80.1	3.04	73.08	88.57	97.83	1.1
2	.../2/xds	89	22.6(1) 22.6(1) 8.89(1) 90 90 90	4537(50)	89.9	4.77	36.94	94.82	91.85	0.9
3	.../3/xds	89	22.22(5) 22.22(5) 9.12(4) 90 90 90	4504(30)	62.7	7.49	43.1	99.13	92.79	1.07
4	.../4/xds	89	22.5(1) 22.5(1) 8.88(2) 90 90 90	4511(40)	83.5	6.2	76.14	91.77	99.0	0.96
5	.../5/xds	89	22.6(2) 22.6(2) 8.892(9) 90 90 90	4543(70)	79.3	5.13	29.67	98.59	96.19	0.87
6	.../6/xds	89	22.5(1) 22.5(1) 9.21(5) 90 90 90	4653(40)	73.9	7.08	23.41	99.16	70.84	0.82
7	.../7/xds	89	22.6(1) 22.6(1) 8.86(1) 90 90 90	4538(30)	95.6	5.73	37.15	98.18	99.89	0.93
8	.../8/xds	89	22.26(7) 22.26(7) 9.46(5) 90 90 90	4685(30)	91.2	9.7	41.08	98.57	96.49	0.87

5.6 Refine Input Parameter

The summary table reveals two noteworthy observations. First, nearly all datasets exhibit high completeness, enabling structure determination using a single dataset. Second, the statistics and the resolution of the datasets is not very satisfying.

- Refine input parameters:
 - Check **Rotation Axis, Divergence & Mosaicity** and **Remove Scale Outlier**
 - Leave the Scale Outlier unchanged, as 2.0.
- Run XDS and View the result:
 - Click **Run XDS with Cell** to perform batch processing electron diffraction data. In the command window, the refinement process will be shown. After input parameter refinement, data reduction process will be started. The processing will be finished with several minutes.
 - After batch processing, AutoLEI will analysis the processing results automatically and update xdsrunner2.xlsx with all result. Click **Show Results** to show the statistics below the button.
 - The table will contain information for all datasets which successful processed:

[A] Serial number, [B] Path of the data, [C] Space group, [D] Symmetry-adopted Unit cell, [E] Index% (calculated from SPOT.XDS), [F] Model ISa from XDS (extracted from CORRECT.LP), [G] R_{meas} , [H] $CC_{1/2}$, [I] Completeness, [J] Resolution.
 - All datasets will now have improved results on better resolution, $CC_{1/2}$ and lower R_{meas} . Datasets 5 and 6 has the highest resolution, however, their completeness are both low than 90%. To achieve a higher completeness with a higher resolution, the data can be merged.

AutoLEI 1.0.0 - /mnt/c/AutoLEI_demo/Part4_batch_data_processing_8datasets - 8 datasets

Input XDSRunner UnitCellCorr XDSRefine MergeData Cluster_Output Tools RealTime

Refine Input Parameters in XDS.INP. Get data reduction result from single dataset.

Note: This step is optional. Use models to refine XDS.INP files in the target folder.

Refine on data as

☒ Rotation Axis ☒ Divergence & Mosaicity ☒ Remove Scale Outlier > IQR ☐ Beam Centre

☐ Refine Index Ratio on datasets with index% < % ☐ Change Resolution to 0.8

Run XDS with Cell Stop Run Change Input Parameters

Show Results Update Results File Open Results File

No.	Path	SG	Unit cell						Vol.	Index%	ISa	Rmeas	CC1/2	Completeness	Reso.
1	.../1/xds	89	22.5(1)	22.5(1)	9.02(5)	90	90	90	4581(50)	80.0	3.44	64.5	95.82	97.07	1.04
2	.../2/xds	89	22.6(1)	22.6(1)	8.89(1)	90	90	90	4536(50)	90.1	7.15	24.24	99.05	91.43	0.9
3	.../3/xds	89	22.22(5)	22.22(5)	9.11(5)	90	90	90	4501(30)	61.6	5.84	41.86	97.91	90.67	1.07
4	.../4/xds	89	22.5(1)	22.5(1)	8.88(1)	90	90	90	4510(40)	83.4	6.52	51.91	95.6	98.81	0.96
5	.../5/xds	89	22.6(2)	22.6(2)	8.889(2)	90	90	90	4558(80)	81.3	6.42	26.19	99.35	88.24	0.8
6	.../6/xds	89	22.5(1)	22.5(1)	9.21(5)	90	90	90	4655(40)	73.8	8.14	22.04	99.0	72.25	0.84
7	.../7/xds	89	22.6(1)	22.6(1)	8.86(1)	90	90	90	4540(30)	95.8	5.96	34.25	98.09	99.1	0.93
8	.../8/xds	89	22.25(7)	22.25(7)	9.48(5)	90	90	90	4693(40)	93.7	11.02	38.28	98.53	96.22	0.87

5.7 Merge Data

For data merging, only datasets of excellent quality should be selected. While there is not a direct indicator to describe the data quality for electron diffraction data, ISa (Metric of Model Fitting), Resolution, R_{meas} and $CC_{1/2}$ can be used for comparison. If another standard is used to filter, **Manually Filter** can be used.

- Filter Data Using $CC_{1/2}$:
 - In DataMerge tab, select **CC1/2** from the drop-down menu. Use the default value (95) and click **Filter Data**. All datasets will be selected for further merging.
- Merge Data:
 - Click **Merge Data**. The selected datasets from xdspicker will be merged.
 - Click **Show Result** to check the statistical table of the data merging process.
 - After this step, a merge subfolder will be generated with the all.HKL file.

AutoLEI 1.0.0 - /mnt/c/AutoLEI_demo/Part4_batch_data_processing_8datasets - 8 datasets

Input XDSRunner UnitCellCorr XDSRefine MergeData Cluster_Output Tools RealTime

Generate and Merge data from xdspicker.xlsx.

Note: Average unit cell parameters will be used during merging. Generate .hkl and .P4P files for SHELX.

I. Filter data for merging

Use the data with CC1/2 better than 95 for merging. Filter Data Manually Filter

II. Merge Data

Merge Data Show Result Open XSCALE.LP * Strongly recommend to cluster before Bus to SHELX

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE >= -3.0 AS FUNCTION OF RESOLUTION	RESOLUTION LIMIT	NUMBER OF REFLECTIONS OBSERVED	UNIQUE POSSIBLE	COMPLETENESS OF DATA	R-FACTOR observed	R-FACTOR expected	R-FACTOR COMPARED	I/SIGMA	R-meas	CC(1/2)	Anomalous Corr	SigAno	Nano
3.58	1192	47	47	100.0%	30.2%	25.8%	1192	13.87	30.8%	94.4*	-34	0.392	11
2.53	2282	66	66	100.0%	27.5%	27.5%	2282	13.69	27.9%	99.8*	-38	0.574	30
2.07	2988	78	78	100.0%	31.5%	30.1%	2988	15.91	32.0%	99.0*	-55	0.550	43
1.79	3618	90	90	100.0%	35.8%	32.5%	3618	14.92	36.4%	98.8*	-44	0.577	55
1.60	4172	100	100	100.0%	45.5%	42.6%	4172	13.26	46.1%	98.9*	-21	0.687	62
1.46	4641	110	110	100.0%	54.1%	49.2%	4641	12.09	54.8%	97.3*	5	0.704	72
1.35	5210	116	116	100.0%	49.6%	47.0%	5210	13.15	50.2%	98.7*	-16	0.746	81
1.27	4910	112	112	100.0%	58.3%	54.9%	4910	12.08	59.0%	98.0*	-25	0.675	80
1.19	6141	137	137	100.0%	71.1%	68.3%	6141	10.93	71.9%	98.5*	-6	0.736	96
1.13	5852	126	126	100.0%	84.7%	82.2%	5852	9.26	85.0%	96.7*	-14	0.610	96
1.08	6895	154	154	100.0%	89.5%	85.6%	6895	8.80	90.5%	97.2*	-16	0.699	111
1.03	6104	144	144	100.0%	125.2%	123.7%	6104	7.16	126.7%	94.6*	-12	0.675	110
0.99	4937	145	145	100.0%	115.4%	112.3%	4937	6.04	117.2%	94.6*	-2	0.762	112
0.96	5034	163	163	100.0%	133.0%	130.8%	5034	4.80	135.2%	92.4*	-17	0.636	124
0.92	3709	165	165	100.0%	148.8%	147.3%	3709	3.67	152.2%	88.7*	10	0.781	131
0.90	2521	161	161	100.0%	160.8%	161.8%	2521	2.64	166.2%	82.0*	-5	0.692	130
0.87	1711	171	180	95.0%	252.8%	249.2%	1708	1.51	267.0%	50.9*	3	0.677	125
0.84	759	155	179	86.6%	184.6%	183.4%	742	1.25	204.5%	40.4*	-3	0.663	88
0.82	362	122	178	68.5%	207.7%	221.3%	332	0.73	244.0%	47.6*	-8	0.591	38
0.80	244	105	183	57.4%	189.2%	186.7%	215	0.59	234.2%	33.8	-3	0.544	20
total	73282	2467	2634	93.7%	51.6%	48.9%	73203	7.22	52.4%	96.9*	-9	0.680	1615

5.8 Intensity-Cluster

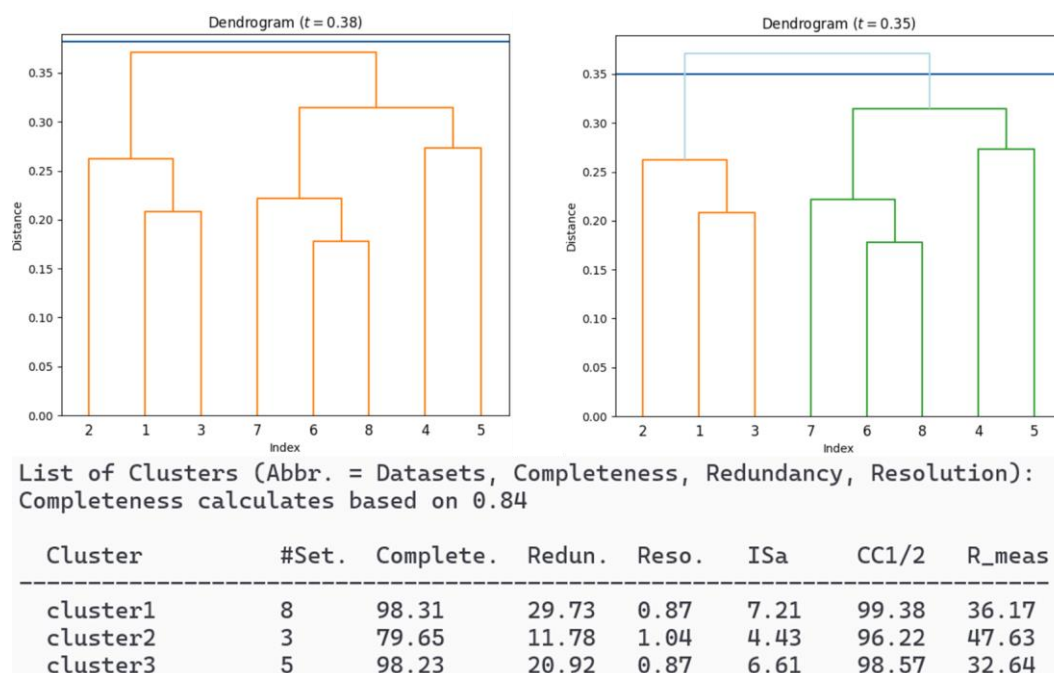
Even with the same cell parameters and symmetry, the contents of the crystals can be distinct. Thus, the “distance” between the datasets needs to be measured as an indicator to remove outliers.

1. Generate Dendrogram:

- Go to the **Cluster_Output** tab and click the **Set Distance from Dendrogram** button. A dendrogram will be saved in the merge subfolder and automatically open in a new window.
- By clicking on the diagram, the result cluster(s) based on the cut-off distance is coloured.

2. Set Cut-off Distance:

- The cut-off distance will be automatically returned to the distance entry box. In this tutorial, we will use two values: the first one is 0.38, the second one is 0.35.
- Enter **0.38** in the distance box or clicking the corresponding distance axis and click **Make Cluster based on Distance**.
- Untick** the checkbox for overwriting (**Important!**), then enter 0.35 in the distance box or clicking the corresponding distance axis and click **Make Cluster**.



3. Show Summary and View Merge Report:

- Click **Refresh and Show Summary** to see the summary from all.hkl for each cluster. In this tutorial, select **cluster1** (with 8 datasets) from the drop-down menu.
- The dendrogram of corresponding cluster can be viewed by clicking **Open Dendrogram**.
- Generate Web Report:
 - Click **Open Report** to generate and open the web report for the data merge result of cluster 1.
 - The web report includes three tables:
 1. Collection information of single datasets
 2. Overall statistics on data reduction results of single datasets
 3. Sub-Resolution-Shell Statistics of merged datasets
 4. Rawdata Path List of single datasets
 - Below the tables, there are two parts of plots, Statistics over Resolution and Dendrogram.

5.9 Generate and Update INS File

1. Generate XDS.INP:

The below tutorial is made based on XPREP. Diverse ways can be used to generate proper .INS file for shelx. If the .INS file is generated, ignore this step and go to the next step.

- To get the INS file for structural solving, click Run XPREP. This will run xprep in Linux or xprep.exe in Windows.
- For electron diffraction data, due to the inevitable dynamic effects, the systematic absence of the space group will not be strictly obeyed. For this tutorial, it is clear that the structure is in the *tP* Bravais lattice, while *-c-* has comparably low statistics on the exceptions (especially on the number of strong points

and mean intensity). Additionally, the $\langle E^2 - 1 \rangle$ value indicates that the non-centric space group. By consulting the ITC vol A, the most likely space group is $P\bar{4}c2$ (No. 116), which can be input manually.

Lattice exceptions:	P	A	B	C	I	F	Obv	Rev	All
N (total) =	0	31178	31205	31141	31216	46762	41519	41561	62403
N (int>3sigma) =	0	6526	6534	6792	6854	9926	8938	8990	13432
Mean intensity =	0.0	5.6	5.8	5.8	5.9	5.7	6.0	6.0	6.0
Mean int/sigma =	0.0	1.6	1.7	1.7	1.7	1.7	1.7	1.7	1.7
Lattice type [P, A, B, C, I, F, O(obv.), R(rev. rhomb. on hex. axes)]									
Select option [P]:									
Mean E-E-1 = 0.732 [expected .968 centrosym and .736 non-centrosym]									
Systematic absence exceptions:									
	41/43	42	n--	-b-	-c-	-n-	-21-	--c	
N	28	21	2475	1893	1898	1859	110	1197	
N I>3s	8	1	1013	515	12	513	73	189	
<I>	31.6	2.1	8.5	17.2	0.9	17.4	24.8	4.3	
<I/s>	1.4	0.8	2.5	1.6	0.6	1.6	3.1	1.4	
Identical indices and Friedel opposites combined before calculating R(sym)									
Option	Space Group	No.	Type	Axes	CSD	R(sym)	N(eq)	Syst. Abs.	CFOM
No acceptable space group - change tolerances or unset chiral flag or possibly change input lattice type, then recheck cell using H-option									
Select option [?]:									

2. Embedded Metadata:

The metadata is lost during the data reduction process, which is crucial to the final report and checkcif procedure.

- After got corresponding , fill in the necessary information in the boxes.
 - From the drop-down menu, select the **TimePix-512-2100_SU**.
 - Tick the **Cryoholder** checkbox it has been used in the experiment.
 - The short name is **CAU-36**.
 - The long name is **CAU-36 with DABCO**.
- Click **Update INS and Metadata** to generate a new .INS file named CAU-36.ins, and the metadata is stored in CAU-36.cif_od.
- Click **Open Folder** and the structure solvation can be started.

Contents of CAU-36.cif_od:

```
data_CAU-36
_symmetry_cell_setting      Tetragonal'
_symmetry_space_group_name_H-M  P-4c2
_cell_measurement_temperature  100(3)
_cell_measurement_reflns_used  14398
_cell_measurement_theta_min    0.0360
_cell_measurement_theta_max    0.8769
_exptl_crystal_description     nano-crystal
_exptl_crystal_colour          ?
_exptl_crystal_size_min        ?
_exptl_crystal_size_mid        ?
_exptl_crystal_size_max        ?
_exptl_crystal_density_meas    ?
_exptl_crystal_density_method  ?
_exptl_absorpt_correction_type none
```

```

_exptl_absorpt_correction_T_min      ?
_exptl_absorpt_correction_T_max      ?
_diffraction_ambient_temperature      100(3)
_diffraction_radiation_type            electron
_diffraction_radiation_source          electron
_diffraction_radiation_monochromator    ?
_diffraction_measurement_device_type    JEOL-2100
_diffraction_measurement_method        'Continuous Rotation Electron Diffraction'
_diffraction_detector_area_resol_mean  18.1818
_diffraction_standards_number          0
_diffraction_standards_interval_count  ?
_diffraction_standards_decay_%         ?
_computing_data_collection             'Instamatic (Wang et al., 2019)'
_computing_cell_refinement             'XDS (Kabsch et al., 2010)'
_computing_data_reduction              'XDS (Kabsch et al., 2010)'
_computing_structure_solution          'SHELXS-97 (Sheldrick 2008)'
_computing_structure_refinement        'SHELXL-2014 (Sheldrick 2014)'
_computing_molecular_graphics          'Bruker SHELXTL'
_computing_publication_material        'Bruker SHELXTL'
_chemical_name_common                 'CAU-36 with DABCO'
_diffraction_ambient_environment       vacuum
_diffraction_radiation_probe           electron
_diffraction_radiation_wavelength      0.0251
_diffraction_source                   'electron microscope'
_diffraction_source_voltage            200
_diffraction_detector                 TimePix-SU
_diffraction_measurement_device         TEM
_diffraction_measurement_details
;

```

List of Samples for Merging (Abbr. = Camera Length, Resolution):

#	Type	Start	End	Step	t~exp~	#Frames	CL	Reso.	ISa	CC1/2
1	/a	-40.4	55.23	0.225	0.096288	425	531.85	1.04	1.34	94.6
2	/a	-44.0	31.6	0.225	0.096288	336	531.85	0.9	3.42	98.8
3	/a	-43.1	28.0	0.225	0.096288	316	531.85	1.07	2.02	98.9
4	/a	-43.5	44.25	0.225	0.096288	390	531.85	0.96	3.08	97.0
5	/a	-44.0	23.5	0.225	0.096288	300	531.85	0.8	3.38	99.4
6	/a	-50.5	38.38	0.225	0.096288	395	531.85	0.84	5.58	99.2
7	/a	-50.5	52.1	0.225	0.096288	456	531.85	0.93	3.45	97.6
8	/a	-42.5	34.67	0.225	0.096288	343	439.13	0.87	3.33	98.7

;

6. Tutorial Example 3: SU-100

6.1 General information

- **Raw Data:** The SU-100 datasets can be downloaded from <https://doi.org/10.5281/zenodo.13835773>
- **Data Collection:** 16 datasets were collected using an ASI TimePix hybrid detector installed on a JEOL JEM-2100 (200 kV) microscope equipped with a LaB₆ filament. Data was collected at cryo temperature with a Gatan 914 cryo-holder.
- **Collection Method:** Data collection utilised Instamatic, with every 20th frame defocused to track the crystal.
- **References:**
 - Instamatic details: [<https://zenodo.org/records/11091622>]
 - Sample and data collection: [AutoLEI paper]
 - Final structure deposited in CCDC: [AutoLEI CCDC]
- **Sample Details:** SU-100 (BiC₁₅H₇O₆), space group: *C2/c* or *I2/a*, unit cell dimensions: 26.44 Å, 10.050 Å, 18.131 Å, 90°, 127.264°, 90° (*mC*); 18.131 Å, 10.050 Å, 21.15 Å, 90°, 95.76°, 90° (*mI*)
- **Software Versions:** Tutorial conducted using AutoLEI version 1.0 and XDS version 2024/07/23.

6.2 Goals

1. Understand the workflow of AutoLEI.
2. Perform batch processing of electron diffraction data using provided XDS.INPs.
3. Finalise data processing by merging dataset.

6.3 Load Working Directory

1. Start AutoLEI:
 - If using a pip-installed version, activate the Python environment in Linux and type **autolei**.
 - Alternatively, navigate to the folder containing **autolei.py** and run it.
 - After version and scaling information appear in the command window, the main AutoLEI window will open.
 - **Note:** If a message appears stating “XDS not existing! Do you wish to continue?”, ensure XDS is accessible to AutoLEI (e.g., added to the system path).
2. Select Input Path:
 - Click **Browse** to select the folder containing the 16 datasets of SU-100.
 - Click **Load**. Once loaded, the title updates to display the selected path and the number of datasets. Input parameters will also reflect communal settings.
 - Proceed to the XDSRunner tab as XDS.INPs are pre-generated.

AutoLEI 1.0.0 - /mnt/c/AutoLEI_demo/SU-100 - 16 datasets

Input XDSRunner UnitCellCorr XDSRefine MergeData Cluster_Output Tools RealTime

Browse and load the work path where the program will load the measurement settings.
For XDS input generation, supply basic parameters and click the 'Save Parameter' button.

Input path: /mnt/c/AutoLEI_demo/SU-100 Browse Load Path Instrument File: Custom Load

I. Instrument Parameters

1. Detector parameters: NX= 516 NY= 516 QX= 0.0550 QY= 0.0550

2. Overloading: OVERLOAD= 130000 3. Wavelength: WAVELENGTH= 0.0251 Å

4. Rotation axis: ROTATION_AXIS= -0.6204 0.7843 0.0000 >>> Use space to segment OR Angle in Degree = -128.35

5. Additional information
(Please copy from XDS):
UNTRUSTED_RECTANGLE= 255 262 0 517
UNTRUSTED_RECTANGLE= 0 517 255 262

II. Measurement Parameters

6. Direct beam position: ORGX= ORGY=

7. Resolution range: INCLUDE_RESOLUTION_RANGE= 20 0.8 >>> Use space to segment

8. Camera Length: DETECTOR_DISTANCE= +439.48 9. Rotation step: OSCILLATION_RANGE=

Save Parameters

6.4 Batch Processing Datasets in P1 Mode

1. Correct the XDS.INP:
 - Since the data is in SMV .img format, no conversion is necessary.
 - Select SMV format and click **Update Instamatic XDS.INP** to batch edit all XDS.INP files, removing illegal characters and obsolete keywords.
2. Run XDS and View the result
 - Click **Run XDS** to perform batch processing electron diffraction data. The processing will be finished with several minutes. Processing time varies with the operating system and hardware performance (e.g., WSL1 is faster than WSL2).
 - The process completes when the command window shows “All information extracted. Note that resolution is ideal,” or the animation in AutoLEI stops. Monitor progress in the command window’s bottom line.
 - After batch processing, AutoLEI will analysis the processing results automatically and generate xdsrunner.xlsx with all result. Click **Show Results** to show the statistics below the button.
 - The table will contain information for all datasets which successful processed:
 - [A] Serial Number
 - [B] Data Path
 - [C] P1 Unit Cell
 - [D] Space Group
 - [E] Symmetry-Adopted Unit Cell

- [F] Index% (from SPOT.XDS)
- [G] Model ISa (from CORRECT.LP)
- [H] $CC_{1/2}$
- [I] Completeness
- [J] Resolution

No.	Path	Integration Cell	SG	Unit cell	Vol.	Index%	ISa	CC1/2	Completeness	Reso.
1	...experiment_1/SMV	10.452 14.062 15.413 75.315 71.905 69.90	5	26.23(9) 10.2(2) 18.047(9) 90 126.97(9) 90	3851(90)	77.5	8.71	98.34	56.99	0.8
2	...experiment_2/SMV	10.057 14.084 15.449 91.21 109.057 110.8	5	26.32(1) 10.012(9) 18.08(1) 90 127.01(5) 90	3804(6)	89.4	5.9	96.41	95.98	0.8
3	...experiment_3/SMV	10.019 14.155 15.337 75.173 70.97 69.368	5	26.52(2) 9.975(6) 18.018(2) 90 127.40(2) 90	3786(4)	88.1	7.21	97.86	68.41	0.8
4	...experiment_4/SMV	9.974 14.088 15.365 75.551 70.955 69.85	5	26.53(1) 10.01(4) 18.10(1) 90 127.61(5) 90	3806(20)	93.4	7.97	97.62	76.68	0.8
5	...experiment_5/SMV	10.423 14.09 15.367 90.854 108.238 110.1	5	26.41(2) 10.1(2) 18.16(3) 90 127.65(4) 90	3829(100)	87.9	14.5	98.9	75.61	0.8
6	...experiment_6/SMV	10.035 14.116 15.407 75.081 71.073 69.39	5	26.47(6) 10.03(1) 18.02(3) 90 127.05(9) 90	3818(10)	95.5	6.53	97.25	62.48	0.8
7	...experiment_7/SMV	9.976 14.252 15.352 91.254 109.695 110.6	1	9.9(5) 14.3(4) 15.1(3) 92(3) 108(2) 110(3)	1887(100)	88.2	4.34	95.98	11.49	0.87
8	...experiment_8/SMV	10.087 14.149 15.391 75.041 71.124 69.23	5	26.461(8) 10.06(1) 18.022(9) 90 127.09(2) 90	3827(5)	88.9	8.37	98.05	54.95	0.8
9	...experiment_9/SMV	9.989 14.148 15.292 75.556 70.975 69.6	5	26.58(3) 9.954(9) 18.07(2) 90 127.79(6) 90	3778(8)	93.2	7.36	96.99	69.16	0.8
10	...experiment_10/SMV	10.071 14.179 15.352 75.248 70.915 69.12	5	26.509(9) 10.068(8) 18.02(2) 90 127.48(3) 90	3816(6)	86.5	9.97	97.31	76.5	0.82
11	...experiment_11/SMV	10.099 10.314 10.568 95.298 90.212 118.8	5	18.082(8) 10.06(2) 10.57(1) 90 96.19(4) 90	1911(5)	52.9	2.62	86.86	57.69	0.82
12	...experiment_12/SMV	10.108 14.138 15.515 75.021 70.884 68.82	5	26.35(2) 10.08(1) 18.09(1) 90 126.89(7) 90	3844(8)	62.1	6.11	95.45	56.65	0.8
13	...experiment_13/SMV	10.054 14.088 15.32 91.106 108.69 110.66	1	10.04(1) 14.09(1) 15.271(6) 91.0(2) 108.5(1) 110.	1895(4)	85.6	7.84	97.47	58.31	0.8
14	...experiment_14/SMV	9.997 14.148 15.422 91.728 108.711 110.2	5	26.57(2) 9.962(6) 18.16(2) 90 127.44(4) 90	3817(7)	43.7	10.92	98.37	75.59	0.8
15	...experiment_15/SMV	9.899 14.112 15.358 75.345 71.688 69.872	5	26.55(3) 9.79(1) 18.05(1) 90 127.03(8) 90	3746(8)	92.5	9.37	98.48	78.62	0.8
16	...experiment_16/SMV	10.028 14.151 15.38 75.798 71.091 69.084	5	26.46(2) 10.01(1) 18.18(1) 90 127.65(1) 90	3814(5)	96.8	14.1	99.22	75.26	0.8

3. Estimate Symmetry:

- Click **Estimate Symmetry & Cell-Cluster**. Results are displayed in the command window and saved as lattice_cluster.txt in the working directory.
- The cell cluster report provides detailed analysis on:
 - Potential Symmetry: Indicates to which symmetry the datasets align.
 - Cluster Analysis: Reveals that 15 of 16 datasets form Cluster 1 (similar unit cell and high symmetry), while the 11th dataset is classified as Cluster 2 due to slightly differing unit cell parameters and low symmetry.
- Based on Cluster 1, the primary unit cell is identified as having a C-centred monoclinic lattice and Laue group $2/m$ with:
 - R_{meas} ratio increased slightly: 1.243
 - CC ratio increased slightly: 1.290
- For further processing, use the refined unit cell parameters: 26.473, 10.087, 18.08 (in Å) and angles 90° , 127.3° , 90° . Adopt space group 5 (the base group for $C2/m$).

```

*****
*      Lattice Symmetry Explorer      *
*****

A script using to estimate lattice symmetry based on reflection data statistics.
The recommended lattice difference should be < 2.5, and the Figure of Merit (FOM) should be < 200.
For a given Bravais lattice, multiple space groups may be available. The R_meas ratio of each space
group to P1 will be provided following the space group name, with a recommended value of < 1.8

      R_meas ratio = R_meas(Space Group) / R_meas(P1)
      CC1/2 ratio = [1 - CC1/2(Space Group)] / [1 - CC1/2(P1)]

Testing Lattice Symmetry: 100%[#####] 16/16 [00:00]

***** Cluster-1: *****
** Data Path: 15 datasets, (93.8%)
/mnt/c/AutoLEI_demo/SU-100/experiment_1/SMV
/mnt/c/AutoLEI_demo/SU-100/experiment_2/SMV
/mnt/c/AutoLEI_demo/SU-100/experiment_3/SMV
/mnt/c/AutoLEI_demo/SU-100/experiment_4/SMV
/mnt/c/AutoLEI_demo/SU-100/experiment_5/SMV
/mnt/c/AutoLEI_demo/SU-100/experiment_6/SMV
/mnt/c/AutoLEI_demo/SU-100/experiment_7/SMV
/mnt/c/AutoLEI_demo/SU-100/experiment_8/SMV
/mnt/c/AutoLEI_demo/SU-100/experiment_9/SMV
/mnt/c/AutoLEI_demo/SU-100/experiment_10/SMV
/mnt/c/AutoLEI_demo/SU-100/experiment_12/SMV
/mnt/c/AutoLEI_demo/SU-100/experiment_13/SMV
/mnt/c/AutoLEI_demo/SU-100/experiment_14/SMV
/mnt/c/AutoLEI_demo/SU-100/experiment_15/SMV
/mnt/c/AutoLEI_demo/SU-100/experiment_16/SMV

** Lattice Choice:
-- Bravais Lattice: mC (setting 1)
Averaged Cell: (26.473, 10.087, 18.08, 90.0, 127.313, 90.0), Lattice Difference: 0.076, FOM: 13.4
SEM of Cell: (0.024, 0.04, 0.021, 0.0, 0.082, 0.0)
Suggested Space Group: C121 (No. 5) : 1.243(R), 1.191(C)

-- Bravais Lattice: aP
Averaged Cell: (10.087, 14.12, 15.367, 82.747, 88.727, 88.58), Lattice Difference: 0.0, FOM: 0.0
SEM of Cell: (0.04, 0.01, 0.018, 2.07, 4.825, 5.277)
Suggested Space Group: P1 (No. 1) : 1.0(R), 1.0(C)

***** Cluster-2: *****
** Data Path: 1 datasets, (6.2%)
/mnt/c/AutoLEI_demo/SU-100/experiment_11/SMV

```

6.5 Batch Processing Datasets in Cell Mode

Shift to **CellCorr** Tab to continue.

1. Supply Unit Cell and Space Group information:

- Enter the space group (5) and unit cell dimensions (26.473, 10.087, 18.08, 90.0, 127.313, 90.0) in the respective fields.
- Click **Update Cell Parameters**, a confirmation message will appear.

2. Run XDS and View the result:

- Click **Run XDS with Cell** to perform batch processing electron diffraction data. The processing will be finished with several minutes.
- After batch processing, AutoLEI will analysis the processing results automatically and generate xdsrunner2.xlsx with all result. Click **Show Results** to show the statistics below the button.

- The table will contain information for all datasets which successful processed:

[A] Serial number, [B] Path of the data, [C] Space group, [D] Symmetry-adopted Unit cell, [E] Index% (calculated from SPOT.XDS), [F] Model ISa from XDS (extracted from CORRECT.LP), [G] R_{meas} , [H] $CC_{1/2}$, [I] Completeness, [J] Resolution.

AutoLEI 1.0.0 - /mnt/c/AutoLEI_demo/SU-100 - 16 datasets

Input XDSRunner UnitCellCorr XDSRefine MergeData Cluster_Output Tools RealTime

Input Space group and unit cell parameters.

Providing unit cell and space group keywords for all datasets is helpful for later data merging. XDS will refine unit cells individually.
Blind unit cell searching should be performed in XDSrunner. Check results in xdsrunner.xlsx / estimate symmetry

Space group: 5 Unit cell: 26.473, 10.087, 18.08, 90.0, 127.313, 90.0

Update Cell Parameters

* Run XDS with updated .inp files.

Run XDS with Cell Stop Run

* Show running result

Show Results Update Results File Open Result File >>> xdsrunner2.xlsx

No.	Path	SG	Unit cell						Vol.	Index%	ISa	Rmeas	CC1/2	Completeness	Reso.
1	.../experiment_1/SMV	5	26.12(8)	10.37(4)	18.05(1)	90	127.18(2)	90	3893(20)	76.8	7.47	16.01	98.4	55.86	0.8
2	.../experiment_2/SMV	5	26.32(7)	10.01(1)	18.08(1)	90	127.01(3)	90	3804(10)	89.4	5.92	32.66	96.08	95.76	0.8
3	.../experiment_3/SMV	5	26.52(8)	9.97(2)	18.02(3)	90	127.40(9)	90	3786(20)	88.2	7.24	20.73	97.95	68.34	0.8
4	.../experiment_4/SMV	5	26.52(7)	10.04(2)	18.09(8)	90	127.6(3)	90	3816(30)	93.2	8.3	20.82	97.77	77.08	0.8
5	.../experiment_5/SMV	5	26.37(8)	10.3(1)	18.08(4)	90	127.6(1)	90	3882(60)	85.9	9.36	21.97	98.33	72.84	0.8
6	.../experiment_6/SMV	5	26.47(8)	10.03(3)	18.02(3)	90	127.0(1)	90	3819(20)	95.8	6.45	17.44	97.32	62.64	0.8
7	.../experiment_7/SMV	5	26.6(4)	10.0(2)	18.037(9)	90	127.3(3)	90	3816(100)	87.5	4.25	30.62	96.39	21.57	0.87
8	.../experiment_8/SMV	5	26.5(1)	10.06(2)	18.02(6)	90	127.1(1)	90	3827(20)	89.6	8.41	18.11	98.04	55.22	0.8
9	.../experiment_9/SMV	5	26.6(1)	9.96(3)	18.07(3)	90	127.8(1)	90	3776(20)	93.1	6.87	32.72	96.93	69.25	0.8
10	.../experiment_10/SMV	5	26.51(4)	10.05(2)	18.03(3)	90	127.46(7)	90	3812(10)	86.8	10.28	33.54	97.39	76.79	0.82
11	.../experiment_11/SMV	5	26.30(5)	10.06(2)	18.072(9)	90	126.89(2)	90	3825(10)	79.6	2.92	48.0	83.88	57.32	0.8
12	.../experiment_12/SMV	5	26.35(5)	10.080(6)	18.09(1)	90	126.90(4)	90	3843(10)	62.1	6.18	30.96	95.98	56.71	0.8
13	.../experiment_13/SMV	5	26.3(1)	10.11(1)	18.15(4)	90	127.7(2)	90	3813(20)	85.8	8.05	35.4	95.54	74.41	0.8
14	.../experiment_14/SMV	5	26.57(9)	9.96(3)	18.16(2)	90	127.53(3)	90	3811(20)	65.7	10.48	19.76	98.27	75.26	0.8
15	.../experiment_15/SMV	5	26.51(9)	9.82(3)	18.04(3)	90	127.0(1)	90	3750(20)	93.5	7.79	22.04	98.15	79.2	0.8
16	.../experiment_16/SMV	5	26.46(4)	10.009(8)	18.18(1)	90	127.65(8)	90	3814(9)	96.8	17.2	16.6	99.23	74.82	0.8

5.6 Refine Input Parameter

The summary table reveals two noteworthy observations. First, nearly all datasets exhibit low completeness, making it implausible to have structure determination using a single dataset. Second, the reported resolution values are commonly at 0.8, which is the upper limit of the resolution range.

1. Refine input parameters:

- Check **Rotation Axis, Divergence & Mosaicity, Remove Scale Outlier** and **Change Resolution**
- Change resolution to “30 0.5”.

2. Run XDS and View the result:

- Click **Run XDS with Cell** to perform batch processing electron diffraction data. In the command window, the refinement process will be shown. After input parameter refinement, data reduction process will be started. The processing will be finished with several minutes.
- After batch processing, AutoLEI will analysis the processing results automatically and update xdsrunner2.xlsx with all result. Click **Show Results** to show the statistics below the button.
- The table will contain information for all datasets which successful processed:
[A] Serial number, [B] Path of the data, [C] Space group, [D] Symmetry-adopted Unit cell, [E] Index% (calculated from SPOT.XDS), [F] Model ISa from XDS (extracted from CORRECT.LP), [G] R_{meas} , [H] $CC_{1/2}$, [I] Completeness, [J] Resolution.

- All datasets will now have improved results on better resolution, $CC_{1/2}$ and lower R_{meas} . To achieve a higher completeness with a higher resolution, the data can be merged.

No.	Path	SG	Unit cell						Vol.	Index%	ISa	Rmeas	CC1/2	Completen	Reso.
1	.../experiment_1/SMV	5	26.24(7)	10.16(1)	18.06(1)	90	127.07(5)	90	3843(10)	79.9	7.73	16.79	98.21	43.18	0.62
2	.../experiment_2/SMV	5	26.34(7)	10.01(1)	18.09(1)	90	127.02(4)	90	3808(10)	89.6	5.86	34.06	96.54	82.6	0.66
3	.../experiment_3/SMV	5	26.53(8)	9.98(3)	18.03(4)	90	127.4(1)	90	3790(20)	88.6	7.21	21.8	98.44	41.38	0.55
4	.../experiment_4/SMV	5	26.52(7)	10.04(2)	18.09(7)	90	127.6(2)	90	3814(30)	92.9	7.93	21.04	97.76	67.92	0.68
5	.../experiment_5/SMV	5	26.38(8)	10.3(1)	18.09(3)	90	127.58(6)	90	3882(40)	85.4	10.89	29.9	98.22	52.31	0.61
6	.../experiment_6/SMV	5	26.48(7)	10.03(3)	18.04(4)	90	127.1(1)	90	3822(20)	95.9	9.63	13.18	98.64	33.47	0.55
7	.../experiment_7/SMV	5	26.6(4)	10.0(2)	18.037(4)	90	127.3(4)	90	3814(100)	88.2	4.95	25.22	96.65	20.79	0.87
8	.../experiment_8/SMV	5	26.5(1)	10.06(2)	18.03(6)	90	127.1(1)	90	3830(30)	89.6	8.72	20.8	98.2	29.71	0.54
9	.../experiment_9/SMV	5	26.6(1)	9.96(3)	18.07(3)	90	127.8(1)	90	3781(20)	93.0	6.32	34.19	97.42	55.69	0.64
10	.../experiment_10/SMV	5	26.51(5)	10.05(2)	18.03(3)	90	127.47(9)	90	3813(20)	87.0	10.42	33.14	97.24	73.97	0.76
11	.../experiment_11/SMV	5	26.30(4)	10.04(1)	18.089(8)	90	126.89(2)	90	3820(9)	82.4	3.4	40.79	86.05	54.66	0.78
12	.../experiment_12/SMV	5	26.36(4)	10.082(6)	18.10(1)	90	126.91(5)	90	3845(8)	62.0	6.36	33.84	95.81	53.45	0.7
13	.../experiment_13/SMV	5	26.3(1)	10.11(2)	18.15(3)	90	127.8(1)	90	3814(20)	85.0	8.6	42.28	95.24	68.81	0.68
14	.../experiment_14/SMV	5	26.57(9)	9.96(3)	18.16(2)	90	127.55(5)	90	3809(20)	66.1	11.28	21.21	98.46	44.08	0.55
15	.../experiment_15/SMV	5	26.5(1)	9.81(3)	18.04(4)	90	127.0(1)	90	3747(20)	93.6	9.14	21.38	98.64	46.19	0.55
16	.../experiment_16/SMV	5	26.46(4)	10.010(7)	18.185(9)	90	127.64(7)	90	3814(8)	96.9	19.54	16.62	99.41	68.28	0.7

6.7 Merge Data

For data merging, only datasets of excellent quality should be selected. While there is not a direct indicator to describe the data quality for electron diffraction data, ISa (Metric of Model Fitting), Resolution, R_{meas} and $CC_{1/2}$ can be used for comparison. If another standard is used to filter, **Manually Filter** can be used.

1. Filter Data Using $CC_{1/2}$:
 - In DataMerge tab, select **ISa** from the drop-down menu. Use the default value (5) and click **Filter Data**. All datasets will be selected for further merging.
2. Merge Data:
 - Click **Merge Data**. The selected datasets from xdspicker will be merged.
 - Click **Show Result** to check the statistical table of the data merging process.
 - After this step, a merge subfolder will be generated with the all.HKL file.

AutoLEI 1.0.0 - /mnt/c/AutoLEI_demo/SU-100 - 16 datasets

Input XDSRunner UnitCellCorr XDSRefine MergeData Cluster_Output Tools RealTime

Generate and Merge data from xdspicker.xlsx.

Note: Average unit cell parameters will be used during merging. Generate .hkl and .P4P files for SHELX.

I. Filter data for merging

Use the data with better than for merging.

II. Merge Data

* Strongly recommend to cluster before

SUBSET OF RESOLUTION LIMIT	INTENSITY OBSERVED	DATA UNIQUE	WITH SIGNAL/NOISE POSSIBLE	COMPLETENESS OF DATA	R-FACTOR observed	AS FUNCTION OF R-FACTOR expected	RESOLUTION COMPARED	I/SIGMA	R-meas	CC(1/2)	Anomalous Corr	SigAno	Nano
2.42	3273	169	175	96.6%	23.5%	26.1%	3273	11.57	24.3%	97.6*	-48	0.567	111
1.71	5871	278	284	97.9%	27.5%	27.1%	5869	11.22	28.1%	97.7*	-46	0.584	217
1.40	7743	358	363	98.6%	33.7%	33.7%	7741	9.75	34.5%	95.8*	-24	0.640	284
1.21	9638	432	435	99.3%	37.9%	35.8%	9634	9.29	38.7%	97.5*	-42	0.584	363
1.08	10705	471	480	98.1%	43.4%	41.8%	10702	8.44	44.5%	97.5*	-31	0.631	394
0.99	12346	533	536	99.4%	54.5%	52.2%	12342	7.33	55.8%	92.9*	-33	0.579	458
0.91	13202	555	559	99.3%	66.6%	63.9%	13200	6.39	68.1%	91.6*	-25	0.639	490
0.85	14140	611	615	99.3%	79.3%	76.2%	14132	5.27	81.1%	88.7*	-30	0.607	523
0.81	13741	657	664	98.9%	86.8%	84.8%	13734	4.40	88.9%	86.5*	-22	0.656	568
0.76	12147	680	686	99.1%	106.4%	105.4%	12141	3.55	109.6%	76.8*	-12	0.669	594
0.73	9799	706	715	98.7%	117.9%	119.8%	9782	2.71	122.5%	70.8*	-17	0.651	581
0.70	6894	735	745	98.7%	131.3%	132.7%	6857	1.79	138.7%	54.8*	-6	0.675	558
0.67	4255	714	766	93.2%	139.4%	142.6%	4184	1.24	152.2%	36.6*	-6	0.658	411
0.65	2690	610	821	74.3%	130.7%	132.2%	2547	1.18	145.0%	34.6*	-5	0.663	246
0.62	1766	510	851	59.9%	154.2%	164.4%	1571	0.87	172.7%	53.3*	-3	0.640	152
0.60	983	327	877	37.3%	191.4%	198.5%	860	0.70	222.5%	23.2	-4	0.650	92
0.59	513	205	891	23.0%	530.7%	531.4%	428	0.52	632.6%	37.3*	-12	0.606	41
0.57	409	179	912	19.6%	186.2%	216.6%	326	0.67	222.2%	26.6	-5	0.690	30
0.55	326	159	952	16.7%	-99.9%	-99.9%	260	0.00	-99.9%	12.8	-34	0.385	13
0.54	90	54	979	5.5%	-99.9%	-99.9%	58	0.00	-99.9%	15.3	-43	0.732	4
total	130531	8943	13306	67.2%	40.9%	40.4%	129641	4.28	42.0%	98.1*	-21	0.636	6130

6.8 Intensity-Cluster

Even with the same cell parameters and symmetry, the contents of the crystals can be distinct. Thus, the “distance” between the datasets needs to be measured as an indicator to remove outliers.

1. Generate Dendrogram:

- Go to the **Cluster_Output** tab and click the **Set Distance from Dendrogram** button. A dendrogram will be saved in the merge subfolder and automatically open in a new window.
- By clicking on the diagram, the result cluster(s) based on the cut-off distance is coloured.

2. Set Cut-off Distance:

- The cut-off distance will be automatically returned to the distance entry box. In this tutorial, we will use three values: the first one is 0.60, the second one is 0.50 and the third is 0.45.
- Enter **0.60** in the distance box or clicking the corresponding distance axis and click **Make Cluster based on Distance**.
- Untick** the checkbox for overwriting (**Important!**), then enter **0.50** in the distance box or clicking the corresponding distance axis and click **Make Cluster**.
- Enter 0.45 in the distance box or clicking the corresponding distance axis and click **Make Cluster**.



3. Show Summary and View Merge Report:

- Click **Refresh and Show Summary** to see the summary from all.hkl for each cluster. In this tutorial, select **cluster2** (with 11 datasets) from the drop-down menu.
- The dendrogram of corresponding cluster can be viewed by clicking **Open Dendrogram**.
- Generate Web Report:
 - Click **Open Report** to generate and open the web report for the data merge result of cluster 2.
 - The web report includes three tables:
 1. Collection information of single datasets
 2. Overall statistics on data reduction results of single datasets
 3. Sub-Resolution-Shell Statistics of merged datasets
 4. Rawdata Path List of single datasets
 - Below the tables, there are two parts of plots, Statistics over Resolution and Dendrogram.

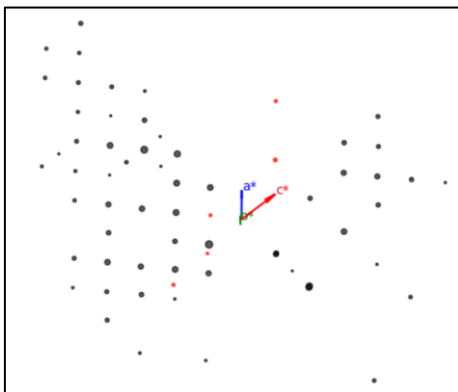
6.9 Generate and Update INS File

1. Generate XDS.INP:

The below tutorial is made based on XPREP. Diverse ways can be used to generate proper .INS file for shelx. If the .INS file is generated, ignore this step and go to the next step.

- To get the INS file for structural solving, click Run XPREP. This will run xprep in Linux or xprep.exe in Windows.
- For electron diffraction data, due to the inevitable dynamic effects, the systematic absence of the space group will not be strictly obeyed. For this tutorial, it is clear that the structure is in the *mC* Bravais lattice (or *mI*), with possible space group of *C2*, *Cm*, *C2/m*, *Cc* and *C2/c*.

- To see if there is really c glide inside the crystal, go to XDSRefine Tab. Select **Single** in the data filter and then select **experiment_13**. Click **View Reciprocal Space** to raise Lattice Visualisation. If the buttons do not appear below the image, resize the window. Select **H0L**, **View [010]**, and **Show Intensity**, the slice of $h0l$ is shown.



- From the plot, there will be some points which violate the rule. However, basically the feature of c glide is maintained. **Note:** the length of the reciprocal vector is doubled! ($2a^*$, $2b^*$, $2c^*$)
- The space group is determined as $C2/c$.

2. Embedded Metadata:

The metadata is lost during the data reduction process, which is crucial to the final report and checkcif procedure.

- After got corresponding , fill in the necessary information in the boxes.
 - From the drop-down menu, select the **TimePix-516-2100_SU**.
 - Tick the **Cryoholder** checkbox it has been used in the experiment.
 - The short name is **SU-100**.
 - The long name is **SU-100-cryo**.
- Click **Update INS and Metadata** to generate a new .INS file named SU-100.ins, and the metadata is stored in SU-100.cif_od.
- Click **Open Folder** and the structure solvation can be started.

AutoLEI 1.0.0 - /mnt/c/AutoLEI_demo/SU-100 - 16 datasets

Input XDSRunner UnitCellCorr XDSRefine MergeData Cluster_Output Tools RealTime

Intensity-Cluster based on Correlation Coefficients in XSCALE.LP

The distance can either gathered from the Dendrogram or manually input.

Set Distance from Dendrogram Distance 0.5985 ☐ Overwrite previous result Make Cluster based on Distance

Process Clusters and Generate .INS

Press Refresh to view the information of all clusters. Run XPREP will raise XPREP in Windows. Set the XPREP path in 'setting.ini' first!

Data Processing Based on cluster2 Refresh and Show Summary

Open Dendrogram Open XSCALE.LP Run XPREP Open Report

Collect and Generate Metadata File

Metadata will be updated with provided information and headers in .img file, and saved in the .CIF_OD file. Olex2 will pick it up automatically.
The compound name in short should be only one word.

Instrument Profile: TimePix-516_2100_SU

TEM Instrument Name JEOL-2100 Detector Name TimePix-SU Temperature 100 K ☒ Cryoholder

Compound Name Short Name: SU-100 Long Name: SU-100_cryo

Update INS and Metadata Open Folder

Contents of SU-100.cif_od:

```
data_SU-100
_symmetry_cell_setting      Monoclinic'
_symmetry_space_group_name_H-M C2/c
_cell_measurement_temperature 100(3)
_cell_measurement_reflns_used 24461
_cell_measurement_theta_min 0.0360
_cell_measurement_theta_max 1.2188
_exptl_crystal_description  nano-crystal
_exptl_crystal_colour      ?
_exptl_crystal_size_min    ?
_exptl_crystal_size_mid    ?
_exptl_crystal_size_max    ?
_exptl_crystal_density_meas ?
_exptl_crystal_density_method ?
_exptl_absorpt_correction_type none
_exptl_absorpt_correction_T_min ?
_exptl_absorpt_correction_T_max ?
_diffraction_ambient_temperature 100(3)
_diffraction_radiation_type  electron
_diffraction_radiation_source  electron
_diffraction_radiation_monochromator ?
_diffraction_measurement_device_type JEOL-2100
_diffraction_measurement_method  'Continuous Rotation Electron Diffraction'
_diffraction_detector_area_resol_mean 18.1818
_diffraction_standards_number 0
_diffraction_standards_interval_count ?
_diffraction_standards_decay_% ?
_computing_data_collection  'Instamatic (Wang et al., 2019)'
_computing_cell_refinement  'XDS (Kabsch et al., 2010)'
_computing_data_reduction  'XDS (Kabsch et al., 2010)'
_computing_structure_solution 'SHELXS-97 (Sheldrick 2008)'
_computing_structure_refinement 'SHELXL-2014 (Sheldrick 2014)'
_computing_molecular_graphics 'Bruker SHELXTL'
_computing_publication_material 'Bruker SHELXTL'
_chemical_name_common        SU-100_cryo
_diffraction_ambient_environment vacuum
_diffraction_radiation_probe  electron
```

```

_diffrn_radiation_wavelength      0.0251
_diffrn_source                    'electron microscope'
_diffrn_source_voltage            200
_diffrn_detector                  TimePix-SU
_diffrn_measurement_device        TEM
_diffrn_measurement_details
;

```

List of Samples for Merging (Abbr. = Camera Length, Resolution):

#	Type	Start	End	Step	t~exp~	#Frames	CL	Reso.	ISa	CC1/2
1	/a	-16.82	40.6374	0.2326	0.5	248	439.48	0.62	3.01	98.2
2	/a	-51.78	57.6676	0.2324	0.5	472	439.48	0.55	3.25	98.4
3	/a	-56.68	35.786	0.2323	0.5	399	439.48	0.68	2.31	98.0
4	/a	-50.47	19.5175	0.2325	0.5	302	439.48	0.55	4.23	98.7
5	/a	-57.04	23.8077	0.2323	0.5	349	439.48	0.54	3.65	98.2
6	/a	-48.75	61.5573	0.2327	0.5	475	439.48	0.64	2.15	97.4
7	/a	-54.26	26.948	0.232	0.5	351	439.48	0.7	2.53	96.0
8	/a	-53.95	54.1775	0.2325	0.5	466	439.48	0.68	2.07	96.1
9	/a	-51.83	52.5575	0.2325	0.5	450	439.48	0.55	2.88	98.5
10	/a	-49.31	53.5672	0.2328	0.5	443	439.48	0.55	2.26	98.6
11	/a	-39.3	47.4574	0.2326	0.5	374	439.48	0.7	3.08	99.4

```
;
```

7. Tutorial Example 4: Lysozyme

7.1 General information

- **Raw Data:** The Lysozyme datasets can be downloaded from <https://doi.org/10.5281/zenodo.13835773>. Please note that the dataset is huge in size, it is recommended to use part of datasets (like 30) to go through the whole tutorial.
- **Data Collection:** 71 datasets were collected using an Ceta-D detector installed on a Titan Krios G2 (300 kV) microscope equipped with the X-FEG. Data was collected at cryo temperature with an autoloader.
- **Collection Method:** Data collection utilised EPU-D.
- **References:**
 - EPU-D details: [<https://www.thermofisher.com/se/en/home/electron-microscopy/products/software-em-3d-vis/epu-software.html>]
 - Sample and data collection: [AutoLEI paper]
 - Final structure deposited in PDB
- **Sample Details:** Lysozyme, space group: $P\bar{1}$, unit cell dimensions: 26.39 Å, 31.01 Å, 33.26 Å, 88.13°, 108.56°, 112.28°
- **Software Versions:** Tutorial conducted using AutoLEI version 1.0 and XDS version 2025/04/30.

7.2 Goals

1. Understand the workflow of AutoLEI.
2. Convert image format.
3. Generate XDS.INPs.
4. Perform batch processing of electron diffraction data using generated XDS.INPs.
5. Finalise data processing by merging dataset.

7.3 Load Working Directory

1. Start AutoLEI:
 - If using a pip-installed version, activate the Python environment in Linux and type **autolei**.
 - Alternatively, navigate to the folder containing **autolei.py** and run it.
 - After version and scaling information appear in the command window, the main AutoLEI window will open.
 - **Note:** If a message appears stating “XDS not existing! Do you wish to continue?”, ensure XDS is accessible to AutoLEI (e.g., added to the system path).
2. Select Input Path:
 - Click **Browse** to select the folder containing the 71 datasets.

- Click **Load**. Once loaded, the title updates to display the selected path and the number of datasets. Input parameters will also reflect communal settings.
- 3. Select Instrument Profile:
 - Use the drop-down menu on the top-right corner to select **Ceta-D-2k-SciLifeLab** as the instrument profile.
 - Click **Load** to load the instrument profile into the current window.
 - Most input fields are fulfilled. In the additional information part, **SIGNAL_PIXEL=7** is the specify parameter for Ceta-D detector.
- 4. Input Parameters:
 - The measurement is performed under same conditions, thus, the measurement condition is
 - Enter **1166** in DETECTOR_DISTANCE box.
 - Enter **0.50** in OSCILLATION_RANGE box.
- 5. Save the parameters through clicking **Save Parameters** button. A window will appear to confirm.

AutoLEI 1.0.0 - /mnt/c/AutoLEI_demo/autolei_lysozyme_71 - 0 datasets

Input XDSRunner UnitCellCorr XDSRefine MergeData Cluster_Output Tools RealTime

Browse and load the work path where the program will load the measurement settings.
For XDS input generation, supply basic parameters and click the 'Save Parameter' button.

Input path: /mnt/c/AutoLEI_demo/autolei_lysozyn Browse Load Path Instrument File: Ceta-D-2k-SciLifeLab Load

I. Instrument Parameters

1. Detector parameters: NX= 2048 NY= 2048 QX= 0.028 QY= 0.028

2. Overloading: OVERLOAD= 65000 3. Wavelength: WAVELENGTH= 0.019687 Å

4. Rotation axis: ROTATION_AXIS= 0.9945 -0.1045 0 >>> Use space to segment OR Angle in Degree = 6.0

5. Additional information (Please copy from XDS): SIGNAL_PIXEL=7

II. Measurement Parameters

6. Direct beam position: ORGX= 1024 ORGY= 1024

7. Resolution range: INCLUDE_RESOLUTION_RANGE= 30 0.8 >>> Use space to segment

8. Camera Length: DETECTOR_DISTANCE= 1166 9. Rotation step: OSCILLATION_RANGE= 0.5

Save Parameters

7.4 Batch Processing Datasets in P1 Mode

1. Convert Image Format:
 - Select **MRC** radio button, while MRC to IMG button will be appeared on the right.
 - Click **MRC to IMG** button, the conversion will be performed. The output on the command window and the animation in the main window will indicate the processing.
 - Check on “**Beam Stop Used**” as beam stop is used in the experiment.

2. Generate XDS.INP:
 - Click **Generate XDS.INP** to generate the input file for XDS. XDS.INP will be generated in all 71 data folders in the subfolder “xds”. (You can monitor the updates in the bash window.)
3. Beam Stop Finding:
 - As an arbitrary beam centre position is provided in the input section, a more precise estimation of the beam centre location for each dataset is needed. Besides, the beam stop area is also needed to be marked as untrusted area.
 - **(Important!)** Ensure “**Beam Stop Used**” is checked.
 - Click **Find Beam Stop**, the information on beam centre and beam stop will be written to the XDS.INP and direct_beam_centers.txt under the image folder.
4. Run XDS and View the result
 - Click **Run XDS** to perform batch processing electron diffraction data. The processing will be finished with several minutes. Processing time varies with the operating system and hardware performance (e.g., WSL1 is faster than WSL2).
 - The process completes when the command window shows “All information extracted. Note that resolution is ideal,” or the animation in AutoLEI stops. Monitor progress in the command window’s bottom line.
 - After batch processing, AutoLEI will analysis the processing results automatically and generate xdsrunner.xlsx with all result. Click **Show Results** to show the statistics below the button.
 - The table will contain information for all datasets which successful processed:
 - [A] Serial Number
 - [B] Data Path
 - [C] *P1* Unit Cell
 - [D] Space Group
 - [E] Symmetry-Adopted Unit Cell
 - [F] Index% (from SPOT.XDS)
 - [G] Model ISa (from CORRECT.LP)
 - [H] $CC_{1/2}$
 - [I] Completeness
 - [J] Resolution

AutoLEI 1.0.0 - /mnt/c:/AutoLEI_demo/autolei_lysozyme_71 - 71 datasets

Input **XDSRunner** **UnitCellCorr** **XDSRefine** **MergeData** **Cluster_Output** **Tools** **RealTime**

XDSrunner aims to perform batch data processing with XDS.
Always perform a demo data processing before batch processing.

1. Select Format and Convert to SMV
☒ SMV ☐ MRC ☐ TIFF ☐ NXS ☐ Beam Stop Used

2. Create and Update XDS.INP

3. Process Data under P1 mode and Estimate Symmetry.

4. View Running Result
 >>> xdsrunner.xlsx

No.	Path	Integration Cell	SG	U	Vol.	Index%	ISa	CC1/2	Compl	Reso.
61	.../lyso_20240128_newmade_64/xds	26.592 31.026 33.156 88.512 71.732 68.343	1	2	23935(50)	98.4	8.5	98.18	8.52	1.04
62	.../lyso_20240128_newmade_65/xds	26.617 30.947 33.299 88.276 71.177 68.239	1	2	23869(100)	99.3	7.33	97.66	8.63	1.04
63	.../lyso_20240128_newmade_66/xds	26.884 31.047 33.263 88.421 71.319 67.782	1	2	24122(200)	99.0	9.57	98.16	8.45	1.0
64	.../lyso_20240128_newmade_67/xds	26.465 31.086 33.609 88.13 70.888 67.844	1	2	23923(70)	98.9	10.2	98.94	8.48	0.96
65	.../lyso_20240128_newmade_68/xds	26.349 30.992 33.5 87.958 71.45 67.821	1	2	23729(300)	96.2	13.47	96.4	7.92	1.23
66	.../lyso_20240128_newmade_69/xds	26.324 31.035 33.224 87.743 71.647 67.697	1	2	23725(200)	96.9	8.35	93.83	8.34	1.35
67	.../lyso_20240128_newmade_70/xds	26.323 31.033 33.324 87.815 71.489 67.714	1	2	23686(100)	99.1	8.47	97.57	8.44	1.1
68	.../lyso_20240128_newmade_71/xds	26.374 31.063 33.376 87.68 71.309 67.623	1	2	23663(100)	99.1	6.94	97.76	8.53	1.07
69	.../lyso_20240128_newmade_72/xds	26.402 31.089 33.302 88.455 71.491 67.634	1	2	23813(50)	93.7	11.35	98.65	8.24	1.28
70	.../lyso_20240128_newmade_73/xds	26.629 30.938 33.229 88.271 71.317 68.377	1	2	24000(90)	86.7	16.08	93.46	8.38	1.18
71	.../lyso_20240128_newmade_75/xds	26.245 31.0 33.448 87.62 70.729 67.961	1	2	23492(100)	98.4	10.59	99.19	8.33	1.04

5. Estimate Symmetry:

- This step is unnecessary for protein, as the rough unit cell / space group is often known. However, it will be recommended to run the estimate symmetry to track the subtle change on the unit cell, given a more accurate starting point for the following step.
- Click **Estimate Symmetry & Cell-Cluster**. Results are displayed in the command window and saved as lattice_cluster.txt in the working directory.
- The cell cluster report provides detailed analysis on:
 - Potential Symmetry: Indicates to which symmetry the datasets align.
 - Cluster Analysis: All datasets are within same cluster.
- The primary unit cell is identified as having an aP lattice, and Laue group $\bar{1}$. For further processing, use the refined unit cell parameters: 26.528, 30.997, 33.309, (in Å) and angles 88.297, 71.341, 68.007. Adopt space group 1 (the only available space group for protein in $\bar{1}$).

```

/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_47/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_48/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_49/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_50/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_51/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_52/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_53/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_54/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_56/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_57/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_58/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_59/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_60/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_61/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_62/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_63/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_64/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_65/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_66/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_67/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_68/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_69/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_70/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_71/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_72/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_73/xds
/mnt/c/AutoLEI_demo/autolei_lysozyme_71/lyso_20240128_newmade_75/xds

** Lattice Choice:
-- Bravais Lattice: aP
Averaged Cell: (26.528, 30.997, 33.309, 88.297, 71.341, 68.007), Lattice Difference: 0.0, FOM: 0.0
SEM of Cell: (0.042, 0.025, 0.034, 0.05, 0.075, 0.065)
Suggested Space Group: P1 (No. 1) : 1.0(R), 1.0(C)

```

7.5 Batch Processing Datasets in Cell Mode

Shift to CellCorr Tab to continue.

1. Supply Unit Cell and Space Group information:

- Enter the space group (1) and unit cell dimensions (26.528, 30.997, 33.309, 88.297, 71.341, 68.007) in the respective fields.
- Click **Update Cell Parameters**, a confirmation message will appear.

2. Run XDS and View the result:

- Click **Run XDS with Cell** to perform batch processing electron diffraction data. The processing will be finished with several minutes.
- After batch processing, AutoLEI will analysis the processing results automatically and generate xdsrunner2.xlsx with all result. Click **Show Results** to show the statistics below the button.
- The table will contain information for all datasets which successful processed:
 [A] Serial number, [B] Path of the data, [C] Space group, [D] Symmetry-adopted Unit cell, [E] Index% (calculated from SPOT.XDS), [F] Model ISa from XDS (extracted from CORRECT.LP), [G] R_{meas} , [H] $CC_{1/2}$, [I] Completeness, [J] Resolution.

AutoELI 1.0.0 - /mnt/c/AutoELI_demo/autolei_lyscyme_71 - 71 datasets

InputXDSRunnerUnitCellCorrXDSRefineMergeDataCluster_OutputToolsRealTime

Input Space group and unit cell parameters.

Providing unit cell and space group keywords for all datasets is helpful for later data merging. XDS will refine unit cells individually.

Blind unit cell searching should be performed in XDSrunner. Check results in xdsrunner.xlsx / estimate symmetry

Space group: Unit cell:

Update Cell Parameters

* Run XDS with updated .inp files.

Run XDS with CellStop Run

* Show running result

Show ResultsUpdate Results FileOpen Result File>>> xdsrunner2.xlsx

No.	Path	Sc	Unit cell					Vol.	Index%	ISa	Rmeas	CC1/2	Comp	Reso.	
60	.../lyso_20240128_newmade_63/xds	1	26.53(6)	31.019(8)	33.20(1)	88.57(2)	71.32(5)	68	23936(70)	99.0	11.7	24.25	98.56	8.53	0.96
61	.../lyso_20240128_newmade_64/xds	1	26.55(5)	31.024(2)	33.13(2)	88.50(7)	71.57(5)	68	23935(50)	98.4	8.5	25.38	98.26	8.52	1.04
62	.../lyso_20240128_newmade_65/xds	1	26.5(1)	30.93(4)	33.26(2)	88.4(1)	71.3(1)	68.3(2)	23869(100)	99.3	7.33	24.42	97.64	8.63	1.04
63	.../lyso_20240128_newmade_66/xds	1	26.8(1)	31.0(1)	33.26(1)	88.42(8)	71.3(1)	67.9(4)	24122(200)	99.0	9.57	26.72	98.19	8.45	1.0
64	.../lyso_20240128_newmade_67/xds	1	26.43(4)	31.03(2)	33.46(6)	88.5(1)	71.09(7)	67.9	23923(70)	98.9	10.2	20.74	98.93	8.48	0.96
65	.../lyso_20240128_newmade_68/xds	1	26.3(2)	30.98(1)	33.3(1)	88.10(2)	71.8(4)	67.87(9)	23729(300)	96.2	13.47	35.96	96.38	7.92	1.23
66	.../lyso_20240128_newmade_69/xds	1	26.3(2)	31.030(3)	33.3(1)	87.8(1)	71.8(4)	67.72(3)	23725(200)	96.9	8.35	35.17	93.9	8.34	1.35
67	.../lyso_20240128_newmade_70/xds	1	26.27(6)	31.03(2)	33.2(1)	87.89(9)	71.7(2)	67.7(1)	23686(100)	99.1	8.47	33.94	97.46	8.44	1.1
68	.../lyso_20240128_newmade_71/xds	1	26.23(6)	31.05(1)	33.2(1)	87.81(5)	71.7(2)	67.65	23663(100)	99.1	6.94	31.54	97.71	8.53	1.07
69	.../lyso_20240128_newmade_72/xds	1	26.39(1)	31.10(5)	33.26(1)	88.5(1)	71.57(8)	67.6	23813(50)	93.7	11.35	26.46	98.66	8.24	1.28
70	.../lyso_20240128_newmade_73/xds	1	26.64(8)	30.96(2)	33.21(3)	88.3(1)	71.4(2)	68.3(1)	24000(90)	86.7	16.08	49.04	93.48	8.38	1.18
71	.../lyso_20240128_newmade_75/xds	1	26.229(8)	31.001(7)	33.3(1)	87.60(7)	70.4(1)	67.4	23492(100)	98.4	10.59	15.76	99.19	8.33	1.04

7.6 Merge Data

For data merging, only datasets of excellent quality should be selected. While there is not a direct indicator to describe the data quality for electron diffraction data, ISa (Metric of Model Fitting), Resolution, R_{meas} and $CC_{1/2}$ can be used for comparison. If another standard is used to filter, **Manually Filter** can be used.

1. Filter Data Using ISa:

- In DataMerge tab, select **ISa** from the drop-down menu. Use the default value (5) and click **Filter Data**. 65 of 71 datasets will be selected for further merging.

2. Merge Data:

- Click **Merge Data**. The selected datasets from xdspicker will be merged.
- A notification may appear stating: “4 volumes deviate by more than 3 sigma. The mean volume is $23,806.5 \pm 232.0$ ($23,110.6 - 24,502.4$). The maximum deviation is 7.98 sigma. Do you wish to continue?” This warning indicates that certain cell volumes deviate significantly from the majority. Open the **Manual Filter** to identify the outlier datasets. Upon inspection, it becomes evident that dataset no. 5 (newmade_7) has a larger unit cell, while datasets 31 (newmade_34/xds) and 32 (newmade_35/xds) exhibit smaller unit cells. Although these deviations are substantial, it is still possible that the datasets are usable. To proceed, **close** the Excel or LibreOffice window.
- Click **Show Result** to check the statistical table of the data merging process.
- After this step, a merge subfolder will be generated with the all.HKL file.

AutoLEI 1.0.0 - /mnt/c/AutoLEI_demo/autolei_lysozyme_T1 - 71 datasets

Input **XDSRunner** **UnitCellCorr** **XDSRefine** **MergeData** **Cluster_Output** **Tools** **RealTime**

Generate and Merge data from xdspicker.xlsx.

Note: Average unit cell parameters will be used during merging. Generate .hkl and .P4P files for SHELX.

I. Filter data for merging

Use the data with better than for merging.

II. Merge Data

* Strongly recommend to cluster before

SUBSET OF RESOLUTION LIMIT	INTENSITY OBSERVED	DATA WITH REFLECTIONS UNIQUE	SIGNAL/NOISE POSSIBLE	COMPLETENESS OF DATA	R-FACTOR observed	FUNCTION OF RESOLUTION expected	R-FACTOR COMPARED	I/SIGMA	R-meas	CC(1/2)	Anomal Corr	SigAno	Nano
4.29	4886	562	637	88.2%	15.5%	17.5%	4866	13.39	16.4%	98.2*	-14	0.731	478
3.04	10552	1124	1147	98.0%	15.6%	18.5%	10543	12.86	16.6%	98.4*	-21	0.671	1011
2.48	13515	1465	1501	97.6%	27.9%	27.8%	13507	9.75	29.7%	91.0*	-18	0.721	1333
2.15	16506	1736	1775	97.8%	38.5%	36.8%	16588	8.25	40.0%	94.5*	-16	0.718	1693
1.92	18762	1955	1982	98.6%	42.6%	41.0%	18750	7.32	45.1%	88.3*	-12	0.732	1812
1.75	21659	2161	2280	98.2%	62.9%	59.8%	21646	5.97	66.6%	83.7*	-9	0.752	1998
1.62	24169	2400	2424	99.0%	85.0%	81.3%	24150	4.97	89.9%	69.4*	-9	0.758	2225
1.52	23258	2523	2587	97.5%	98.0%	93.4%	23226	4.14	104.2%	49.9*	-10	0.738	2243
1.43	22909	2690	2777	96.9%	94.4%	91.0%	22868	3.56	100.6%	57.7*	-6	0.746	2311
1.36	21772	2767	2870	96.4%	101.0%	99.0%	21727	3.01	108.3%	48.6*	-5	0.751	2319
1.29	18691	2853	3077	92.7%	96.6%	94.2%	18601	2.70	105.2%	46.6*	-6	0.731	2190
1.24	18538	2882	3180	90.6%	115.3%	111.8%	18437	2.32	126.0%	40.2*	-6	0.715	2170
1.19	16877	2873	3323	86.5%	119.0%	116.3%	16767	2.14	130.9%	41.3*	-6	0.717	2042
1.15	15134	2930	3449	85.0%	123.0%	120.4%	14984	1.87	137.1%	37.6*	-4	0.725	1890
1.11	14772	3006	3624	82.9%	139.4%	136.3%	14595	1.69	156.7%	33.5*	-3	0.687	1891
1.07	12223	2824	3704	76.2%	154.8%	148.7%	12025	1.45	176.5%	28.6*	1	0.697	1384
1.04	10244	2790	3857	72.3%	158.0%	155.0%	10008	1.20	184.4%	26.5*	-6	0.673	1153
1.01	5277	1589	3921	40.5%	184.9%	180.4%	5098	1.09	217.6%	21.5*	-6	0.643	551
0.98	3494	1295	4112	31.5%	180.1%	176.6%	3321	0.93	221.4%	21.8*	3	0.650	290
0.96	1599	825	4186	19.7%	245.5%	244.2%	1443	0.58	337.6%	13.6	24	0.651	37
total	294927	43250	56333	76.8%	36.5%	36.8%	293150	3.70	39.0%	98.1*	-8	0.723	30931