Prepare Experiment

Contents

[1.1 List of terms 1](#_Toc320706117)

[1.1.1 Samples 1](#_Toc320706118)

[1.1.2 Data classification hierarchy 2](#_Toc320706119)

[1.1.3 Data acquisition 2](#_Toc320706120)

[1.1.4 Data processing 2](#_Toc320706121)

[1.2 Definitions 3](#_Toc320706122)

[1.2.1 Macromolecule 3](#_Toc320706123)

[1.2.2 Sample 3](#_Toc320706124)

[1.2.3 Buffer 4](#_Toc320706125)

[1.2.4 Complex 4](#_Toc320706126)

[1.2.5 Additive 5](#_Toc320706127)

[1.2.6 Session 5](#_Toc320706128)

[1.2.7 Stock solutions 5](#_Toc320706129)

[1.2.8 Sample Plate (MX Container) 5](#_Toc320706130)

[1.2.9 Sample Case (MX Dewar) 5](#_Toc320706131)

[1.2.10 Shipment 5](#_Toc320706132)

[1.3 Example 6](#_Toc320706133)

[1.4 Screenshots 7](#_Toc320706134)

[1.4.1 Session List 7](#_Toc320706135)

[1.4.2 Experiment List 7](#_Toc320706136)

[1.4.3 Create / Edit Experiment 8](#_Toc320706137)

[1.4.4 Create / Edit Buffer 9](#_Toc320706138)

[1.4.5 Create / Edit Macromolecule 10](#_Toc320706139)

[1.4.6 Create / Edit Samples 11](#_Toc320706140)

[1.4.7 Create / Edit Complex 12](#_Toc320706141)

## List of terms

### Samples

**Macromolecule** – Biological construct in solution for investigation

**Sample** –specific measurment details for the macromolecule (concentration, HPLC or position in SC) for all alliquots needed for the data collection. *Thus a Macromolecule should have at least 3 samples.*

**Buffer** – The matched solution in which a sample is suspended

**Additive** – Any component of the buffer which will be varied in an experiment (salts, Ligands, detergents lipids, deuteration)

**Complex** – Description (size, composition) of an assembly containing multiple (list of contained) macromolecules (different combinations of individual macromolecules are considered as a separate macromlecules themselves (see example))

**Project** – Group of related Complexes

### Data classification hierarchy

**Frame** – One individual exposure of the detector (refers to both both the 2D detector image .edf and the reduced 1D curve .dat) *(should be linked with the logged Beamline parameters)*

**Measurment / Run** – all frames for an individual acquisition (buffer or individual macromolecule at one concentration) *(should be linked with the average of the logged Beamline parameters)*

**Series** – combination of runs (buffer\_before, macromolecule at one concentration and buffer\_after).

**Data collection** – combined series for one macromolecule (minimum 3 concentrations). *In a sequential data collection, by deffinition all samples are in the same buffer there is only one buffer measurment between samples which is classed as both the buffer\_after in the first series and the buffer\_before in the subsequent series.*

**Experiment** – The group of data collections that are required to answer a biological question.

**Session** – a slot of beamtime allocated to a particular proposal (may contain multiple experiments)

Depending on the complexity of the system for some cases (structure validation for example) an experiment is an individual data collection. However, for others the experiment could be in a number of related parts (see example)

### Data acquisition

**Number of Frames** – The desired number of **frames** to be aquired for each **measurment / run**

**Time per frame** – The duration of each individual **frame** in a **measurment / run**

**Transmission** – The % X-ray transmission to be used for each **measurment / run**

**Energy** – the X-ray energy used for each **measurment / run**

Acquisition parameters should normally be kept the same for all **measurments** in a **data collection** and all **data collections** in an **experiment**. However, as some **samples** my have different radiation sensitivity it is possible that some **macromolecules** have to be treated differently. Furthermore as the deffinition of radiation sensitivity is impirical starting from the first **sample measurment** in the **series**, based on analysis it could be seen that the acquisition parameters (including sample volume) need to be modified to improve data quality / confidence. Thus it is possible (though not ideal) that the data acquisition parameters will be modified for the subsequent **measurment / run**. However, it should be strongly encouraged that the acquisition parameters should be the same for a data collection! As exposure time/intensity are normalized for and variations in energy should be corrected for by a change in the scattering vector (s) scaling, data acquisition variations should not block downstream processing.

### Data processing

**Subtraction** – Background corrected (averaged buffer subtracted) macromolecule measurement.

**Merged curve** – combination of all subtractions from a data collection (corrected for interparticle scattering effects) which corresponds to the scattering from the macromolecule free from concentration effects.

**Structure** – complimentary structural information from other techniques (MX, NMR, EM, etc.) in PDB format for comparison to experimental data

**Shipping** (Note: in brackets are the terms used for MX experiments)

**Stock solutions** – set of samples and buffers (perhaps not diluted) for final preparation onsite in case of additional procedures necessary

**Sample Case (MX Dewar)** – Box with Sample Plates or stock solutions that will be physically sent to the ESRF or brought by the User.

**Sample plate (MX Container)** – Prepared samples (buffers and or samples) in a 96 wellplate (or other deffined option)

**Shipment** –Set of Sample Cases sent to the ESRF (or brought at the ESRF by the User) at the same time with the same courier company.

## Definitions

### Macromolecule

* A name
  + An acronym
* Biological details
  + Molecular mass (kD)
  + Extinction coefficient (M-1 cm-1) (for concentration scaling for UV Spectrophotometer in either HPLC or SC)
  + Number of individual:
    - Protein chains and residues in each
    - DNA strands and bases in each
    - RNA strands and bases in each
  + Sequence
  + List of possible **structures** (as there could be a number of homologues)
    - Definition of which part of the macromolecule the structure is *(to minimize repetition should be linked with complex definitions)*
      * Could be from sequence information if available or
      * Residues (from – to)
      * Bases (from – to)
* Complex formation details
  + Which **Complex**(s) this **macromolecule** belongs to:
  + Definition of which part of the complex the **macromolecule** is (sequence, residues from - to or Bases if DNA/RNA) *(multiple regions must be allowed as it could be the combination of two chains in an assembly)*
  + Protonation/duteration (labeling)

### Sample

* Its macromolecule acronym
* An associated buffer (by definition it is the conditions in which the sample is suspended and is thus the descriptor of its specific conditions and the buffer which should be measured to enable background subtractions) *i.e. if the buffer conditions change for a project each is defined as a new sample (copy paste) with a new buffer (with the correct additives) associated.*
* Concentration (Mass or Protein/DNA/RNA) ( **Non 0** mgml-1)
* Measurement temperature (°C),
  + Set point
  + Measured value for each frame
* Its default name will then be:

**MACROMOLECULE acronym\_BUFFER acronym\_Concentration\_TempC \***

(if there are conflicts users should be prompted to distinguish the individual samples)

* Type of measurement,
  + HPLC
  + sample changer
    - viscosity (high/medium/low)
    - Locations (in large projects multiple wells might be needed to provide the volume necessary for all measurements
      * plate name
      * row and column number
      * volume in well (ml)
* Volume to load (ml)
  + Flow (Yes / No)
* Comments: (text describing the sample)
* As required by the safety , each sample has to be described in a sample sheet in the proposal submission form (User Portal / SMIS system)

\* As a sample could be measured multiple times (for instance to find the best acquisition parameters for radiation damage) the output data (**frames**) will take the sample name and be appended by run number and frame number (or \_ave).

**MACROMOLECULE acronym\_BUFFER acronym\_Concentration\_TempC\_Run\_(Frame or ave)**

Thus following the example, Aase (A) in the base buffer (B1) at 10mg/ml and 4 °C if data was acquired at run 015 with 10 frames would give, 11 frames named sequentially: **A\_B1\_10\_4C\_015\_*x*** (where ***x*** is the incrementing frame number (1 to 10) and the average file (ave)

*Parameters from all Beamline monitoring devices (machine current, beamstop diode, SC temp, etc.) should be logged for each* ***frame*** *individually and the average.*

### Buffer

* A name
  + An acronym
* Concentration (Mass or Protein/DNA/RNA is **BY DEFFINITION = 0** mgml-1) **Non modifiable value**

*Value of 0 mgml-1 needs to be saved with the measurement details for downstream processing and verification.*

* Measurement Temperature
  + Set point **(Non modifiable by USER).**  *For each buffer measurement it is set by the sample thus you only need to define one buffer for a temperature series.*
  + Measured value for each frame
* Type of measurement,
  + HPLC
  + sample changer
    - viscosity (high/medium/low)
    - Locations (in large projects multiple wells might be needed to provide the volume necessary for all measurements
      * plate name
      * row and column number
      * volume in well (ml)
* Volume to load (ml) (by default should match that of the sample)
  + Flow (Yes / No) (by default should match that of the sample)
* Composition
  + Type
    - PBS
    - TRIS
    - HEPES
    - Etc.
  + pH
* List of additives

Note: A buffer has no measurement temperature defined but the temperature it was measured at (defined by the sample it matches) must be recorded for verification. This temperature should be used in the naming for the frames collected for the buffer in the same way for samples. For a buffer the naming convention should be:

**BUFFER acronym\_TempC\_Run\_(Frame or ave)**

Thus all like buffers can be easily grouped and checked for consistency and will always be associated with the correct samples as the run numbers will increment.

*Parameters from all Beamline monitoring devices (machine current, beamstop diode, SC temp, etc.) should be logged for each* ***frame*** *individually and the average.*

### Complex

* A name
  + An acronym
* Biological details
  + Molecular mass (kD)
  + Number of individual:
    - Protein chains and residues in each
    - DNA strands and bases in each
    - RNA strands and bases in each
    - Each chain / strand should be given a unique name to allow it to be used for selection for macromolecules part of the complex
  + Sequence
  + List of associated structures
    - Definition of which part of the complex the structure is *(to minimize repetition should be linked with Macromolecule definitions)* (additionally there could be a number of homologues for the same structure)
      * Could be from sequence information if available or
      * Residues (from – to)
      * Bases (from – to)
* A list of Macromolecules which are part of the complex
* As required by the safety, each complex has to be described by a sample sheet in the proposal submission form (User Portal / SMIS system).

Note: In the simple case of a structure validation the complex is the macromolecule (which has three samples for the individual concentrations) as it will be defined as being the full number of residues for the complex.

### Additive

* A name
* A Quantity (mM or if the additive is deuterium %)

1. A **structure** is defined by

* A PDB file or accession code
* List of what data this structure is associated with
  + Complex
    - Macromolecule

### Session

* Dates
  + start
  + end
* a number of shifts
* a beamline
* a local contact

Note: 1 Beamtime proposal may have multiple **sessions**, each session could have multiple **Experiments**

### Stock solutions

* Will need instructions for storage once they have arrived
  + Room temp (20°C)
  + Fridge (4°C)
  + Freezer (-20°C)
  + Cryo-cooling (-80°C)
  + Comments for special instructions
* Extra instructions for preparation of dilutions ( to make individual samples / Sample plate)

### Sample Plate (MX Container)

* Storage Temperature (°C) should be defined for each plate to set SC thermostat for correct conditions (as 3 plates can be stored at once if there is a difference in the plates in the SC a warning should be given for the users to decide to either override (set new temp for all plates in SC) or to remove plates from SC)

### Sample Case (MX Dewar)

* List of Stock solutions and Sample plates contained
* Will need instructions for storage once they have arrived
  + Room temp (aprox 20°C)
  + Fridge (4°C)
  + Freezer (-20°C)
  + Cryo-cooling (-80°C)
  + Comments for special instructions

### Shipment

* User (return/Contact) details
* List of Sample cases included in the shipment

## Example

An enzyme which in its functional form is comprised of 3 individual subunits (A, B and C). The first part of the Experiment ( P1) is to determine how the subunits fit together.

P1 will comprises data collections of:

A, B and C individually

The dimeric complexes of AB and BC

The trimeric complex ABC

Thus the **complex** will have 6 individual **macromolecules** associated with it (A, B, C, AB, BC and ABC). Each with a minimum of 3 samples, all with the same buffer conditions.

P2 is to understand how the enzyme functions

P2 will comprises data collections of:

The trimeric complex ABC (same conditions as for P1)

Plus a data colelction for each buffer condition (Ligands and additives required for activity and /or non hydrolysable homologues to isolate the various stages of the reaction).

Thus the macromolecule ABC will have an additional n\*3 samples where n is the number of buffer conditions that need to be measured.

If structures for A, B, and C are available they can be validadted (with CRYSOL) against the macromlecule data or if they are only partial structures the missing residues modeld ( with BUNCH). They can be used to make rigid body models for AB, BC and ABC (with SASREF or BUNCH / CORAL if missing protions need to be added). The Ab-initio models (automatically calculated by EDNA pipeline) can be overlayed with the rigid bidy models for comparison (SUPCOMB20) and a table of the goodness of fit to both the experimental data and the other models can be produced aiding validation.

## Screenshots

### Session List

This screen shows all the sessions that are available for the given proposal. Clicking on the “Edit” button of a session brings you to the Experiment List of this session (see next screen).

Note: This session list will contain all MX and BioSAXS sessions – by selecting a particular session the users interface will change to best accommodate the MX or BioSAXS information displayed. It might be good to have a toggle to display MX sessions or BioSAXS sessions alone.



### Experiment List

This screen shows all the experiments of a given session. You can Delete, Edit an existing experiment or Create a new one (see next screen).

Also should include pdf creator buttons for exporting the list of experiments from a particular session.

Action

New



Experiment List

Experiment

EXP

01

EXP

02

Comments

First experiment

Second experiment



Comments

This is a comment



Session start date

ID

14

-

3

Beamline

02

/

01

/

2010 14

:

00

Daniel Martin

Operator

Edit the

Experiment

Create a new

Experiment in

this session

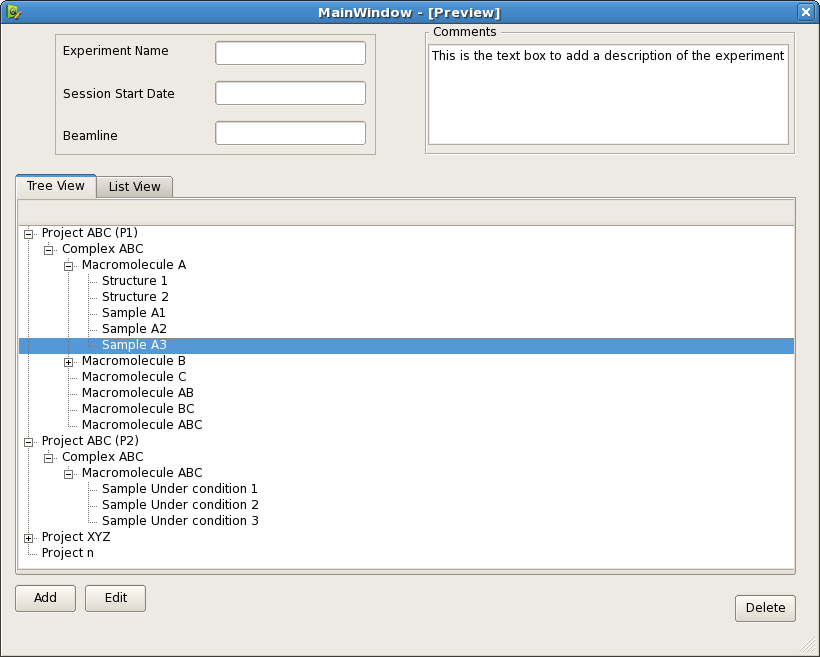
### Create / Edit Experiment

An experiment is defined a number of data collections which in turn consists of a number of measurements of buffers or macromolecule solutions. An experiment needs to hold the information for all required samples, their buffers the macromolecules and in turn the complex the samples belong to.

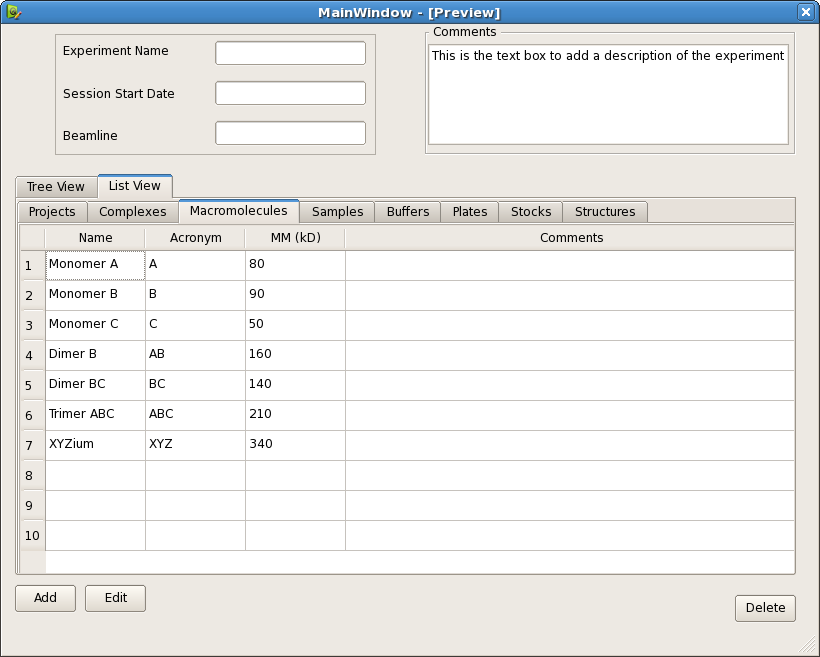
As all of these variable are nestled there should be a view to highlight that (such as a tree) and the lists of all individual samples etc. In the experiment

Pressing Add or edit should take the user to the appropriate dialog depending what (list,or part of the tree) is active.

Note: Extra blank measurements of buffers with excessive additives can be measured to ensure the additive has no effect on the scattering. Thus it is possible for there to be buffers that are not related to any sample.



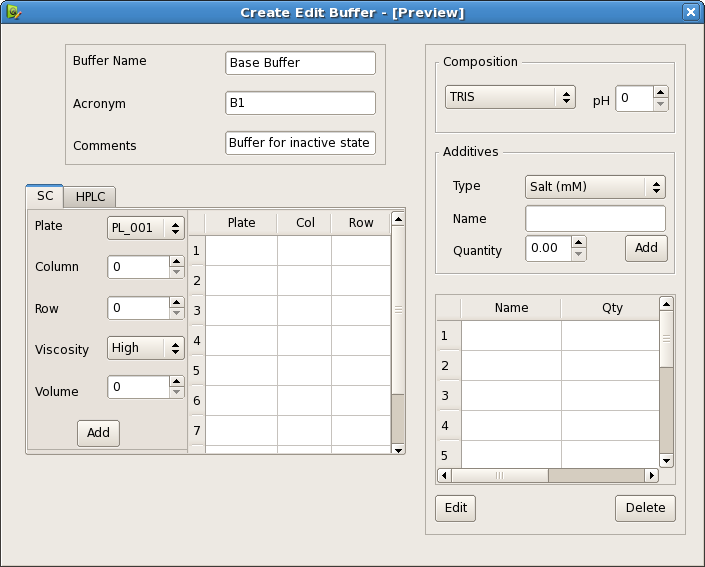
Tree View



List View

### Create / Edit Buffer

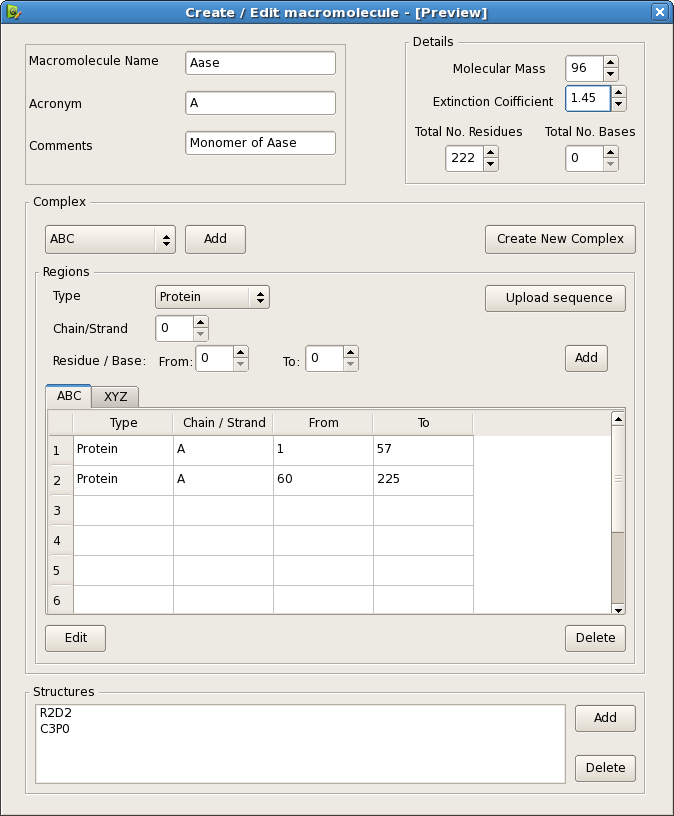
A buffer is now the specific conditions and additives in which a sample will be and is associated with. It also has the positions in which all like buffers can be found.



Multiple defined positions in SC are required

### Create / Edit Macromolecule

The molecule contains all the information regarding the samples context to the complex and biological details. In the MX version we have a pull-down menu to populate the ‘sample’ field – where the pull-down menu is composed of all macromolecules with sample sheets.



### Create / Edit Samples

Sample now contains just the specific information for the specific measurement of a macromolecule under 1 conditions and individual concentration, as the only variable changing are Concentration, temperature and the position of the sample the definition of samples should allow defining multiple samples (different concentrations / temperatures quickly and easily).

As was observed from previous use a visual representation of the SC and sample positions in a plate would be very helpful.

Note: Perhaps a variable in the sample information would be useful for protonation/duteration (labeling) for contrast variation experiments but as it is irrelevant for X-rays it could also be included in the comments

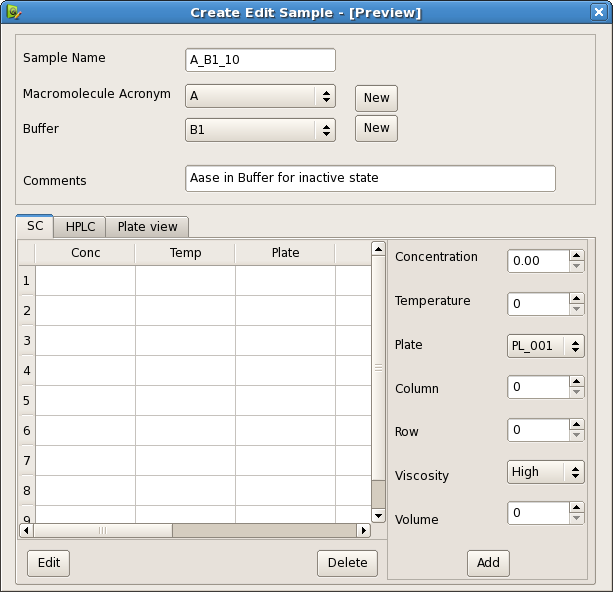


Table View

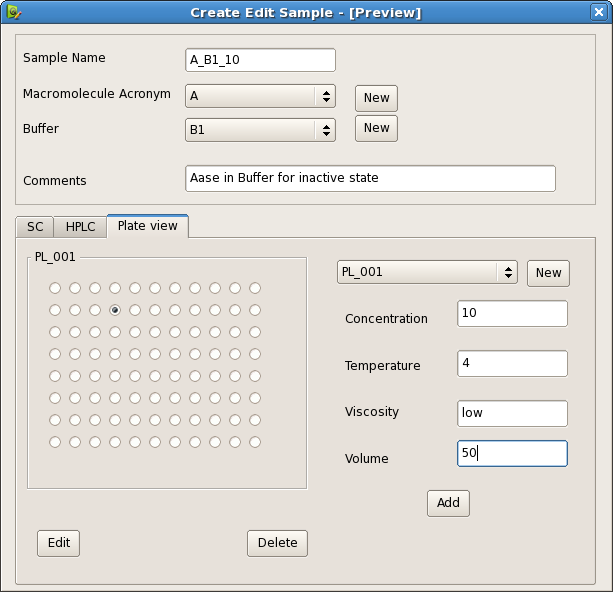


Plate view

### Create / Edit Complex

A complex not only holds the list of all macromolecules it contains but the biological details of the whole assembly\*

\*Specific information on how the macromolecules are related (fit in) to the complex is held in the macromolecule itself

