LHSG-CT-2003-503420

BioXHIT

A European integrated project to provide a highly effective technology platform for Structural Genomics.

Life Sciences, Genomics and Biotechnology for Health

WP5.2: De 5.2.13 Updating of XIA2 to deliver all possible information required for deposition

Due date of deliverable: **30.06.2008**
Actual submission date: **30.06.2008**

Start date of project: **1.1.2004**   **Duration: 60 months**

Organisation name of lead contractor for this deliverable: CCP4/STFC Daresbury **Author:** Graeme Winter

De 5.2.13 Updating of XIA2 to deliver all possible information required for deposition

The key programs used in xia2, from the perspective of deposition, are XDS, Mosflm, Scala & Truncate. Of these only XDS is not capable of producing harvest files though a request has been added to the wiki http://strucbio.biologie.uni-konstanz.de/xdswiki/index.php/Main_Page to allow this, and Scala is used with XSCALE to merge giving much of the necessary information. Although the harvesting information is usually indexed by PROJECT/CRYSTAL/DATASET it has been necessary to append a sweep name to the DATASET for Mosflm, as multiple sweeps may exist for a given data set. To support the use of MAD experiments, all data are reprocessed / remerged to the finally decided resolution limit to ensure consistent results.

The harvest files from xia2 are placed in the following subdirectories of the xia2 working directory:

```
Harvest/
Harvest/DepositFiles
Harvest/DepositFiles/TS01
Harvest/DepositFiles/TS01/NATIVE_NATIVE_LR.mosflm_run_1_90
Harvest/DepositFiles/TS01/INFL.truncate
Harvest/DepositFiles/TS01/LREM.truncate
Harvest/DepositFiles/TS01/NATIVE_NATIVE_HR.mosflm_run_1_180
Harvest/DepositFiles/TS01/LREM.scala
Harvest/DepositFiles/TS01/INFL_INFL.mosflm_run_1_180
Harvest/DepositFiles/TS01/NATIVE.truncate
Harvest/DepositFiles/TS01/NATIVE.scala
Harvest/DepositFiles/TS01/LREM_LREM.mosflm_run_1_180
Harvest/DepositFiles/TS01/INFL.scala
```

The results may then be directly imported into AutoDep. The corresponding reflection files are found in the DataFiles directory:

```
DataFiles/
DataFiles/TS01_12847_unmerged.sca
DataFiles/TS01_13140_free.mtz
DataFiles/TS01_13140_unmerged_INFL.sca
DataFiles/TS01_12847_free.mtz
DataFiles/TS01_13140_unmerged_LREM.sca
```

Both MTZ and unmerged scalepack files are provided to give the widest support for downstream processing. Finally, the log files from optimised processing can be found in the LogFiles directory:

```
LogFiles/
LogFiles/bioxhit.xml
LogFiles/INFL_TS01_13140_INFL_mosflm_integrate.log
LogFiles/LREM_TS01_13140_LREM_mosflm_integrate.log
LogFiles/TS01_13140_INFL_truncate.log
LogFiles/TS01_13140_LREM_truncate.log
LogFiles/TS01_12847_NATIVE_truncate.log
LogFiles/TS01_12847_scala.log
```

```
LogFiles/TS01_13140_scala.log
LogFiles/NATIVE_LR_TS01_12847_NATIVE_mosflm_integrate.log
LogFiles/NATIVE_HR_TS01_12847_NATIVE_mosflm_integrate.log
```

The version of xia2 supporting this functionality will be included in the final release of CCP4 6.1 and is available from http://www.ccp4.ac.uk/xia.

Graeme Winter