

Standard Operating Procedure (SOP)

Acquisition, backup and archiving of HPLC-MS datasets

Version 2.0; 2026.31.08

Christof Regl, Aleksandar Paunovic and Christian G. Huber

University of Salzburg
Department of Biosciences and Medical Biology
Bioanalytical Research Labs

Table of Contents

1. Introduction and Purpose	3
2. Scope and Application	3
3. Instrument names	3
4. Correct naming of files and sequences	3
4.1 Chromeleon based instruments	3
4.2 Xcalibur	4
4.3 Correct naming of sequences and files	4
5. Exporting the files	4
5.1 Chromeleon based instruments (Q Exactive and Q Exactive Plus)	4
5.2 Xcalibur based instruments (Orbitrap Ascend Tribrid Mass Spectrometer)	7
6. Archiving of files	7
6.1 Archiving and registry to SQL database	7
6.2 Retrieving data from the archive	8
7. APPENDIX	8

List of Figures:

Figure 1: Export of raw data from Chromeleon step 1.	5
Figure 2: Export of raw data from Chromeleon step 2.	5
Figure 3: Export of raw data from Chromeleon step 3	6
Figure 4: Export of raw data from Chromeleon step 4	7
Figure 5: Searching files in the SQL database.	8

1. Introduction and Purpose

This SOP describes how datasets should be labeled during HPLC-MS data acquisition and downstream. Furthermore, the SOP describes how the HPLC-MS data are to be backed up and archived, including metadata and results.

2. Scope and Application

The SOP was designed to ensure traceable data acquisition and storage in the Bioanalytical Research Labs within the Department of Biosciences and Medical Biology of the University of Salzburg. It should be applicable for all data sets acquired in the Lab.

3. Instrument names

Each device in the Bioanalytical Research Labs has a specific designation. Mass spectrometers are designated with a capital “M” and a serial number (e.g. M10 for the Thermo Fisher Scientific Orbitrap Ascend Tribrid Mass Spectrometer), and HPLC instruments are designated with a capital H and a serial number (e.g. H16 for the Thermo Fisher Scientific Vanquish Neo).

4. Correct naming of files and sequences

Depending on the instrument, data are acquired in two different ways:

4.1 Chromeleon based instruments

If samples are analyzed by HPLC only or on one of the hybrid quadrupole-Orbitrap mass spectrometers (Q Exactive and Q Exactive Plus), then data acquisition is within the Chromeleon software. Here the first step is to generate a folder with the **users name** in the Chromeleon Data Vault of the PC (e.g. BIO24216) assigned to the respective HPLC (-MS) instrument (e.g. M8): “**BIO24216\DataVaultE_M8\Christof**”. Within the folder, the data are ordered in sequences, which have to be named as described in 4.3.

4.2 Xcalibur

If samples are analyzed on the Orbitrap Ascend Tribrid Mass Spectrometer, then data acquisition is within the Xcalibur software. This works without dedicated Data Vaults but within Windows. Again, each user has to generate a dedicated folder assigned to him on the D hard drive, e.g. “D:\Data\Christof”. Within the folder, the data are ordered in sequences, which have to be named as described in 4.3.

4.3 Correct naming of sequences and files

The sequences have to be named easily understandable, containing the **date**, whereby the order is **YYYYMMDD**, the **users abbreviation**, the used **instruments**, the **column** and the **samples** and who delivered them. All separated by “_”. E.g.: “20250801_CR_M8H14_Aurora258_PMU_MigraineEVs”.

Injections within the sequence are named accordingly to the **samples** by another extension of “_” and ideally the **sequential number**. E.g.: “20250801_CR_M8H14_Aurora258_PMU_MigraineEVs_Control_001”.

5. Exporting the files

Based on the used instrument there are two different data exporting protocols. In both cases, after exporting the raw data from the instruments, also all associated metadata and results have to be stored in the same folders or subfolders as the raw data. This includes e.g. Word documents with the used chemicals and methods, Excel files with protein lists or abundances, R Scripts and other metafiles generated during data evaluation (e.g. Proteome Discoverer, Maxquant, FragPipe metafiles). However, the user needs to ensure a short folder structure (not too many subfolders with long names), as the subsequent archiving which will be described in 6.0, does not allow more than **256 characters**. For example the following folder would have 53 characters “Y:\AG backups\regl\data\2025\PMU\MigraineEVs\rawdata\”, in addition with the 55 characters used for the file name “20250801_CR_M8H14_Aurora258_PMU_MigraineEVs_Control_001” already 108 characters are used.

5.1 Chromeleon based instruments (Q Exactive and Q Exactive Plus)

The generated raw data are saved within the specified Data Vault and have to be exported by the user: Within the sequence all runs need to be marked, followed by “right-click” and “Export...” as shown in Figure 1.

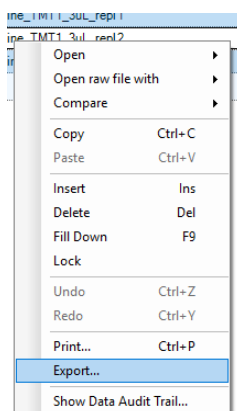


Figure 1: Export of raw data from Chromeleon step 1.

Next the Export Formats & Filenames need to be checked as shown in Figure 2. Raw file format (*.raw)” exports the mass spectrometry data, the “Chromeleon data file format (*.cmbx) contains all HPLC relevant parameters including pressure curves, temperature, method file, etc. The files should **not** be exported to the local PC but directly into the users personal folder in **GFS_BIOMEDBIO_AGHUBER2/AG Backups 2022** to ensure back up.

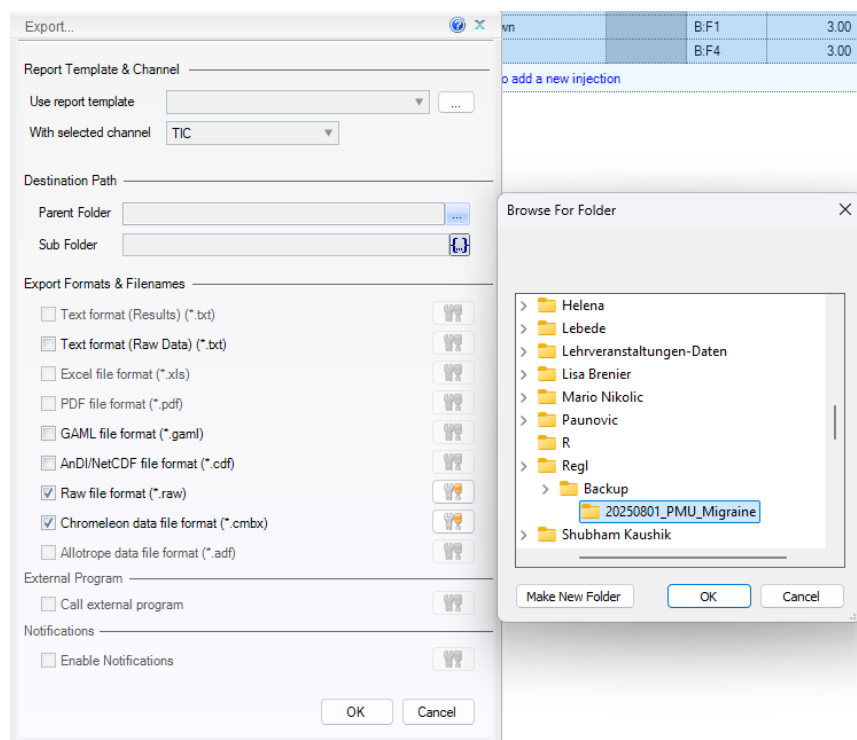
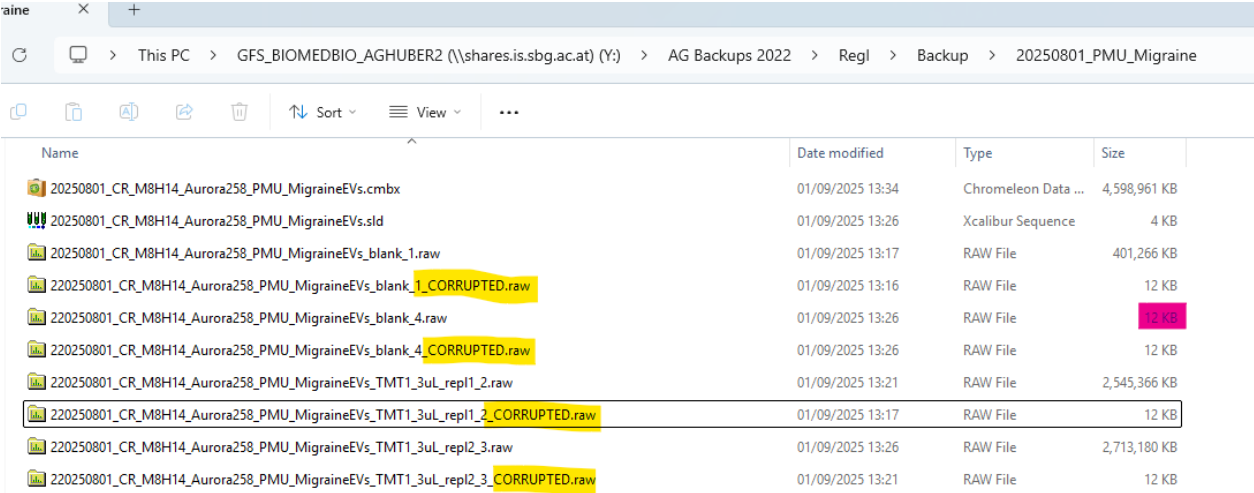


Figure 2: Export of raw data from Chromeleon step 2.

After this, the user has to check if the data export was complete. The folder should contain one .cmbx file and a .raw file for each injected sample (.raw is a Thermo Fisher Scientific specific data format). E.g. 20250801_CR_M8H14_Aurora258_PMU_MigraineEVs_Control_001.raw. Sometimes .raw data are not **completely exported**. Unfortunately, this is not obvious, but only by the fact that the .raw file is not larger than **12 KB**. In this case the sequence has to be reloaded by pressing “F5” and the affected files need to be exported again. Furthermore during the export, an additional .raw file with the extension “**_CORRUPTED**” is generated, these files can be deleted by the user (Figure 3).

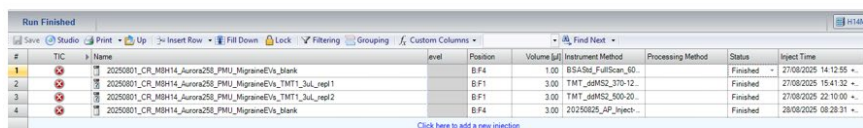


Name	Date modified	Type	Size
20250801_CR_M8H14_Aurora258_PMU_MigraineEVs.cmbx	01/09/2025 13:34	Chromeleon Data ...	4,598,961 KB
20250801_CR_M8H14_Aurora258_PMU_MigraineEVs.sld	01/09/2025 13:26	Xcalibur Sequence	4 KB
20250801_CR_M8H14_Aurora258_PMU_MigraineEVs_blank_1.raw	01/09/2025 13:17	RAW File	401,266 KB
220250801_CR_M8H14_Aurora258_PMU_MigraineEVs_blank_1_CORRUPTED.raw	01/09/2025 13:16	RAW File	12 KB
220250801_CR_M8H14_Aurora258_PMU_MigraineEVs_blank_4.raw	01/09/2025 13:26	RAW File	12 KB
220250801_CR_M8H14_Aurora258_PMU_MigraineEVs_blank_4_CORRUPTED.raw	01/09/2025 13:26	RAW File	12 KB
220250801_CR_M8H14_Aurora258_PMU_MigraineEVs_TMT1_3uL_rep11_2.raw	01/09/2025 13:21	RAW File	2,545,366 KB
220250801_CR_M8H14_Aurora258_PMU_MigraineEVs_TMT1_3uL_rep11_2_CORRUPTED.raw	01/09/2025 13:17	RAW File	12 KB
220250801_CR_M8H14_Aurora258_PMU_MigraineEVs_TMT1_3uL_rep12_3.raw	01/09/2025 13:26	RAW File	2,713,180 KB
220250801_CR_M8H14_Aurora258_PMU_MigraineEVs_TMT1_3uL_rep12_3_CORRUPTED.raw	01/09/2025 13:21	RAW File	12 KB

Figure 3: Export of raw data from Chromeleon step 3

After ensuring correct export of all data, the user has to remove the data from the instrument PC to keep storage space on the instrument PC available. This is done by setting the Status from “Finished” to “Idle” and click save, which will remove the data from the instrument PC. Last the Status needs to be set to “Finished” again. The “empty” sequence can be left on the instrument PC, as it just takes a few KB of storage space (Figure 4).

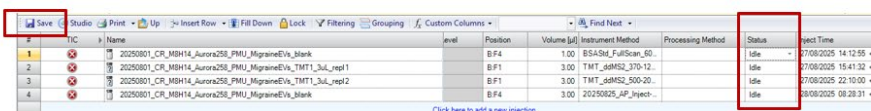
Before



#	TIC	Name	Level	Position	Volume [μl]	Instrument Method	Processing Method	Status	Inject Time
1		20250801_CR_MBH14_Aurora250_PMU_MigraineEva_Blank		B.F4	1.00	BSAStd_FullScan_60		Finished	27/08/2025 14:12:55
2		20250801_CR_MBH14_Aurora250_PMU_MigraineEva_TMT1_3ul_rep1		B.F1	3.00	TMT_dMS2_370-12		Finished	27/08/2025 15:41:32
3		20250801_CR_MBH14_Aurora250_PMU_MigraineEva_TMT1_3ul_rep2		B.F1	3.00	TMT_dMS2_370-12		Finished	27/08/2025 22:10:00
4		20250801_CR_MBH14_Aurora250_PMU_MigraineEva_Blank		B.F4	3.00	20250825_AP_Inject...		Finished	28/08/2025 08:28:31

Local Disk (C:) 53.4 GB free of 953 GB

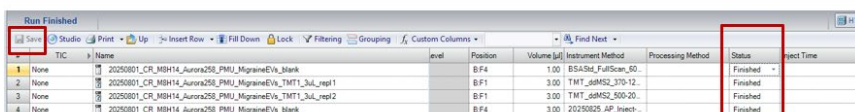
Set to Idle and save



#	TIC	Name	Level	Position	Volume [μl]	Instrument Method	Processing Method	Status	Inject Time
1		20250801_CR_MBH14_Aurora250_PMU_MigraineEva_Blank		B.F4	1.00	BSAStd_FullScan_60		Idle	27/08/2025 14:12:55
2		20250801_CR_MBH14_Aurora250_PMU_MigraineEva_TMT1_3ul_rep1		B.F1	3.00	TMT_dMS2_370-12		Idle	27/08/2025 15:41:32
3		20250801_CR_MBH14_Aurora250_PMU_MigraineEva_TMT1_3ul_rep2		B.F1	3.00	TMT_dMS2_370-12		Idle	27/08/2025 22:10:00
4		20250801_CR_MBH14_Aurora250_PMU_MigraineEva_Blank		B.F4	3.00	20250825_AP_Inject...		Idle	28/08/2025 08:28:31

Chromleon You have reset one or more injections to idle. The raw data for these injections will be deleted if you choose to continue.

After



#	TIC	Name	Level	Position	Volume [μl]	Instrument Method	Processing Method	Status	Inject Time
1		20250801_CR_MBH14_Aurora250_PMU_MigraineEva_Blank		B.F4	1.00	BSAStd_FullScan_60		Finished	27/08/2025 14:12:55
2		20250801_CR_MBH14_Aurora250_PMU_MigraineEva_TMT1_3ul_rep1		B.F1	3.00	TMT_dMS2_370-12		Finished	27/08/2025 15:41:32
3		20250801_CR_MBH14_Aurora250_PMU_MigraineEva_TMT1_3ul_rep2		B.F1	3.00	TMT_dMS2_370-12		Finished	27/08/2025 22:10:00
4		20250801_CR_MBH14_Aurora250_PMU_MigraineEva_Blank		B.F4	3.00	20250825_AP_Inject...		Finished	28/08/2025 08:28:31

Local Disk (C:) 66.5 GB free of 953 GB

Figure 4: Export of raw data from Chromeleon step 4

5.2 Xcalibur based instruments (Orbitrap Ascend Tribrid Mass Spectrometer)

On Xcalibur based instruments the generated .raw data are automatically saved within the specified folder of the sequence, and besides mass spectra also contain HPLC parameters. They do not require further user interaction for an export, and can directly be transferred to the users personal folder in **GFS_BIOMEDBIO_AGHUBER2/AG Backups 2022** to ensure back up.

6. Archiving of files

As also the storage on the **GFS_BIOMEDBIO_AGHUBER2** is limited, the data that is no longer needed on a regular basis (e.g. accepted publication, finished thesis, etc.) has to be archived. For this the data has to be moved by the respective user to the folder "**GFS_BIOMEDBIO_AGHUBER2\AG Backups 2022\ARCHIV\TBA**".

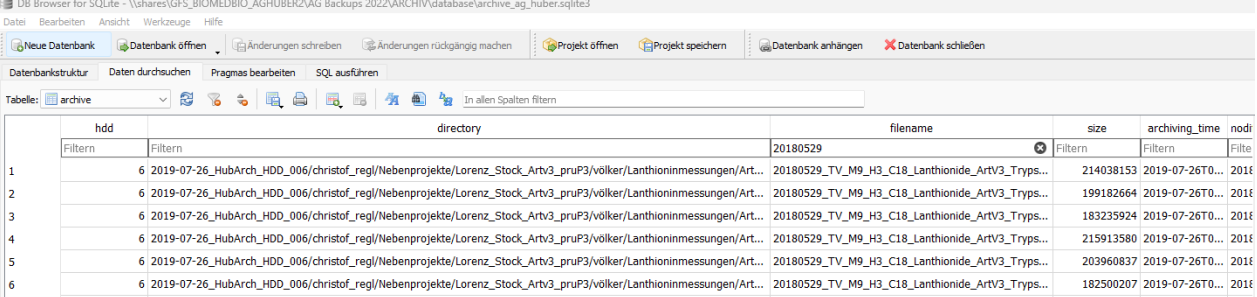
6.1 Archiving and registry to SQL database

The archiving is done as copies on two external hard drives stored in two different offices (EE-062 "Aleks office" and EE-085 "Veronikas office") by Christof Regl (as of 2025) or an instructed person using a dedicated R script ("20250828_CR_Archiving_Script.R"; see in 7.0 Appendix). Using this script the data in the TBA folder are moved into a folder with the current date and

time (e.g. YYYYMMDD_HHMMSS), and may then be moved to the two external hard drives. Additionally all files are registered into the SQL database “archive_ag_huber.sqlite3”.

6.2 Retrieving data from the archive

If already archived data are required, they can be searched in the SQL database using the program “DB Browser for SQLite”, or as of 2025, it is easiest to approach Christof Regl or Aleksandar Paunovic and ask them to retrieve the wanted data. Figure 5 shows an example, how data of a Master student from 2018 are retrieved.



hdd	directory	filename	size	archiving_time	nodf
6	2019-07-26_HubArch_HDD_006/christof_regl/Nebenprojekte/Lorenz_Stock_Artv3_pruP3/völker/Lanthioninmessungen/Art...	20180529_TV_M9_H3_C18_Lanthionide_ArtV3_Tryps...	214038153	2019-07-26T0...	2018
6	2019-07-26_HubArch_HDD_006/christof_regl/Nebenprojekte/Lorenz_Stock_Artv3_pruP3/völker/Lanthioninmessungen/Art...	20180529_TV_M9_H3_C18_Lanthionide_ArtV3_Tryps...	199182664	2019-07-26T0...	2018
6	2019-07-26_HubArch_HDD_006/christof_regl/Nebenprojekte/Lorenz_Stock_Artv3_pruP3/völker/Lanthioninmessungen/Art...	20180529_TV_M9_H3_C18_Lanthionide_ArtV3_Tryps...	183235924	2019-07-26T0...	2018
6	2019-07-26_HubArch_HDD_006/christof_regl/Nebenprojekte/Lorenz_Stock_Artv3_pruP3/völker/Lanthioninmessungen/Art...	20180529_TV_M9_H3_C18_Lanthionide_ArtV3_Tryps...	215913580	2019-07-26T0...	2018
6	2019-07-26_HubArch_HDD_006/christof_regl/Nebenprojekte/Lorenz_Stock_Artv3_pruP3/völker/Lanthioninmessungen/Art...	20180529_TV_M9_H3_C18_Lanthionide_ArtV3_Tryps...	203960837	2019-07-26T0...	2018
6	2019-07-26_HubArch_HDD_006/christof_regl/Nebenprojekte/Lorenz_Stock_Artv3_pruP3/völker/Lanthioninmessungen/Art...	20180529_TV_M9_H3_C18_Lanthionide_ArtV3_Tryps...	182500207	2019-07-26T0...	2018

Figure 5: Searching files in the SQL database.

7. APPENDIX

20250828_CR_Archiving_Script.R:

```
suppressMessages({
  library(tidyverse)
  library(lubridate)
  library(fs)
  library(DBI)
  library(cli)
  library(RSQLite)
  library(bit64)
})
```

```
getwd()
setwd("Y:/AG Backups 2022/ARCHIV")
```

Standard Operating Procedure: Acquisition, backup and archiving of HPLC-MS datasets

```
# =====
# archivR - Archiving system
# =====

cli_h1("archivR - Archiving system for the Labs")
cli_h2("Preparing files for archiving")

# ---- MANUAL HDD NUMBER ----
hdd_number <- 11 # <- change this manually when moving to next HDD
if (is.na(hdd_number)) {
  cli_alert_danger("Invalid HDD number. Aborting.")
  quit(save = "no", status = 1)
}

# ---- Backup database ----
cli_alert_info("Creating backup of database in {.file database/archive_ag__backup.sqlite3}")
file_copy(
  "database/archive_ag_huber.sqlite3",
  "database/archive_ag_huber_backup.sqlite3",
  overwrite = TRUE
)
cli_alert_success("Backup created.")

# ---- Prepare folder ----
archiving_time <- now()
top_folder <- format(archiving_time, "%Y%m%d_%H%M%S")

cli_alert_info("Moving files from {.file TBA} to {.file {top_folder}} ...")
dir_create(top_folder)
file_move(dir_ls("TBA"), top_folder)
cli_alert_success("Files moved.")

# ---- Index files ----
cli_alert_info("Indexing files ..")
files <- dir_info(top_folder, all = TRUE, recurse = TRUE, type = "file")

files_db <- files %>%
  transmute(
    hdd = hdd_number,
    directory = path_dir(path),
    filename = path_file(path),
    size = as.numeric(size) %>% bit64::as.integer64(),
    archiving_time = archiving_time,
    modification_time = format_ISO8601(modification_time),
    birth_time = format_ISO8601(birth_time)
  )

total_size <- sum(files_db$size) %>% as_fs_bytes()
```

Standard Operating Procedure: Acquisition, backup and archiving of HPLC-MS datasets

```
cli_alert_success("Indexed {nrow(files_db)} files (total {total_size}).")

# ---- Write to database ----
db_file <- "database/archive_ag_huber.sqlite3"
cli_alert_info("Appending entries to database {.file {db_file}} ...")
con <- dbConnect(SQLite(), db_file)

if (!("archive" %in% dbListTables(con))) {
  # Create table if it doesn't exist
  dbWriteTable(con, "archive", files_db, overwrite = FALSE, append = FALSE)
  cli_alert_success("Created table 'archive' and inserted first entries.")
} else {
  # Always append
  dbAppendTable(con, "archive", files_db)
  cli_alert_success("Appended entries to table 'archive'.")
}

dbDisconnect(con)

# ---- Log file ----
dir_create("log")
log_file <- str_glue("log/{format(archiving_time, '%Y%m%d_%H%M%S')}.csv")

cli_alert_info("Creating log file {.file {log_file}} ...")
write_lines(str_glue("# Archived files on {archiving_time}"), log_file)
write_csv(files_db, log_file, col_names = TRUE, append = TRUE)
cli_alert_success("Log file created.")

# ---- Final steps ----
cli_h2("Further steps")
cli_ol()
cli_li("Copy the folder {.file {top_folder}} to HDD {hdd_number}.1.")
cli_li("Copy the folder {.file {top_folder}} to HDD {hdd_number}.2.")
cli_li("Delete the folder once verified.")
cli_end()
```