Package 'chromConverter'

October 12, 2022

Title Chromatographic File Converter

Version 0.2.1

Maintainer Ethan Bass <ethanbass@gmail.com>

Description Reads chromatograms from binary formats into R objects. Currently supports conversion of 'Agilent ChemStation', 'Agilent MassHunter', and 'ThermoRaw' files as well as various text-based formats. Utilizes file parsers from external libraries, such as 'Aston' <https://github.com/bovee/aston>, 'Entab' <https://github.com/bovee/entab>, and 'ThermoRawFileParser' <https://github.com/compomics/ThermoRawFileParser>.

License GPL (>= 3)

URL https://github.com/ethanbass/chromConverter, https://ethanbass.github.io/chromConverter/

BugReports https://github.com/ethanbass/chromConverter/issues

Imports purrr, readr, readxl, reticulate, stringr, tidyr, utils, magrittr, xml2

Suggests entab, mzR, testthat (>= 3.0.0)

Config/reticulate list(packages = list(list(package = ``scipy''), list(package=``numpy''), list(package = ``pandas''), list(package=``aston'', pip = TRUE)))

Encoding UTF-8

Language en-US

RoxygenNote 7.2.0

Additional_repositories https://ethanbass.github.io/drat/

Config/testthat/edition 3

NeedsCompilation no

Author Ethan Bass [aut, cre] (<https://orcid.org/0000-0002-6175-6739>)

Repository CRAN

Date/Publication 2022-07-11 08:40:02 UTC

R topics documented:

`attach_metadata` *Attaches metadata to chromatogram*

Description

Attaches metadata to chromatogram

Usage

```
attach_metadata(x, meta, format_in, format_out, format_data, parser = NULL)
```

Arguments

`x`	chromatogram
`meta`	List object containing metadata.
`format_in`	Chromatogram format
`format_out`	R format. Either `matrix` or `data.frame`.
`format_data`	Whether data is in wide or long format.
`parser`	What parser was used to decode the data

Value

A chromatogram with attached metadata.

Author(s)

Ethan Bass

`call_entab` *Entab parsers*

Description

Entab parsers

Usage

```
call_entab(
  file,
  format_data = c("wide", "long"),
  format_in = "",
  format_out = c("matrix", "data.frame"),
  read_metadata = TRUE
)
```

Arguments

`file`	path to file
`format_data`	Whether to output data in wide or long format.
`format_in`	Format of input.
`format_out`	R format. Either `matrix` or `data.frame`.
`read_metadata`	Whether to read metadata from file.

Value

A chromatogram in the format specified by `format_out` (retention time x wavelength).

`call_openchrom` *Parse files with OpenChrom*

Description

Writes `xml` batch-files and calls OpenChrom file parsers using a system call to the command-line interface. To use this function OpenChrom must be manually installed.

Usage

```
call_openchrom(
  files,
  path_out,
  format_in,
  export_format = c("csv", "cdf", "mzml", "animl"),
  return_paths = FALSE
)
```

Arguments

files	files to parse
path_out	directory to export converted files.
format_in	Either msd for mass spectrometry data, csd for flame ionization data, or wsd for DAD/UV data.
export_format	Either csv, cdf, mzml, animl.
return_paths	Logical. If TRUE, the function will return a character vector of paths to the newly created files.

Details

The call_openchrom works by creating an xml batchfile and feeding it to the OpenChrom command-line interface. OpenChrom batchfiles consist of InputEntries (the files you want to convert) and ProcessEntries (what you want to do to the files). The parsers are organized into broad categories by detector-type and output format. The detector-types are msd (mass selective detectors), csd (current selective detectors, such as FID, ECD, NPD), and wsd (wavelength selective detectors, such as DAD, and UV/VIS). Thus, when calling the OpenChrom parsers, you must select one of these three options for the input format (format_in).

Note: Turning on the OpenChrom command-line will deactivate the graphical user interface (GUI). Thus, if you wish to continue using the OpenChrom GUI, it is recommended to create a separate command-line version of OpenChrom to call from R.

Value

If return_paths is TRUE, the function will return a vector of paths to the newly created files. If return_paths is FALSE and export_format is csv, the function will return a list of chromatograms in data.frame format. Otherwise, it will not return anything.

Side effects

Chromatograms will be exported in the format specified by export_format in the folder specified by path_out.

Author(s)

Ethan Bass

---

configure_aston *Configure Aston*

Description

Configures reticulate to use Aston file parsers.

Usage

```
configure_aston()
```

Value

No return value.

Author(s)

Ethan Bass

---

read_chemstation_csv *Chemstation CSV reader*

Description

Chemstation CSV reader

Usage

```
read_chemstation_csv(file, format_out = c("matrix", "data.frame"))
```

Arguments

`file`	path to file
`format_out`	R format. Either `matrix` or `data.frame`.

Value

A chromatogram in the format specified by `format_out` (retention time x wavelength).

Author(s)

Ethan Bass

read_chromeleon *Chromeleon ascii reader*

Description

Chromeleon ascii reader

Usage

```
read_chromeleon(
  file,
  format_out = c("matrix", "data.frame"),
  read_metadata = TRUE
)
```

Arguments

file	path to file
format_out	R format. Either matrix or data.frame.
read_metadata	Whether to read metadata from file.

Value

A chromatogram in the format specified by format_out (retention time x wavelength).

Author(s)

Ethan Bass

read_chroms *Read Chromatograms*

Description

Reads chromatograms from specified folders or vector of paths using file parsers from Aston, Entab, and ThermoRawFileParser.

Usage

```
read_chroms(
  paths,
  find_files,
 format_in = c("chemstation_uv", "chemstation_csv", "masshunter_dad", "shimadzu_fid",
  "shimadzu_dad", "chromeleon_uv", "thermoraw", "mzml", "waters_arw", "msd", "csd",
   "wsd", "other"),
```

```
  pattern = NULL,
  parser = c("", "chromconverter", "aston", "entab", "thermoraw", "openchrom"),
  format_out = c("matrix", "data.frame"),
  export = FALSE,
  path_out = NULL,
  export_format = c("csv", "cdf", "mzml", "animl"),
  read_metadata = TRUE,
  dat = NULL
)
```

Arguments

paths	paths to files or folders containing files
find_files	Logical. Set to TRUE (default) if you are providing the function with a folder or vector of folders containing the files. Otherwise, set toFALSE.
format_in	Format of files to be imported/converted. The current options are: chemstation_uv, chemstation_csv, masshunter_dad, shimadzu_fid, shimadzu_dad, chromeleon_uv, thermoraw, mzml, waters_arw, msd, csd, wsd, or other.
pattern	pattern (e.g. a file extension). Defaults to NULL, in which case file extension will be deduced from format_in.
parser	What parser to use. Current option are chromconverter, aston, entab, thermoraw, or openchrom.
format_out	R object format (i.e. data.frame or matrix).
export	Logical. If TRUE, will export files as csvs.
path_out	Path for exporting files. If path not specified, files will export to current working directory.
export_format	Export format. Currently the only option is .csv, unless you are using OpenChrom parsers, where you could have csv, cdf, mzml, or animl.
read_metadata	Logical, whether to attach metadata (if it's available). Defaults to TRUE.
dat	Existing list of chromatograms to append results. (Defaults to NULL).

Details

Currently recognizes Agilent ChemStation '.uv', MassHunter '.dad' files, and ThermoRaw files. To use Entab and the ThermoRawFileParser, they must be manually installed. Please see the instructions in the Read Me.

Value

A list of chromatograms in matrix or data.frame format, according to the value of format_out.

Side effects

If export is TRUE, chromatograms will be exported in the format specified by export_format in the folder specified by path_out. Currently, the only option for export is csv unless the parser is openchrom.

Author(s)

Ethan Bass

Examples

```
path <- "tests/testthat/testdata/dad1.uv"
chr <- read_chroms(path, find_files = FALSE, format_in = "chemstation_uv")
```

---

read_mzml *Extract UV data from mzML files*

Description

Extracts UV data from mzML files

Usage

```
read_mzml(path, format_out = c("matrix", "data.frame"))
```

Arguments

path	path to file
format_out	R format. Either matrix or data.frame.

Value

A chromatograms in matrix format.

Author(s)

Ethan Bass

---

read_shimadzu *Shimadzu ascii reader*

Description

Shimadzu ascii reader

Usage

```
read_shimadzu(
  file,
  format_in,
  read_metadata = TRUE,
  format_out = c("matrix", "data.frame"),
  what = c("chromatogram", "peak_table", "both")
)
```

Arguments

file	path to file
format_in	Format of files. fid or dad.
read_metadata	Whether to read metadata from file.
format_out	R format. Either matrix or data.frame.
what	Whether to extract chromatogram, peak_table or both.

Value

A chromatogram in the format specified by format_out (retention time x wavelength).

Author(s)

Ethan Bass

---

read_thermoraw *Read ThermoRaw files into R using ThermoRawFileParser*

Description

Converts ThermoRawFiles to mzmL by calling the ThermoRawFileParser from the command-line.

Usage

```
read_thermoraw(
  path_in,
  path_out,
  format_out = c("matrix", "data.frame"),
  read_metadata = TRUE
)
```

Arguments

path_in	path to file
path_out	directory to export mzML files.
format_out	R format. Either matrix or data.frame.
read_metadata	Whether to read metadata from file.

Details

To use this function, the ThermoRawFileParser must be manually installed.

Value

A chromatogram in the format specified by `format_out`.

Side effects

Exports chromatograms in `mzml format` to the folder specified by `path_out`.

Author(s)

Ethan Bass

References

Hulstaert Niels, Jim Shofstahl, Timo Sachsenberg, Mathias Walzer, Harald Barsnes, Lennart Martens, and Yasset Perez-Riverol. "=ThermoRawFileParser: Modular, Scalable, and Cross-Platform RAW File Conversion." *Journal of Proteome Research* **19**, no. 1 (January 3, 2020): 537–42. doi: 10.1021/acs.jproteome.9b00328.

Examples

```
## Not run:
read_thermoraw(path)

## End(Not run)
```

`read_waters_arw` *Waters ascii (.arw) reader*

Description

Waters ascii (.arw) reader

Usage

```
read_waters_arw(
  file,
  read_metadata = TRUE,
  format_out = c("matrix", "data.frame")
)
```

Arguments

`file`	path to file
`read_metadata`	Whether to read metadata from file.
`format_out`	R format. Either `matrix` or `data.frame`.

Value

A chromatogram in the format specified by `format_out` (retention time x wavelength).

Author(s)

Ethan Bass

`sp_converter` *Converter for Agilent MassHunter UV files*

Description

Converts a single chromatogram from MassHunter `.sp` format to R `data.frame`.

Usage

```
sp_converter(
  file,
  format_out = c("matrix", "data.frame"),
  read_metadata = TRUE
)
```

Arguments

`file`	path to file
`format_out`	R format. Either `matrix` or `data.frame`.
`read_metadata`	Logical. Whether to read metadata and attach it to the chromatogram.

Details

Uses the Aston file parser.

Value

A chromatogram in `data.frame` format (retention time x wavelength).

uv_converter *Converter for Agilent ChemStation UV files*

Description

Converts a single chromatogram from ChemStation `.uv` format to R `data.frame`.

Usage

```
uv_converter(
  file,
  format_out = c("matrix", "data.frame"),
  correction = TRUE,
  read_metadata = TRUE
)
```

Arguments

`file`	path to file
`format_out`	R format. Either `matrix` or `data.frame`.
`correction`	Logical. Whether to apply empirical correction. Defaults is TRUE.
`read_metadata`	Logical. Whether to read metadata and attach it to the chromatogram.

Details

Uses the Aston file parser.

Value

A chromatogram in `data.frame` format (retention time x wavelength).

Index