

# Package ‘msdata’

May 14, 2026

**Version** 0.53.0

**Title** Various Mass Spectrometry raw data example files

**Author** Steffen Neumann <sneumann@ipb-halle.de>, Laurent Gatto  
<laurent.gatto@uclouvain.be> with contributions from Johannes Rainer

**Maintainer**

Steffen Neumann <sneumann@ipb-halle.de>, Laurent Gatto <laurent.gatto@uclouvain.be>

**Depends** R (>= 2.10)

**Suggests** xcms, mzR, MSnbase

**ZipData** no

**Description** Ion Trap positive ionization mode data in mzML file format. Subset from 500-850 m/z and 1190-1310 seconds, incl. MS2 and MS3, intensity threshold 100.000. Extracts from FTICR Apex III, m/z 400-450. Subset of UPLC - Bruker micrOTOFq data, both mzML and mz5. LC-MSMS and MRM files from proteomics experiments. PSI mzIdentML example files for various search engines.

**biocViews** ExperimentData, MassSpectrometryData

**License** GPL (>= 2)

**Encoding** UTF-8

**git\_url** <https://git.bioconductor.org/packages/msdata>

**git\_branch** devel

**git\_last\_commit** 56531dd

**git\_last\_commit\_date** 2026-04-28

**Repository** Bioconductor 3.24

**Date/Publication** 2026-05-14

## Contents

CE-MS . . . . .	2
msdata . . . . .	2
proteomics . . . . .	3
sciexdata . . . . .	4
<b>Index</b>	<b>5</b>

---

CE-MS

*CE-MS test data*

---

### Description

The CE-MS test files consist of two files, i.e. "CEMS\_10ppm.mzML" and "CEMS\_25ppm.mzML". The data contains CE-MS runs of a standard mixture that contains e.g. Lysine (at 10 ppm and 25 ppm, respectively) and the neutral EOF marker Paracetamol (50 ppm). The data was acquired on a 7100 capillary electrophoresis system from Agilent Technologies, coupled to an Agilent 6560 IM-QToF-MS. CE Separation was performed using a 80 cm fused silica capillary with an internal diameter of 50  $\mu\text{m}$  and external diameter of 365  $\mu\text{m}$ . The Background Electrolyte was 10 % acetic acid and separation was performed at +30 kV and a constant pressure of 50 mbar. MS detection was performed in positive ionization mode.

The raw data were then converted to an open-source ".mzML" format (Proteowizzard) and load into R via the `MSnBase::readMSData()` function. In order to reduce data size, the test data was subsequently cutted in migration time and m/z range using `filterRt(rt = c(400, 900))` and `filterMz(mz = c(147.1, 152.0))` from `MSnBase`

### Author(s)

Liesa Salzer

---

msdata

*Sample FTICR, LC/MS and MS<sup>n</sup> data*

---

### Description

x object containing a subset of LC/MS raw data from a Thermo Finnigan LCQ Deca XP The data is a subset from 500-850 m/z and 1190-1310 seconds, incl. MS2 and MS3, intensity threshold 100.000. It was collected in positive ionization mode.

xs object containing a subset of FTICR data from a Bruker APex III FTICR. The data is a subset from 400-450 m/z, collected in positive ionization mode.

### Usage

```
data(xs)
```

### Format

The format is:

```
xs
```

### Details

The corresponding raw mzML files are located in the `fticr-mzML` and `iontrap` subdirectory of this package.

**See Also**

[xcmsSet](#), [xcmsRaw](#)

**Examples**

```
## The directory with the mzML LC/MS files
data(xs)
mzMLpath <- file.path(find.package("msdata"), "iontrap")
mzMLpath
files <- list.files(mzMLpath, recursive = TRUE, full.names = TRUE)
files

if (require(xcms)) {

  ## xcmsSet Summary
  show(xs)

  ## Access raw data file
  x <- xcmsRaw(files[1])
  x
}
}
```

---

proteomics

*Proteomics data in msdata*

---

**Description**

This function returns proteomics mass spectrometry files. These files are all stored in the `proteomics` directory in the `msdata` package. Each file/data is described in more details below.

**Usage**

```
proteomics(...)
```

**Arguments**

... Additional arguments passed to [list.files](#).

**Value**

A character with file names.

**Author(s)**

Laurent Gatto <laurent.gatto@uclouvain.be>

**See Also**

For more access to mass spectrometry-based proteomics data, see the `rpx` and `MsDataHub` packages.

---

sciexdata

*AB Sciex LC-MS data files*

---

## Description

The mzML files in the sciex directory in the msdata package represent profile-mode LC-MS data of pooled human serum samples (the same pool being measured). The samples were analyzed by ultra high-performance liquid chromatography (UHPLC; Agilent 1290) coupled to a Q-TOF mass spectrometer (TripleTOF 5600+ AB Sciex). The chromatographic separation was based in hydrophilic interaction liquid chromatography (HILIC) and performed using an Waters Acquity BEH Amide, 100 x 2.1 mm column.

The mass spectrometer was operated in full scan mode in the mass range from 50 to 1000 m/z and with an accumulation time of 250 ms. The files represent a subset of spectra/scans from m/z 105 to 134 and from retention time 0 to 260 seconds. The files were generated in the same LC-MS run, but from different injections. Details on the individual files are provided below.

## Details

- 20171016\_POOL\_POS\_1\_105-134.mzML profile-mode LC-MS data of pooled human serum samples. Injection index: 1.
- 20171016\_POOL\_POS\_3\_105-134.mzML profile-mode LC-MS data of pooled human serum samples. Injection index: 19.

## Author(s)

Sigurdur Smarason, Giuseppe Paglia and Johannes Rainer

## Examples

```
## List the files in the sciex folder
dir(system.file("sciex", package = "msdata"))
```

# Index

## \* datasets

msdata, 2

sciexdata, 4

CE-MS, 2

fdms3tmt11 (proteomics), 3

ident (proteomics), 3

list.files, 3

msdata, 2

proteomics, 3

quant (proteomics), 3

sciexdata, 4

x (msdata), 2

xcmsRaw, 3

xcmsSet, 3

xs (msdata), 2