SCIEX

The Power of Precision

SCIEX MS Data Converter 1.3.1 Utility

User Guide

RUO-IDV-05-14383-A

February 2022

AB Sciex Pte. Ltd.
Blk33, #04-06 Marsiling Industrial Estate Road 3
Woodlands Central Industrial Estate, Singapore 739256

Contents

Introduction 1

The SCIEX MS Data Converter utility converts wiff data acquired by a ZenoTOF 7600 mass spectrometer to the open data mzML and MGF formats. The utility can convert data directly, or it can process the data during the conversion, producing a peak list.

Supported Formats

The mzML Format

The mzML format is the single XML standard mass spectrometry format that was created by the merger of the older mzXML and mzData formats. The mzML data exported by the SCIEX MS Data Converter utility obeys the version 1.1 specification for the mzML format. For more information on mzML, go to http://www.psidev.info/index.php?q=wiki/Mass_Spectrometry.

Note: The XML format uses much more file space than wiff format to store the same data. As a result, the conversion of data in unreduced profile mode results in mzML output files that are many times larger than the original wiff data files. The utility uses compression functions to reduce file size, but the files are still very large.

The MGF Format

The Mascot generic format (MGF) was created by Matrix Science as input for the Mascot search engine, but it is now a widely used open format for proteomics applications. For more information about MGF, go to http://www.matrixscience.com/help/data_file_help.html.

The MGF format produced by this utility includes MS2 level data. If MS1 level data is required, then convert to mzML as an intermediate format, because it can contain the survey level data.

Enhancements and Fixed Issues

- The utility supports the Windows 10 operating system.
- The utility supports wiff files acquired by the ZenoTOF 7600 mass spectrometer. It does not support wiff2 files.

Known Issues and Limitations

- MRM HR data is not supported.
- Peptide mass fingerprinting (PMF) and other survey level data cannot be written to the MGF format.

Installation 2

Requirements

Operating System Requirements

- Windows 10
- Microsoft Visual C++ 2010 SP1 Redistributable Package (x86), available at: https://www.microsoft.com/en-us/download/details.aspx?id=26999. Select the **vcredist_x86.exe** option.

Install the SCIEX MS Data Converter

Prerequisites
• Remove any previous versions of the SCIEX MS Data Converter

1. Log on to the computer as a user with Administrative access.
2. Download the SCIEX MS Data Converter utility from sciex.com/software-support/software-downloads to a local folder.

   Tip! To prevent potential installation issues, save the file to a location other than the computer desktop and then disconnect any external USB storage devices before starting the installation.
3. After the download is complete, right-click the `sciex_ms_data_converter_V1.3.1.zip` file.
4. Click **Extract All**, browse to and select the destination folder, and then click **Extract**.
5. After the extraction is complete, browse to the extracted folder and then double click the `MSDataConverter.msi` file. Follow the on-screen instructions.

   The installation program creates the `C:\Program Files (x86)\AB SCIEX\MS Data Converter` folder.

Convert MS Data to mzML or MGF Format 3

Use this procedure to convert MS data using a batch file. This approach is recommended for users who are unfamiliar with command line tools.

Note: This procedure uses an example batch file, named `Example batch file.bat`.

1. Create a text file with a bat (batch) extension. For example, `Example batch file.bat`.

   Note: If the computer is configured to hide file extensions for known file types, then, in File Explorer, open the View tab and click **Options**. Go to the View tab and clear **Hide extensions for known file types**.

2. Add the converter commands to the batch file. Refer to the section: Command Line Syntax.

   Tip! At the end of the batch file, include a `pause` command, to keep the command window open after execution completes.

   ```
   AB_SCIEX_MS_Converter WIFF "D:\Data\File A.wiff"
      -proteinpilot MGF "D:\Data\File A.mgf"
   pause
   ```

   This example batch converts the wiff file `File A.wiff` to an MGF file with the same file name but a different extension. The peak list in the new file is the peak list that would be found by the ProteinPilot 4.5 software.

   To convert multiple input files in the same batch file, include multiple conversion commands.

Tip! On a computer with multiple cores, mutliple conversions can be run in parallel, significantly reducing processing time. To do so, start each conversion as a separate instance of the converter utility, as in the following example.

```
start "<instance1>" /i AB_SCIEX_MS_Converter WIFF
  "D:\File A.wiff" -proteinpilot MGF "D:\File A.mgf"
timeout 10
start "<instance2>" /i AB_SCIEX_MS_Converter WIFF
  "D:\File B.wiff" -proteinpilot MGF "D:\File B.mgf"
timeout 10
start "<instance3>" /i AB_SCIEX_MS_Converter
  WIFF "D:\File C.wiff" -proteinpilot MGF "D:\File C.mgf"
timeout 10
start "<instance4>" /i AB_SCIEX_MS_Converter WIFF
  "D:\File D.wiff" -proteinpilot MGF "D:\File D.mgf"
timeout 10
pause
```

Where *<instance1>* to *<instance4>* are user-defined names for the different instances of the utility.

3. Save the batch file.
4. Copy the batch file to the `C:\Program Files (x86)\AB SCIEX\MS Data Converter` folder.
5. Double-click the batch file to run it.

Command Line Syntax 4

```
AB_SCIEX_MS_Converter <input format> <input data>
  <output content type> <output format> <output file>
```

Table 4-1 Parameters

Parameter	Description
<input format>	The format of the input data. Use the following value: • `wiff`: Data acquired by SCIEX OS from a ZenoTOF 7600 system **Note:** The utility does not support wiff2 data.
<input data>	The path to the input data file, including the file name and extension. For example, `C:\Data\File A.wiff`. If the input file is in the folder that contains the utility files, then the path can be omitted. For wiff data types with multiple file components, such as scan, specify the wiff file only. Make sure that the other files are in the same folder as the wiff file. **Note:** The path must be a local path. The utility does not support conversion across a network.
<output content type>	The information to be included in the output data. Use one of the following values: • `-profile`: Includes the full X-Y trace recorded by the mass spectrometer, without any reduction in the information content. This mode cannot be used for conversion to the MGF format. • `-centroid`: Includes the centroided peak list, as calculated in real time, during acquisition, by the mass spectrometer control software. • `-proteinpilot`: Includes the peak list that would be searched by the ProteinPilot software version 4.5. The raw data is processed by a slower but higher quality signal processing approach that produces better results for protein identification applications. **Note:** The results are not the same as those generated by the most recent version of the ProteinPilot software.

Table 4-1 Parameters (continued)

Parameter	Description
<output format>	The format of the output data. Use one of the following values: • `MGF`: Exports MGF data. Only MS/MS spectra are converted. This format type cannot be used with the `-profile` option, because MGF data is in peak list format. • `MZML`: Exports mzML data. All data levels are converted (MS1, MS2, and so on).
<output file>	The path to the output data file, including the file name and extension. For example, `C:\Data\File A.mgf`. **Note:** The data file cannot be written to the folder that contains the utility files. **Note:** The path must be a local path. The utility does not support conversion across a network.

Appendix A

mzML Output

Table A-1 Conversion Options

Option	Description
Profile mode	Profile data is written as it is measured by the ZenoTOF 7600 mass spectrometer, and therefore is not regularly spaced.
Centroid mode	Unlike in the MGF export where fragment peaks are converted to their +1 charge equivalent *m/z* value, the *m/z* values are written as the actual observed *m/z* value.
ProteinPilot mode	Unlike in the MGF export where fragment peaks are converted to their +1 charge equivalent *m/z* value, the *m/z* values are written as the actual observed *m/z* value.

MGF Output

The data written during conversion to MGF format obeys the MGF specification as much as possible. There are few fine details not clearly prescribed by this specification. When writing the MS/MS fragments for a spectrum, this utility provides three columns of data to describe fragment peaks.

Table A-2 Fragment Peak Data

Column	Description
Column 1	The *m/z* value of the +1 charge state of a fragment, regardless of the charge state that is actually observed. Mascot prefers searching fragments in their +1 equivalent *m/z* value, rather than the observed *m/z* value. If the charge state is unknown, then the observed *m/z* value is listed.
Column 2	The intensity of the fragment peak measured as the apex intensity of the monoisotopic peak. These are real measured intensity, not normalized intensity, measurements.

Table A-2 Fragment Peak Data (continued)

Column	Description
Column 3	The charge state at which the fragment was actually observed. Fragment ions are listed as their corresponding +1 charge state *m/z* value and this charge value allows for the reconstruction of the observed spectrum. If a fragment is observed at multiple charge states in a spectrum, there there will be multiple rows with very similar *m/z* values and different charge states. If the charge cannot be determined, as is the case with low resolution data, a value of 0 is listed.

Contact Us

Customer Training

- In North America: NA.CustomerTraining@sciex.com
- In Europe: Europe.CustomerTraining@sciex.com
- Outside the EU and North America, visit sciex.com/education for contact information.

Online Learning Center

- SCIEX Now Learning Hub

SCIEX Support

SCIEX and its representatives maintain a staff of fully-trained service and technical specialists located throughout the world. They can answer questions about the system or any technical issues that might arise. For more information, visit the SCIEX website at sciex.com or contact us in one of the following ways:

- sciex.com/contact-us
- sciex.com/request-support

CyberSecurity

For the latest guidance on cybersecurity for SCIEX products, visit sciex.com/productsecurity.

Documentation

This version of the document supercedes all previous versions of this document.

To view this document electronically, Adobe Acrobat Reader is required. To download the latest version, go to https://get.adobe.com/reader.

The latest versions of the documentation are available on the SCIEX website, at sciex.com/customer-documents.

Note: To request a free, printed version of this document, contact sciex.com/contact-us.