LipidXplorer: Software for Quantitative Shotgun Lipidomics Compatible with Multiple Mass Spectrometry Platforms

Ronny Herzog,[1,2] Dominik Schwudke,[3] and Andrej Shevchenko[1]

[1]MPI of Molecular Cell Biology and Genetics, Dresden, Germany
[2]Lipotype GmbH, Dresden, Germany
[3]Research Center Borstel, Borstel, Germany

ABSTRACT

LipidXplorer is an open-source software kit that supports the identification and quantification of molecular species of any lipid class detected by shotgun experiments performed on any mass spectrometry platform. LipidXplorer does not rely on a database of reference spectra: instead, lipid identification routines are user defined in the declarative molecular fragmentation query language (MFQL). The software supports batch processing of multiple shotgun acquisitions by high-resolution mass mapping, precursor and neutral-loss scanning, and data-dependent MS/MS lending itself to a variety of lipidomics applications in cell biology and molecular medicine. *Curr. Protoc. Bioinform.* 43:14.12.1-14.12.30.

Keywords: mass spectrometry • shotgun lipidomics • MS/MS • LipidXplorer • lipid identification • lipid quantification • lipidomics screen

INTRODUCTION

Lipdomics is a systematic effort toward the quantitative characterization of the entire lipid complement of organelles, cells, tissues, or entire organisms (reviewed in Dennis, 2009; Shevchenko and Simons, 2010; Wenk, 2010, Gross and Han, 2011; Harkewicz and Dennis, 2011). It addresses a task of enormous analytical complexity—according to different estimates, a eukaryotic lipidome might comprise tens of thousands of molecular species from a few hundred lipid classes (reviewed in Shevchenko and Simons, 2010). Lipids are typically extracted using chloroform/methanol (Folch et al., 1957; Bligh and Dyer, 1959) or methyl-*tert*-butyl ether (Matyash et al., 2008), and extracts are either analyzed by LC-MS or LC-MS/MS or directly infused into a mass spectrometer. The latter approach is termed shotgun lipidomics (reviewed in Han and Gross, 2005; Han et al., 2012), and implies that lipid species are either identified by accurately determined intact masses (top-down shotgun lipidomics; Schwudke et al., 2007a) or (in bottom-up shotgun lipidomics) all plausible precursor ions are subjected to MS/MS and individual species identified by structure specific fragment ions (Schwudke et al., 2006; Schuhmann et al., 2011). Shotgun analyses provide a rapid and sufficiently comprehensive quantitative overview of the bulk composition of lipidomes. Because of its sensitivity as well as vastly simplified sample preparation and spectra-acquisition methods, it lends itself to high-throughput full-lipidome screens. However, shotgun spectra datasets are exceedingly complex and cannot be interpreted manually. Indeed, a typical lipidomics study might encompass hundreds of unique samples, from each of which ∼100 MS and 1000 MS/MS spectra are acquired in several technical and biological replicates.

The LipidXplorer software is designed for supporting shotgun lipidomics profiling of series of samples of any lipid composition using any mass spectrometry platform

Current Protocols in Bioinformatics 14.12.1-14.12.30, October 2013
Published online October 2013 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/0471250953.bi1412s43

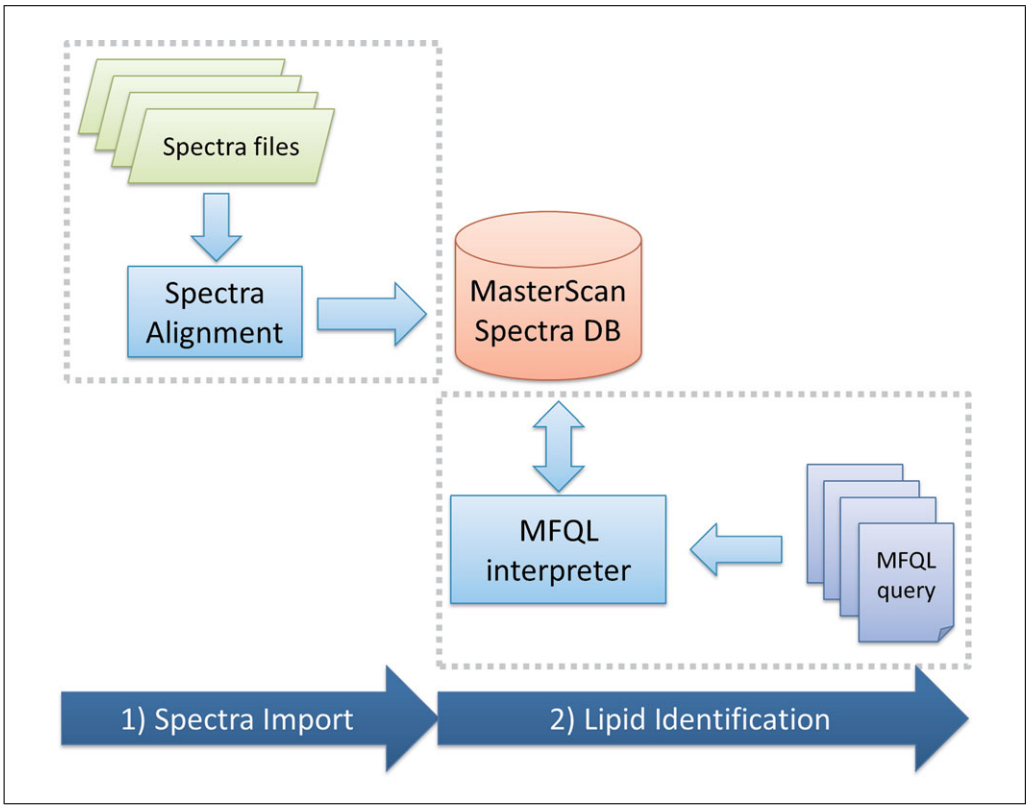


Figure 14.12.1 Basic LipidXplorer workflow. (1) Shotgun MS and MS/MS spectra acquired from all samples in the experiment series are imported into a MasterScan database. All scans of each acquisition are averaged into a single representative spectrum. All related spectra are subsequently aligned in an experiment- and mass spectrometer–dependent manner and stored in the MasterScan database. (2) Lipids are identified by probing MasterScan with queries written in the Molecular Fragmentation Query Language (MFQL). In this way, lipid identification relies on specific reporter ion(s) detected in MS and/ or MS/MS spectra; in the latter scenario, writing MFQLs requires the knowledge of major fragmentation pathways expected for a given lipid class; however, it does not require a resource of reference spectra.

(Fig. 14.12.1). First, it organizes acquired MS and MS/MS spectra into a flat-file database (MasterScan). Then, the molecular fragmentation query language (MFQL) formalizes fragmentation pathways specific for analyzed lipid classes and implements them into customized spectra interpretation methods for probing the MasterScan, identifying lipid species and reporting the abundances of user-defined ions for subsequent lipid quantification.

The LipidXplorer interpretation routine starts with importing MS and MS/MS spectra into the MasterScan, recognizing the spectra related to each other, and aligning them in a mass resolution–dependent manner, while the associations of MS/MS spectra to their precursor ions are maintained. Users control importing and aligning of spectra by defining instrument- and experiment-dependent settings. A single MasterScan file retains essential mass spectrometric data acquired from all samples within the series of experiments. The same MasterScan file can be screened by any combination of individual MFQL queries, each of which defines the identification routine for all species of a lipid class [for example, all ester (PE) or ether (PE-*O*) phosphatidylethanolamines], or their specific subclass (e.g., only PE-*O* plasmalogens). For identifying lipids, LipidXplorer does not rely on a reference spectra resource, and the identification confidence is not compromised if several isobaric species of different lipid classes were co-selected for MS/MS and contributed structure-specific fragments to the mixed spectrum. Altogether, this informatics solution helps to utilize the full content of all MS and MS/MS spectra;

for example, species of the same lipid class can be identified in several independent ways by selecting different combinations of fragment ions, or solely by matching accurate masses of intact precursors. This is achieved by probing the same MasterScan file with multiple MFQL queries, and there is no need to re-acquire the mass spectra dataset.

The identified lipids are then reported for quantification. The report file content is user defined: it is compatible with any arbitrary convention for naming the identified lipid species and can output the abundance(s) of any ion(s) relevant to the quantification. Note that LipidXplorer has no fixed quantification routine since, in our experience, this is very much project dependent.

This unit is a detailed guide encompassing key steps of the lipid-identification procedure from importing files with shotgun mass spectra and applying MFQL scripts from supplied libraries (Basic Protocol 2) to compiling customized MFQL scripts (Basic Protocols 3, 4, and 5) and critical assessment and troubleshooting (Basic Protocol 6) of lipid identifications. Here we rely on our >5 years experience of using LipidXplorer in practical lipidomics and valuable critical feedback provided by the worldwide community of users.

STRATEGIC PLANNING: SUPPORTED FILE FORMATS

LipidXplorer supports shotgun lipidomics experiments across multiple instrumentation platforms and is not associated with any particular type of mass spectrometer or any instrument vendor. It supports the generic mass spectra file formats mzXML and mzML (Deutsch, 2008), and also peak lists in the formats of `*.csv` for MS spectra and `*.dta` for MS/MS spectra.

Peak lists should be provided according to the directory structure in Figure 14.12.2. The top-level directory `<project name>\` assigns the project name. All samples included in this directory are imported. Each subdirectory `<[neg_]Sample 1 -- n>` (where the option `neg_` indicates negative acquisition polarity, which otherwise is regarded as positive) contains all spectra acquired from the sample. The survey (MS) spectra should be provided as `*.csv` files containing a list of `<mass>`, `<intensity>` pairs. MS/MS spectra (if any were acquired) should be in `*.dta` format. It is also possible to supply only MS/MS spectra: in this case, the survey (MS) spectrum will be recreated from precursor masses targeted for MS/MS as specified in the corresponding `*.dta` files.

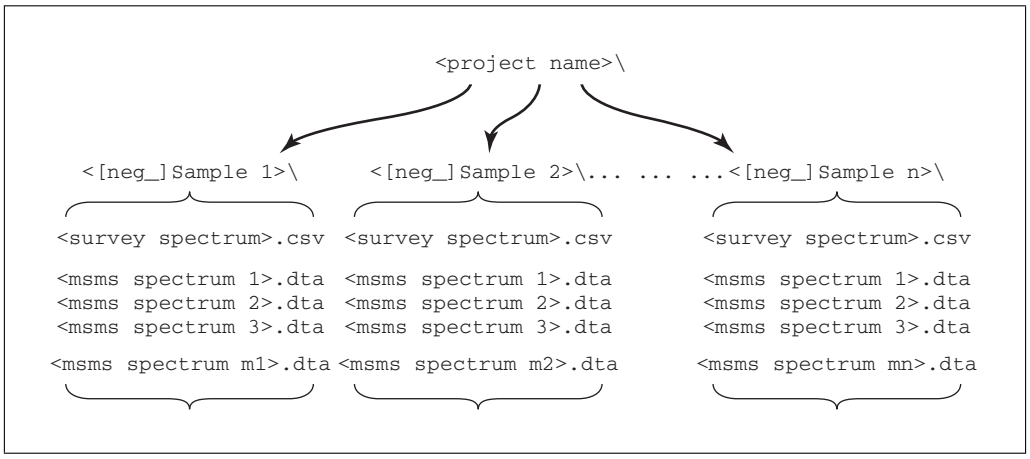


Figure 14.12.2 File structure for importing spectra in peak list format.

BASIC PROTOCOL 1

INSTALLING LipidXplorer

LipidXplorer is written in the Python programming language. Python is an interpreted language, and it is necessary to install the Python interpreter and some additional modules prior installing LipidXplorer. The LipidXplorer wiki: *https://wiki.mpi-cbg.de/lipidx/LipidXplorer_Installation* provides step-by-step guidelines and links to all required external modules.

Necessary Resources

Hardware

Windows-compatible computer with 32-bit (x86) or 64-bit processor; >100 MB free hard disk space, and minimum 1 GB RAM

Software

Microsoft Windows XP, Vista, or Windows 7

1. Install Python and all necessary modules using links provided at the LipidXplorer wiki (*https://wiki.mpi-cbg.de/lipidx/LipidXplorer_Installation*; the links to additional modules are also provided there) according to your computer setup (32- or 64-bit environment).
2. Once Python is installed, proceed with installing LipidXplorer: download the LipidXplorer version 1.2.5 zip package from *http://sourceforge.net/projects/lipidxplorer/files/LipidXplorer-1.2.5.zip/download* and save it anywhere on your computer.
3. Unzip the file: right-click on the `LipidXplorer-1.2.5.zip` file and select "Extract All...". Chose the directory where you want to install LipidXplorer and click on Extract. The whole folder contains everything required for executing LipidXplorer. There is no installation routine to follow, and, at any time, you can move the LipidXplorer folder to another directory or another computer without uninstalling it. Optionally, you may now delete the `LipidXplorer-1.2.5.zip` file.
4. Move to the directory where LipidXplorer is installed and start the program by clicking on `LipidXplorer.py`. The program modules for LipidXplorer are located in the `lx\` directory. The `mfql\` directory contains example MFQL queries, which were used in a few recent articles. The URL files with the `infos_` prefix locate these articles at the LipidXplorer wiki site. The folders with the `neg_` and `pos_` prefixes contain actual MFQL queries for the lipid identification in negative and positive acquisition mode, respectively. Folder names refer to the articles in which these MFQLs were used. Example MFQLs, spectra datasets, and corresponding lipid identification reports are also available on the wiki page and can be used for training or testing LipidXplorer's performance.

SUPPORT PROTOCOL

FILE CONVERSION USING MSConvert

LipidXplorer (version 1.2.4 and later) imports data in mzXML and mzML formats. To convert original spectra saved in the vendor's format (such as `*.wiff` from Sciex or `*.raw` from Thermo Fisher Scientific) to mzXML/mzML, users can apply MSConvert, a utility program from the ProteoWizard software tool kit, available at *http://proteowizard.sourceforge.net/downloads.shtml*.

Below, we describe how to use MSConvert for converting files from proprietary formats into generic `*.mzXML` or `*.mzML` formats readable by LipidXplorer. However, please note that file conversion still requires the original vendor's software to access the spectra. Although `*.mzXML` files are usually more compact, in our experience, this

format might cause problems with certain instruments or spectra-acquisition methods, particularly when importing precursor ion scan spectra acquired on a triple quadrupole mass spectrometer. We therefore recommend that, whenever possible, the more advanced `*.mzML` format should be used instead.

Necessary Resources

Hardware

Windows-compatible computer with 32-bit (x86) or 64-bit processor; >100 MB free hard disk space, and minimum 1 GB RAM

Software

Microsoft Windows XP, Vista, or Windows 7

1. Download and unzip MSConvert from the URL above.
2. Start MSConvert by clicking on `MSConvertGUI.exe`.
3. Click on Browse and select the spectra that you want to convert. The selected files are listed in the text box below the Browse button. By default, the output directory is the same as the import directory. To change this, click on Browse next to "Output Directory:" (see Fig. 14.12.3).

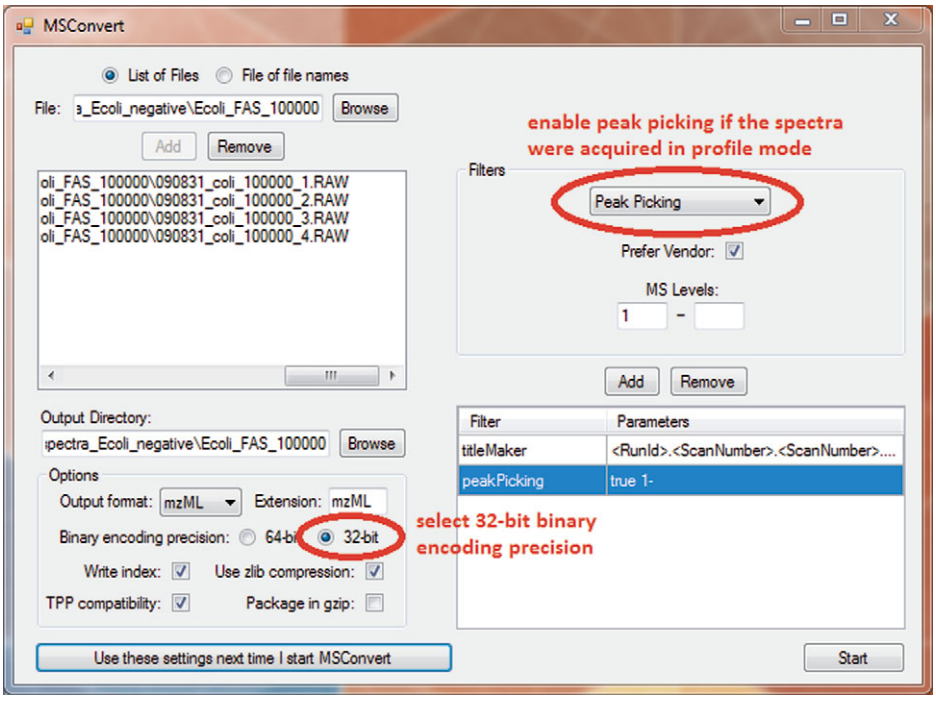


Figure 14.12.3 A screenshot of the MSConvert GUI. MSConvert is the preferred conversion tool to generate `*.mzXML` or `*.mzML` data files that are compatible with LipidXplorer. Essential settings are circled. Select the files to convert using Browse and target data format by "Output format:". Choose 32-bit encoding precision. If spectra were acquired in the profile mode, select the Peak Picking option and click Add. The 1- under the filter name indicates that all spectra should be converted.

4. Select “mzML” from the pull-down menu next to “Output format:”. The extension is automatically filled with “mzML”.
5. Select “32-bit” for the “Binary encoding precision”. Currently, LipidXplorer only supports 32-bit encoding.
6. Make sure that “Write index” is checked and “Package in gzip” is unchecked. Applying “Use zlib compression” decreases the file size by ∼50%.
7. LipidXplorer only imports centroided spectra, i.e., peak lists formatted as (*m*/*z*, peak intensity) pairs. If source spectra were acquired in profile mode, the Peak Picking function of MSConvert should be applied to produce centroided spectra: select Peak Picking from the filters at the left side of the GUI and click Add. The peak picker shows as a filter in the table below. The parameter should be `true 1-` to enable peak picking in all MS and MS/MS spectra.
8. Click Start to start the conversion. A window appears showing the conversion progress. Once the conversion is completed, close the window.
9. Now, spectra are ready for importing by LipidXplorer.

BASIC PROTOCOL 2

BASIC LIPID IDENTIFICATION WORKFLOW

This protocol exemplifies a spectra-interpretation routine that is typically adopted by new users. It employs a supplied library of MFQL scripts that were developed for and tested on a common type of a tandem mass spectrometer, the LTQ Orbitrap. Therefore, the identification of major species of most common lipid classes (such as, in this example, PA, PE, PG) does not require adjusting MFQL scripts or fine tuning spectra import settings. Experienced users may also use this protocol with supplied MFQL and spectra datasets to test if their installation of LipidXplorer is operating correctly or that their spectra acquisition methods have provided data of required quality.

Necessary Resources

Hardware

PC computer under Microsoft Windows XP, Vista, or 7

Software

LipidXplorer, downloaded and installed as described in Basic Protocol 1

Files

Download the collection of test spectra from *https://wiki.mpi-cbg.de/lipidx/File:Spectra_article.zip*. Open the `*.zip` file and store the content of `spectra_article\*` to a place of your choice. LipidXplorer comes with a suggested folder structure: for example, you can store the spectra under: `LipidXplorer\spectra\negative\`

The folder `spectra_article\spectra_Ecoli_negative\` includes sub-folders containing spectra acquired from an *E.coli* total lipid extract on an LTQ Orbitrap mass spectrometer under different mass-resolution settings. The name of each folder refers to the applied resolution. All spectra were acquired under data-dependent acquisition, i.e., each acquisition consists of a large number of MS and MS/MS spectra (see Herzog et al., 2011, for further details).

Select spectra

1. Start LipidXplorer and open Import Source panel. Click on the green Browse button and select the folder `<your_chosen_folder>\spectra_Ecoli_negative\Ecoli_FAS_100000`. The folder contains all spectra that should be

imported into the MasterScan database. In this case, the folder contains four independent acquisitions of *E.coli* extract on a LTQ Orbitrap in negative mode.

2. Make sure that you selected the correct file format (next to the green Browse button). In this case it should be "mzML".

Configure spectra import settings

3. Switch to the panel Import Settings. In this menu, the spectra import settings should be configured such that they match the expected performance of your mass spectrometer (Fig. 14.12.4).

4. All individual settings in this panel can be stored as an `*.ini` file and reloaded under Select a Configuration. To load the `*.ini` file with the configurations required by this protocol, click on Browse and select `lpdxImportSettings_tutorial.ini`, which is located in your LipidXplorer folder. Now, select the configuration appropriate for processing the spectra (in this case it is FAS_LTQ_100000MS1). Note that the correct selection of import settings is crucial for the accurate spectra interpretation, and both mass resolution and mass resolution gradient are having a particularly strong impact. Do not use arbitrary values, as they might not match the actual performance of your instrument. The full list of import options and their detailed description is provided in the appendix at the end of this unit.

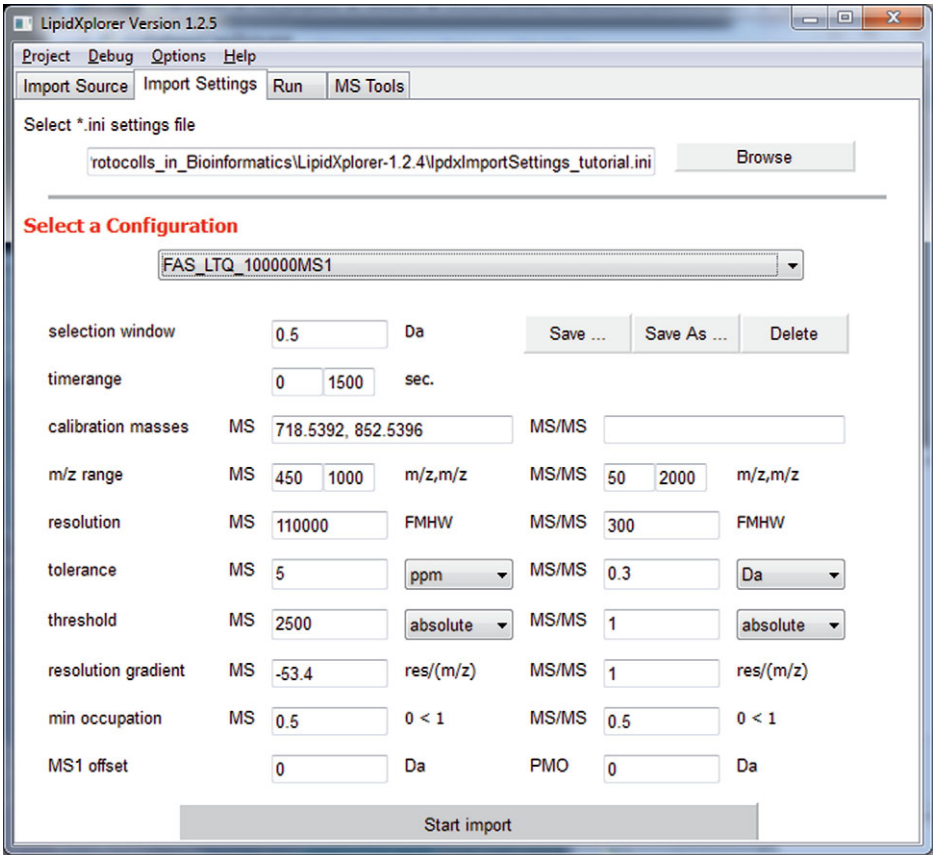


Figure 14.12.4 Import settings panel. All necessary settings for importing spectra are controlled here; they are explained in detail in the appendix at the end of the unit.

5. Hit the Start import button at the bottom of the panel. The Debug window will report on the processing progress. Upon import completion, a pop-up window "Task completed" appears.

Lipid identification

6. Turn to the panel Run. Here you select the MFQL queries for probing the MasterScan database (Fig. 14.12.5).

7. Click on "Add MFQL directory" to select all MFQL files in a folder. Select `LipidXplorer\mfql\negative_mode`. This is a collection of MFQL queries for identifying *E.coli* lipids using MS/MS spectra acquired negative mode. Once selected, the queries are listed in the query window. Another way to load queries is to "drag and drop" the folder or individual files.

8. Make sure that both the "Isotopic correction MS" and "Isotopic correction MS/MS" boxes are checked. The "Type I" correction adjusts the intensity of monoisotopic peaks by summing up the intensities of peaks within the isotopic cluster. The "Type II" correction resolves the intensities of isotopic peaks overlapping between neighboring isotopic clusters, (see Herzog et al., 2011, 2012, for the algorithm's details).

9. Click on Run LipidXplorer to start the lipid identification.

10. The Debug Window will open and inform about the current tasks performed by LipidXplorer.

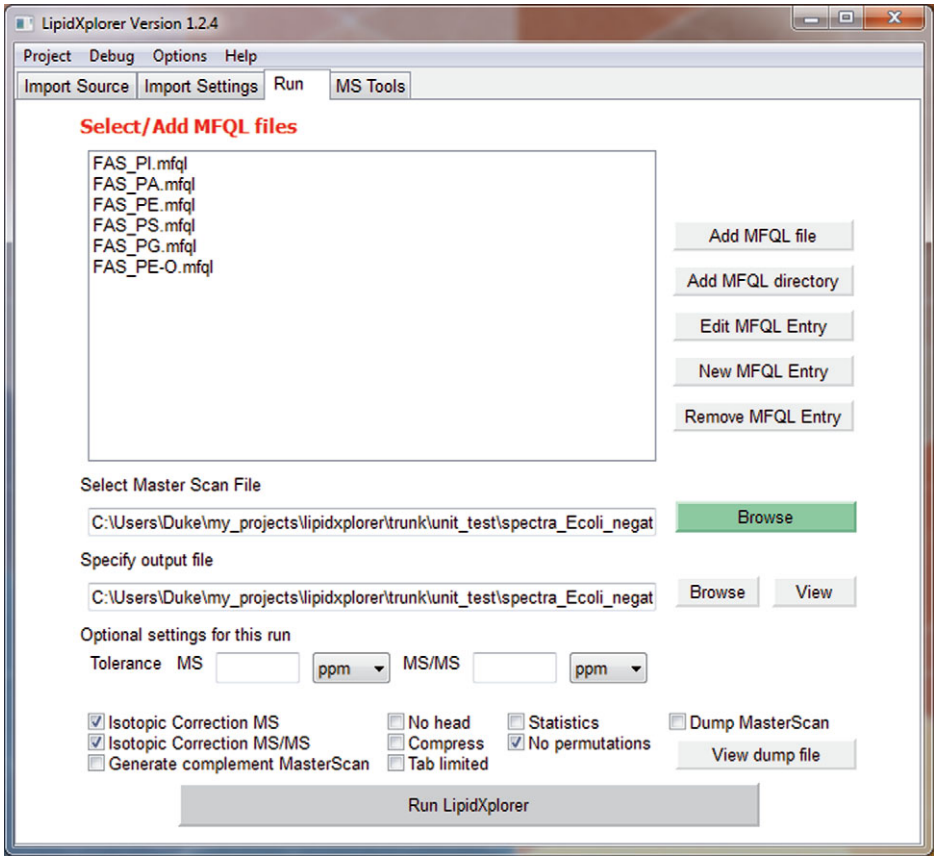
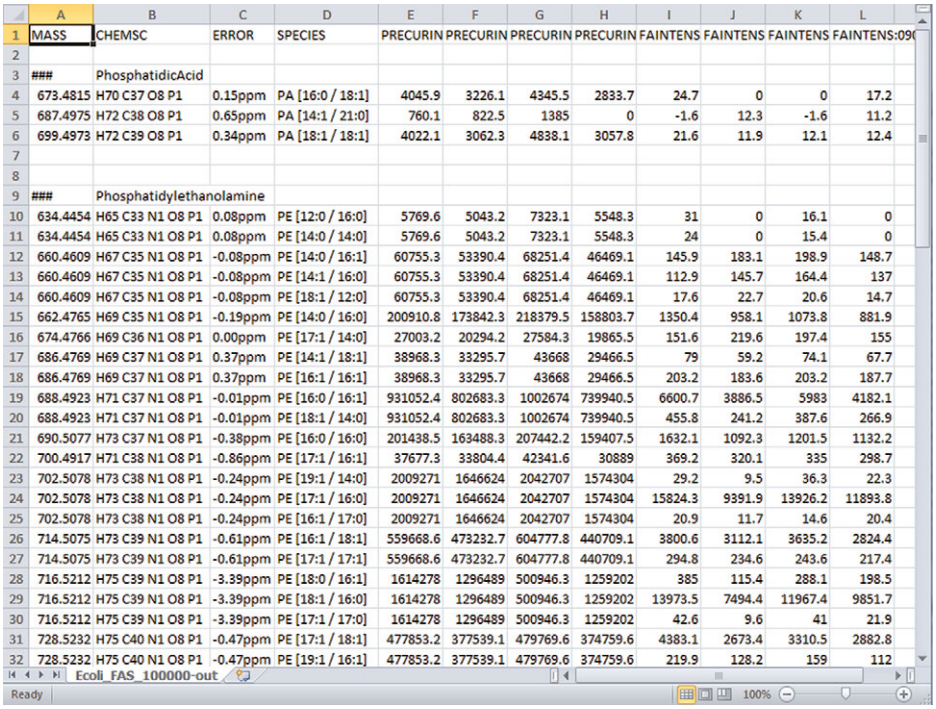


Figure 14.12.5 Run panel controls screening the MasterScan with MFQL queries. Several MFQLs can be applied in parallel.



	A	B	C	D	E	F	G	H	I	J	K	L
1	MASS	CHEMSC	ERROR	SPECIES	PRECURIN	PRECURIN	PRECURIN	PRECURIN	FAINTENS	FAINTENS	FAINTENS	FAINTENS:09(
2												
3	###	PhosphatidicAcid										
4	673.4815	H70 C37 O8 P1	0.15ppm	PA [16:0 / 18:1]	4045.9	3226.1	4345.5	2833.7	24.7	0	0	17.2
5	687.4975	H72 C38 O8 P1	0.65ppm	PA [14:1 / 21:0]	760.1	822.5	1385	0	-1.6	12.3	-1.6	11.2
6	699.4973	H72 C39 O8 P1	0.34ppm	PA [18:1 / 18:1]	4022.1	3062.3	4838.1	3057.8	21.6	11.9	12.1	12.4
7												
8												
9	###	Phosphatidylethanolamine										
10	634.4454	H65 C33 N1 O8 P1	0.08ppm	PE [12:0 / 16:0]	5769.6	5043.2	7323.1	5548.3	31	0	16.1	0
11	634.4454	H65 C33 N1 O8 P1	0.08ppm	PE [14:0 / 14:0]	5769.6	5043.2	7323.1	5548.3	24	0	15.4	0
12	660.4609	H67 C35 N1 O8 P1	-0.08ppm	PE [14:0 / 16:1]	60755.3	53390.4	68251.4	46469.1	145.9	183.1	198.9	148.7
13	660.4609	H67 C35 N1 O8 P1	-0.08ppm	PE [14:1 / 16:0]	60755.3	53390.4	68251.4	46469.1	112.9	145.7	164.4	137
14	660.4609	H67 C35 N1 O8 P1	-0.08ppm	PE [18:1 / 12:0]	60755.3	53390.4	68251.4	46469.1	17.6	22.7	20.6	14.7
15	662.4765	H69 C35 N1 O8 P1	-0.19ppm	PE [14:0 / 16:0]	200910.8	173842.3	218379.5	158803.7	1350.4	958.1	1073.8	881.9
16	674.4766	H69 C36 N1 O8 P1	0.00ppm	PE [17:1 / 14:0]	27003.2	20294.2	27584.3	19865.5	151.6	219.6	197.4	155
17	686.4769	H69 C37 N1 O8 P1	0.37ppm	PE [14:1 / 18:1]	38968.3	33295.7	43668	29466.5	79	59.2	74.1	67.7
18	686.4769	H69 C37 N1 O8 P1	0.37ppm	PE [16:1 / 16:1]	38968.3	33295.7	43668	29466.5	203.2	183.6	203.2	187.7
19	688.4923	H71 C37 N1 O8 P1	-0.01ppm	PE [16:0 / 16:1]	931052.4	802683.3	1002674	739940.5	6600.7	3886.5	5983	4182.1
20	688.4923	H71 C37 N1 O8 P1	-0.01ppm	PE [18:1 / 14:0]	931052.4	802683.3	1002674	739940.5	455.8	241.2	387.6	266.9
21	690.5077	H73 C37 N1 O8 P1	-0.38ppm	PE [16:0 / 16:0]	201438.5	163488.3	207442.2	159407.5	1632.1	1092.3	1201.5	1132.2
22	700.4917	H71 C38 N1 O8 P1	-0.86ppm	PE [17:1 / 16:1]	37677.3	33804.4	42341.6	30889	369.2	320.1	335	298.7
23	702.5078	H73 C38 N1 O8 P1	-0.24ppm	PE [19:1 / 14:0]	2009271	1646624	2042707	1574304	29.2	9.5	36.3	22.3
24	702.5078	H73 C38 N1 O8 P1	-0.24ppm	PE [17:1 / 16:0]	2009271	1646624	2042707	1574304	15824.3	9391.9	13926.2	11893.8
25	702.5078	H73 C38 N1 O8 P1	-0.24ppm	PE [16:1 / 17:0]	2009271	1646624	2042707	1574304	20.9	11.7	14.6	20.4
26	714.5075	H73 C39 N1 O8 P1	-0.61ppm	PE [16:1 / 18:1]	559668.6	473232.7	604777.8	440709.1	3800.6	3112.1	3635.2	2824.4
27	714.5075	H73 C39 N1 O8 P1	-0.61ppm	PE [17:1 / 17:1]	559668.6	473232.7	604777.8	440709.1	294.8	234.6	243.6	217.4
28	716.5212	H75 C39 N1 O8 P1	-3.39ppm	PE [18:0 / 16:1]	1614278	1296489	500946.3	1259202	385	115.4	288.1	198.5
29	716.5212	H75 C39 N1 O8 P1	-3.39ppm	PE [18:1 / 16:0]	1614278	1296489	500946.3	1259202	13973.5	7494.4	11967.4	9851.7
30	716.5212	H75 C39 N1 O8 P1	-3.39ppm	PE [17:1 / 17:0]	1614278	1296489	500946.3	1259202	42.6	9.6	41	21.9
31	728.5232	H75 C40 N1 O8 P1	-0.47ppm	PE [17:1 / 18:1]	477853.2	377539.1	479769.6	374759.6	4383.1	2673.4	3310.5	2882.8
32	728.5232	H75 C40 N1 O8 P1	-0.47ppm	PE [19:1 / 16:1]	477853.2	377539.1	479769.6	374759.6	219.9	128.2	159	112

Ecoli_FAS_100000-out

Ready

100%

Figure 14.12.6 Lipids identified in the collection of tutorial spectra. The lipid naming convention and the content of each column is defined in a MFQL query. The results of the queries are separated by ### and the query name.

View the results

11. Once the MasterScan has been screened with all selected MFQL queries, the pop-up window "Task completed" will appear. Click on View next to the Browse button to view the results, or just open the output file `<your_chosen_folder>\spectra_Ecoli_negative\Ecoli_FAS_100000\Ecoli_FAS_10000-out.csv` with spreadsheet software like Microsoft Excel or OpenOffice. The output should match Figure 14.12.6.

COMPOSING MFQL QUERIES FOR TOP-DOWN SHOTGUN LIPIDOMICS

BASIC PROTOCOL 3

This protocol describes the basic organization of MFQL queries and explains how to create queries for the top-down identification of lipids by matching their accurately determined intact masses. As an example, we will consider top-down profiling of phosphatidylethanolamines (PE) in negative mode using LTQ Orbitrap.

The Molecular Fragmentation Query Language (MFQL) supports basic data types like strings and floating-point numbers. Further, there are specific data types: elemental compositions (also often called sum compositions) and sum composition constraints (*sc*-constraints). A sum composition specifies the exact number of corresponding atoms in the intact precursor ion or its fragment and is placed within single quotation marks—for example: `'C35 H67 N1 O8 P1'`.

Sc-constraints describe molecules or fragments whose sum composition may vary within the specified ranges, For example: `'C[31..49] H[30..200] N[1] O[8] P[1]'`.

Lipid molecules may contain a few double bonds. LipidXplorer expresses the unsaturation of a molecule as its double-bond equivalent (Sparkman, 2006). In *sc*-constraints, the

equivalent may vary within a specified range termed DBR for Double Bond Range, for example: `'C[31..49] H[30..200] N[1] O[8] P[1]' WITH DBR = (2.5,9.5).`

Any MFQL query is organized in three or four sections:

a. The DEFINE section specifies sum compositions, *sc*-constraints, or exact *m*/*z* of intact lipid molecules or their fragments, which will be further looked up in MS and MS/MS spectra; they are associated with user-defined names.
b. The IDENTIFY section determines in which spectra (MS or MS/MS) and in which way these defined variables are looked up.
c. The SUCHTHAT section is optional. Additional constraints are applied here as mathematical expressions, inequalities or functions of sum compositions, and other data types.
d. The REPORT section specifies the format and content of the output. Upon MasterScan search, every query returns a table, whose column headers are defined here. This section also explains the convention for naming the identified lipid species.

Necessary Resources

Hardware

PC computer under Microsoft Windows XP, Vista, or 7.

Software

LipidXplorer, downloaded and installed as described in Basic Protocol 1

Files

MasterScan file produced according to Basic Protocol 2

Composing an MFQL query

1. Start LipidXplorer and go to the Run panel.
2. Click on New MFQL Entry.
3. Give the name and location for your MFQL query file and click Save.
4. The newly created file is opened in a new panel in LipidXplorer. Turn to this panel.
5. Write the first line:

   ```
   QUERYNAME = Phosphatidylethanolamine;
   ```

   This line is mandatory; every MFQL query starts with the name (note that in this example we will be identifying PE, hence we give it this name). The command ends with the semicolon symbol `;`.
6. Compile the DEFINE section:

   ```
   DEFINE prPE = 'C[31..49] H[30..200] N[1] O[8] P[1]'
   WITH DBR = (2.5,9.5), CHG = 1;
   ```

   Here we defined a variable `prPE` (precursor of PE), whose *sc*-constraint covers the range of expected sum compositions of PE species; effectively, we defined that they should contain exactly one atom of nitrogen, eight oxygen atoms, and one phosphorus atom. At the same time, the number of carbon and hydrogen atoms, as well as double bonds, may vary within a relatively broad range accommodating the compositional variability of their fatty acid moieties.

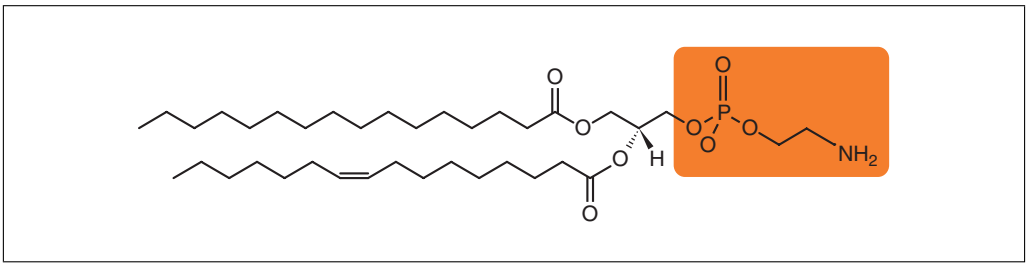


Figure 14.12.7 Chemical structure of phosphatidylethanolamine (PE) species (PE 16:0/16:1 is shown here as an example). The phosphoethanolamine head group (highlighted) defines the lipid class, while the structural variability within the class originates from varying the lengths (number of carbon atoms) and unsaturation (number of double bonds) at the two fatty acid moieties.

The general syntax of the DEFINE command is:

```
DEFINE <variable name> = <sum composition> | <sc-
constraint> | <m/z value> WITH <options>;
```

The keyword `WITH` follows the *sc*-constraint and makes it possible to impose other optional constraints. The option `DBR` (Double Bond Range) defines the lower and upper limits of the double-bond equivalent allowed for species of this lipid class. For PE, two double-bond equivalents are required because of two C=O groups in the fatty acid moieties (Fig. 14.12.7). Since here the *sc*-constraints are defined for ions (rather than neutral molecules), the range of double-bond equivalents is increased by 0.5 for singly charged anions and decreased by 0.5 for singly charged cations. Here we expect that PEs were detected as singly charged anions, and therefore DBR 2.5 to 9.5 presumes that PE anions comprise zero to the maximum of seven double bonds in both fatty acid moieties. The option CHG defines the ion charge state (see Fig. 14.12.8).

7. Compile the IDENTIFY section:

```
IDENTIFY
  prPE IN MS1-;
```

Here, we requested LipidXplorer to look for prPE variable in all MS spectra acquired in negative ion mode.

8. Compile the REPORT section:

```
REPORT
  MASS = prPE.mass;
  CHEMSC = prPE.chemsc;
  ERROR = "%2.2f" % "(prPE.errppm)";
  SPECIES = "PE [%d:%d]" % (prPE.chemsc[C] -- 5, prPE.
  chemsc[db] -- 2.5);
  INTENS = prPE.intensity;
;
```

In the REPORT section, each line defines a column in the output table. Above we defined five columns containing the mass (MASS), the chemical sum composition (CHEMSC), the difference between the experimental and calculated *m/z* values (ERROR), the lipid species name (SPECIES), and its abundance (INTENS). The values filling these columns are obtained from peak attributes of the defined variable, which is in this case `prPE`. The attribute `.mass` contains *m/z* of the identified lipid;

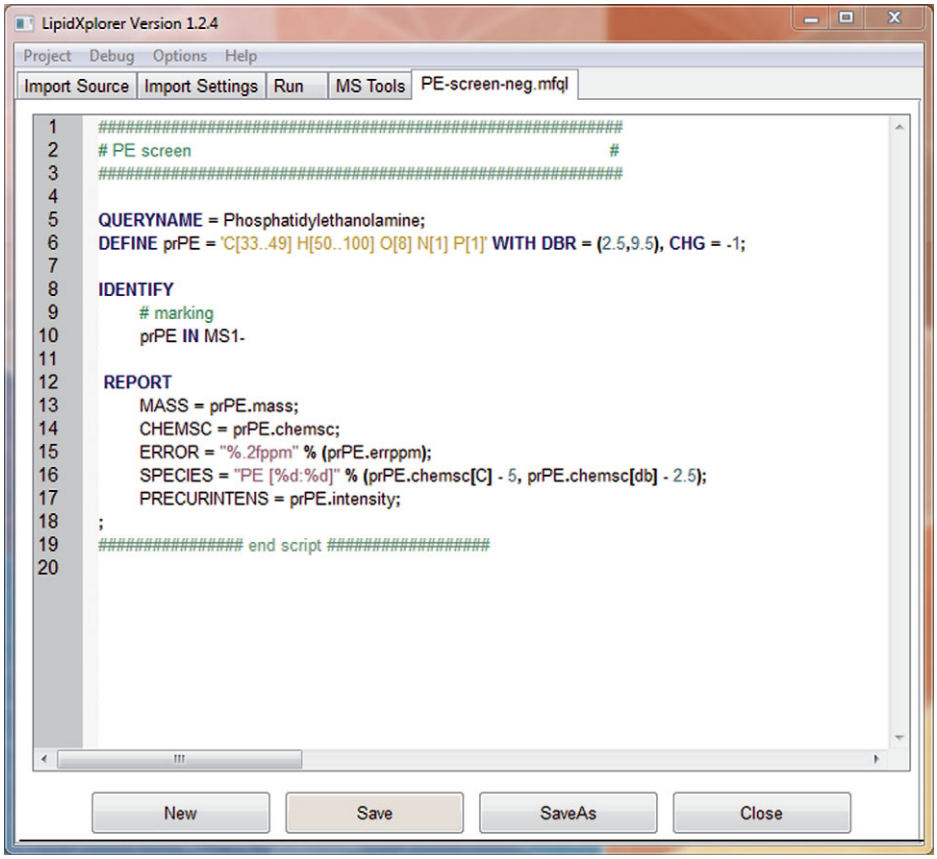


Figure 14.12.8 Screenshot of the MFQL query in the LipidXplorer editor. This query identifies PE anions by their accurate masses and is typically used in top-down lipidomic screens.

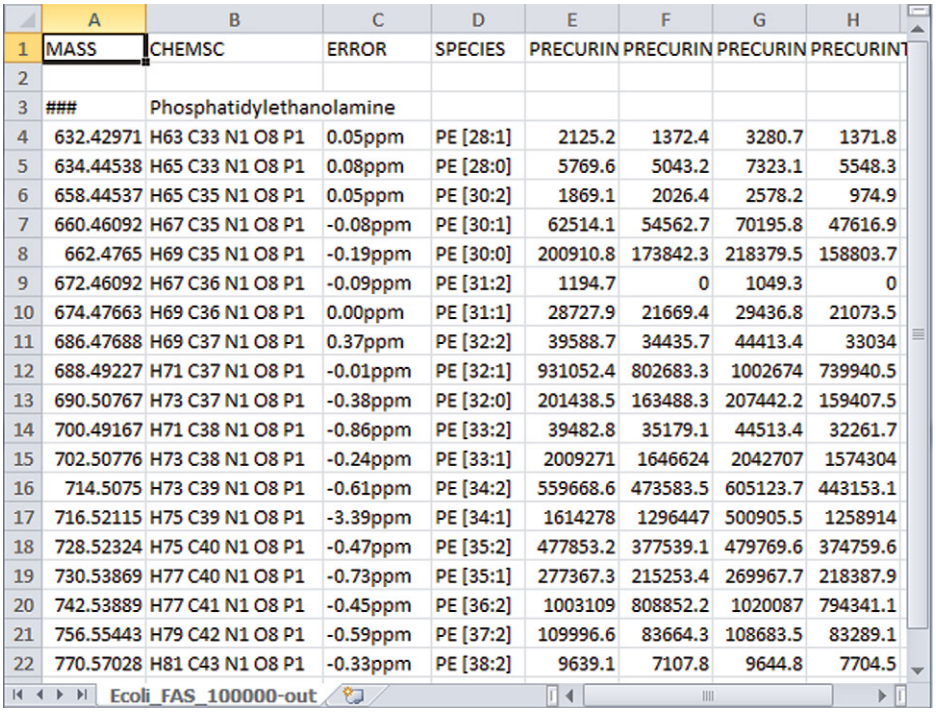
`.chemsc` is the chemical sum composition; `.errppm` is the deviation from the calculated mass (in parts-per-million); and `.intensity` is the intensity of the lipid peak. The name for the lipid species is compiled from values of the chemical sum composition of the lipid species using the operator `%` for formatting strings. It is mainly used for composing lipid species names that typically consist of the lipid class name and the length (as the number of carbon atoms) and unsaturation (as the number of double bonds) of the fatty acid moieties. Its general syntax is:

```
<column name> = "<string with placeholder>" % (<list
of content for the placeholders>);
```

There three types of placeholders: `%f` is for floating point numbers, `%d` for decimal values, and `%s` for strings. The placeholder content can be formatted—for example, the maximum number of decimals `n` for floating point numbers: `%.nf`. Here, to output the mass error in ppm with just one decimal we use:

```
ERROR = "%.1f" % (prPE.errppm);
```

In this protocol, we use this operator for the lipid name and the string `PE [%d:%d]` with two placeholders for decimal values (`%d`). The values for these placeholders are the number of carbon atoms minus five (`prPE.chemsc[C] − 5`) and the number of double bonds minus 2.5 (`prPE.chemsc[db] − 2.5`). Note that the PE head group and glycerol backbone consist of five carbon atoms, and



	A	B	C	D	E	F	G	H
1	MASS	CHEMSC	ERROR	SPECIES	PRECURIN	PRECURIN	PRECURIN	PRECURINT
2								
3	###	Phosphatidylethanolamine						
4	632.42971	H63 C33 N1 O8 P1	0.05ppm	PE [28:1]	2125.2	1372.4	3280.7	1371.8
5	634.44538	H65 C33 N1 O8 P1	0.08ppm	PE [28:0]	5769.6	5043.2	7323.1	5548.3
6	658.44537	H65 C35 N1 O8 P1	0.05ppm	PE [30:2]	1869.1	2026.4	2578.2	974.9
7	660.46092	H67 C35 N1 O8 P1	-0.08ppm	PE [30:1]	62514.1	54562.7	70195.8	47616.9
8	662.4765	H69 C35 N1 O8 P1	-0.19ppm	PE [30:0]	200910.8	173842.3	218379.5	158803.7
9	672.46092	H67 C36 N1 O8 P1	-0.09ppm	PE [31:2]	1194.7	0	1049.3	0
10	674.47663	H69 C36 N1 O8 P1	0.00ppm	PE [31:1]	28727.9	21669.4	29436.8	21073.5
11	686.47688	H69 C37 N1 O8 P1	0.37ppm	PE [32:2]	39588.7	34435.7	44413.4	33034
12	688.49227	H71 C37 N1 O8 P1	-0.01ppm	PE [32:1]	931052.4	802683.3	1002674	739940.5
13	690.50767	H73 C37 N1 O8 P1	-0.38ppm	PE [32:0]	201438.5	163488.3	207442.2	159407.5
14	700.49167	H71 C38 N1 O8 P1	-0.86ppm	PE [33:2]	39482.8	35179.1	44513.4	32261.7
15	702.50776	H73 C38 N1 O8 P1	-0.24ppm	PE [33:1]	2009271	1646624	2042707	1574304
16	714.5075	H73 C39 N1 O8 P1	-0.61ppm	PE [34:2]	559668.6	473583.5	605123.7	443153.1
17	716.52115	H75 C39 N1 O8 P1	-3.39ppm	PE [34:1]	1614278	1296447	500905.5	1258914
18	728.52324	H75 C40 N1 O8 P1	-0.47ppm	PE [35:2]	477853.2	377539.1	479769.6	374759.6
19	730.53869	H77 C40 N1 O8 P1	-0.73ppm	PE [35:1]	277367.3	215253.4	269967.7	218387.9
20	742.53889	H77 C41 N1 O8 P1	-0.45ppm	PE [36:2]	1003109	808852.2	1020087	794341.1
21	756.55443	H79 C42 N1 O8 P1	-0.59ppm	PE [37:2]	109996.6	83664.3	108683.5	83289.1
22	770.57028	H81 C43 N1 O8 P1	-0.33ppm	PE [38:2]	9639.1	7107.8	9644.8	7704.5

Ecoli_FAS_100000-out

Figure 14.12.9 Screenshot of the output file with PE identified by top-down shotgun analysis according to Basic Protocol 3. Since PE were only identified by matching their masses in survey (MS) spectra, fatty acid moieties were not determined, and species are annotated by the lipid class (PE) and total number of carbon atoms and double bonds in both fatty acid moieties (see Figure 14.12.7 for the PE chemical structure). For the subsequent quantification, LipidXplorer reports isotopically corrected intensities of monoisotopic peaks of their molecular anions.

therefore subtracting them from the sum composition gives the total number of carbon atoms in the two fatty-acid moieties. In the same way, subtracting the double-bond equivalent of 2.5 (for two double bonds contained in C=O groups of fatty acid moieties, and 0.5 for the anion form) will give the total number of double bonds in the same moieties (Fig. 14.12.7). In this way, each identified lipid species will be named as PE [`<total number of carbon atoms>: <total number of double bonds>`] (Fig. 14.12.8).

Do not forget to add the semicolon (`;`) at the end of the query.

9. Save the query and close the editor panel.

Identification of PE

10. Go to the Run panel.
11. Click on the green Browse button to upload the MasterScan file produced according to Basic Protocol 2.
12. Make sure that Isotopic Correction MS is checked, and click on the Run LipidXplorer button.
13. Once the identification is completed, click on the View panel and compare the output with the screenshot in Figure 14.12.9.

COMPOSING MFQL QUERIES FOR BOTTOM-UP SHOTGUN LIPIDOMICS

Top-down shotgun analyses identify lipids by their exact masses, while fatty acid moieties in individual molecular species remain unknown. Therefore, the reported lipid abundances may represent the combined abundance of several species sharing the same head group, number of carbon atoms double bonds in both fatty acid moieties. To distinguish individual lipid molecules, precursor ions should be subjected to MS/MS; upon collisional fragmentation, molecular anions of glycerophospholipids (including PE considered here) produce abundant acyl anions of their fatty acid moieties, which, together with precursor *m/z*, unequivocally identify corresponding molecular species.

This protocol describes how to compile an MFQL query for the bottom-up shotgun identification of individual molecular species, which relies upon accurate masses determined in MS (survey) spectra and characteristic fragment masses (here, of acyl anions of fatty acid moieties) recognized in MS/MS spectra acquired from individual precursors.

Necessary Resources

Hardware

PC computer under Microsoft Windows XP, Vista, or 7

Software

LipidXplorer, downloaded and installed as described in Basic Protocol 1

Files

The MasterScan file produced according to Basic Protocol 2 and the MFQL query compiled according to Basic Protocol 3.

1. Start LipidXplorer and turn to the Run panel.
2. Click on “Add MFQL file” to upload the query from Basic Protocol 3.
3. The file name appears in the MFQL text box on the left side. Click on the file name to highlight it, and then click on Edit MFQL Entry to open it with the built-in MFQL editor. A new panel having the same name as the MFQL file will appear. Turn to this panel.
4. Modify this query such that, additionally to matching intact precursor masses in MS spectra, it will also recognize fatty acid anions in MS/MS spectra and use both precursor and fragment ions for identifying individual molecular species. Add the following to the DEFINE section:

```
DEFINE FA1 ='C[12..22] H[20..50] O[2]' WITH DBR =
(1.5,7.5), CHG = −1;
DEFINE FA2 ='C[12..22] H[20..50] O[2]' WITH DBR =
(1.5,7.5), CHG = −1;
```

Here, we defined two new variables covering the full range of possible elemental compositions of expected fatty acid anions, one for each fatty acid moiety. It is presumed that a fatty acid may only have two atoms of oxygen and up to 6 double bonds in its hydrocarbon chain (note that it already has one double bond in its carboxyl group). Since we expect to detect it as a deprotonated molecular anion, we increase the DBR by 0.5.

Then modify the IDENTIFY section:

```
IDENTIFY
  prPE IN MS1- AND
```

```
  FA1 IN MS2- AND
  FA2 IN MS2-
```

Here LipidXplorer is requested to identify the intact precursor ion prPE in negative MS spectra (MS1-). Then, if such a precursor was found, LipidXplorer should further examine the negative-ion mode MS/MS spectrum acquired from this precursor for a pair of fragment ions, whose masses are matching the variables FA1 and FA2. The logical operator AND bundles the matching terms, requesting that all three conditions should be simultaneously met, and if one of the three is missing, LipidXplorer will not consider it as a hit.

5. Add the SUCHTHAT section:

```
SUCHTHAT
  FA1 + FA2 + 'C5 H11 O4 N1 P1'  == prPE
```

Effectively, this requires that the sum composition of both acyl anions together with the sum composition of the head group of PE (C5H11O4N1P1) match the sum composition of the intact ion. Applying the additional compositional constraint drastically reduces the false-positive rate, especially if other acyl anions originating from co-fragmented, yet unrelated lipids, or from chemical background, are detectable in the MS/MS spectrum (Fig. 14.12.10).

6. Modify the REPORT section as follows:

```
REPORT
  MASS = prPE.mass;
```

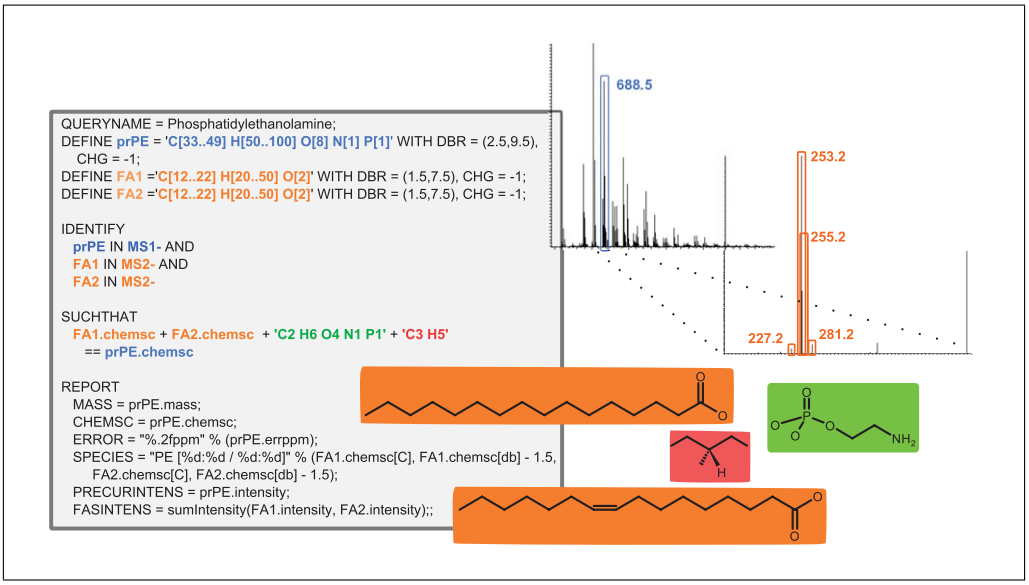


Figure 14.12.10 MFQL query for the identification of molecular species of phosphatidylethanolamines (PE) by bottom-up shotgun lipidomics. The cartoon at the top right shows a survey (MS) spectrum; the precursor ion with *m/z* 688.5 (boxed) was subjected to MS/MS (spectrum below) in which four peaks can be assigned to acyl anions of fatty acids: *m/z* 227.2: FA 14:0; *m/z* 253.2: FA 16:1; *m/z* 255.2: FA 16:0; and *m/z* 281.2: FA 18:1. However, only two pairs of fatty acids: (FA 16:1;FA 16:0) and (FA 14:0; FA18:1) might satisfy the SUCHTHAT requirement (Figure 14.12.11). For clarity, "building blocks" of the PE 16:0/16:1 chemical structure are shown at the bottom right: note that fatty acid moieties are readily detectable as acyl anions, whereas the phosphoethanolamine head group and glycerol backbone are not. Their combined sum composition (C5 H11 O4 N1 P1) is used in the SUCHTHAT section to complement the sum compositions of FA1 and FA2 moieties.

```
  CHEMSC = prPE.chemsc;
  ERROR = "%2.2fppm" % (prPE.errppm);
  SPECIES = "PE [%d:%d / %d:%d]" % (FA1.chemsc[C],
            FA1.chemsc[db] - 1.5,
            FA2.chemsc[C], FA2.chemsc[db] - 1.5);
  INTENS = prPE.intensity;
  FASINTENS = sumIntensity(FA1.intensity,
              FA2.intensity);
;
```

This looks generally similar to the REPORT section in Basic Protocol 3 and contains the same columns—MASS, CHEMSC, ERROR, and INTENS. However, here the MFQL query determines the exact fatty acid composition of individual PE species, and therefore we should also modify the naming convention. We put four placeholders after the lipid class name for the SPECIES column. We then fill the placeholders with the number of carbon atoms and double bonds in each fatty acid moiety, and subtract 1.5 from their DBRs. The column FASINTENS will report the combined abundance of both acyl anions. However, if a lipid comprises the same fatty acid at both positions of the glycerol backbone (as, for example, PE 18:0/18:0), the abundance of 18:0 acyl anion will be erroneously doubled. Therefore, to provide a quantitatively accurate

	A	B	C	D	E	F	G	H	I	J	K	L
1	MASS	CHEMSC	ERROR	SPECIES	PRECURIN	PRECURIN	PRECURIN	PRECURIN	FASINTEN	FASINTEN	FASINTEN	FASINTENS
2												
3	###	Phosphatidylethanolamine										
4	634.4454	H65 C33 N1 O8 P1	0.08ppm	PE [12:0 / 16:0]	5769.6	5043.2	7323.1	5548.3	31	0	16.1	0
5	634.4454	H65 C33 N1 O8 P1	0.08ppm	PE [14:0 / 14:0]	5769.6	5043.2	7323.1	5548.3	24	0	15.4	0
6	660.4609	H67 C35 N1 O8 P1	-0.08ppm	PE [14:0 / 16:1]	62514.1	54562.7	70195.8	47616.9	145.9	185.2	201.6	148.7
7	660.4609	H67 C35 N1 O8 P1	-0.08ppm	PE [14:1 / 16:0]	62514.1	54562.7	70195.8	47616.9	112.9	149.4	168.3	137
8	660.4609	H67 C35 N1 O8 P1	-0.08ppm	PE [18:1 / 12:0]	62514.1	54562.7	70195.8	47616.9	17.6	22.7	20.6	14.7
9	662.4765	H69 C35 N1 O8 P1	-0.19ppm	PE [14:0 / 16:0]	200910.8	173842.3	218379.5	158803.7	1350.4	958	1073.7	881.9
10	674.4766	H69 C36 N1 O8 P1	0.00ppm	PE [17:1 / 14:0]	28727.9	21669.4	29436.8	21073.5	151.6	219.6	197.4	155
11	686.4769	H69 C37 N1 O8 P1	0.37ppm	PE [14:1 / 18:1]	39588.7	34435.7	44413.4	33034	79	59.2	74.1	67.7
12	686.4769	H69 C37 N1 O8 P1	0.37ppm	PE [16:1 / 16:1]	39588.7	34435.7	44413.4	33034	203.2	183.6	203.2	187.7
13	688.4923	H71 C37 N1 O8 P1	-0.01ppm	PE [16:0 / 16:1]	931052.4	802683.3	1002674	739940.5	6619.6	3908.1	5995	4206.7
14	688.4923	H71 C37 N1 O8 P1	-0.01ppm	PE [18:1 / 14:0]	931052.4	802683.3	1002674	739940.5	465.3	254.8	396.5	277.8
15	690.5077	H73 C37 N1 O8 P1	-0.38ppm	PE [16:0 / 16:0]	201438.5	163488.3	207442.2	159407.5	1631.8	1092.1	1201.3	1131.9
16	700.4917	H71 C38 N1 O8 P1	-0.86ppm	PE [17:1 / 16:1]	39482.8	35179.1	44513.4	32261.7	371.2	320.1	335	299.9
17	702.5078	H73 C38 N1 O8 P1	-0.24ppm	PE [19:1 / 14:0]	2009271	1646624	2042707	1574304	29.2	9.5	36.3	22.3
18	702.5078	H73 C38 N1 O8 P1	-0.24ppm	PE [17:1 / 16:0]	2009271	1646624	2042707	1574304	15824.3	9391.9	13926.2	11893.8
19	702.5078	H73 C38 N1 O8 P1	-0.24ppm	PE [16:1 / 17:0]	2009271	1646624	2042707	1574304	20.9	11.7	14.6	20.4
20	714.5075	H73 C39 N1 O8 P1	-0.61ppm	PE [16:1 / 18:1]	559668.6	473583.5	605123.7	443153.1	3800.6	3112.1	3635.2	2824.4
21	714.5075	H73 C39 N1 O8 P1	-0.61ppm	PE [17:1 / 17:1]	559668.6	473583.5	605123.7	443153.1	294.8	234.6	243.6	217.4
22	716.5212	H75 C39 N1 O8 P1	-3.39ppm	PE [18:0 / 16:1]	1614278	1296447	500905.5	1258914	392.3	117.6	293.5	202.3
23	716.5212	H75 C39 N1 O8 P1	-3.39ppm	PE [18:1 / 16:0]	1614278	1296447	500905.5	1258914	14239.2	7636.9	12195	10039.1
24	716.5212	H75 C39 N1 O8 P1	-3.39ppm	PE [17:1 / 17:0]	1614278	1296447	500905.5	1258914	43.4	9.7	41.8	22.3
25	728.5232	H75 C40 N1 O8 P1	-0.47ppm	PE [17:1 / 18:1]	477853.2	377539.1	479769.6	374759.6	4383.1	2673.4	3310.5	2882.8
26	728.5232	H75 C40 N1 O8 P1	-0.47ppm	PE [19:1 / 16:1]	477853.2	377539.1	479769.6	374759.6	219.9	128.2	159	112
27	730.5387	H77 C40 N1 O8 P1	-0.73ppm	PE [16:1 / 19:0]	277367.3	215253.4	269967.7	218387.9	14.7	12.2	25.9	14.4
28	730.5387	H77 C40 N1 O8 P1	-0.73ppm	PE [19:1 / 16:0]	277367.3	215253.4	269967.7	218387.9	3580.4	2176.3	2480.9	2157.7
29	730.5387	H77 C40 N1 O8 P1	-0.73ppm	PE [18:1 / 17:0]	277367.3	215253.4	269967.7	218387.9	43.6	62	52.1	28.8
30	730.5387	H77 C40 N1 O8 P1	-0.73ppm	PE [17:1 / 18:0]	277367.3	215253.4	269967.7	218387.9	147.2	125.2	94.4	59.8
31	742.5389	H77 C41 N1 O8 P1	-0.45ppm	PE [20:1 / 16:1]	1003109	808852.2	1020087	794341.1	36.2	15.2	18.6	17
32	742.5389	H77 C41 N1 O8 P1	-0.45ppm	PE [18:1 / 18:1]	1003109	808852.2	1020087	794341.1	9470	5097.9	7131.6	6455
33	742.5389	H77 C41 N1 O8 P1	-0.45ppm	PE [19:1 / 17:1]	1003109	808852.2	1020087	794341.1	223.8	87.6	165.1	121.7
34	756.5544	H79 C42 N1 O8 P1	-0.59ppm	PE [19:1 / 18:1]	109996.6	83664.3	108683.5	83289.1	1170.9	656.9	709.4	826.7
35	770.5703	H81 C43 N1 O8 P1	-0.33ppm	PE [18:1 / 20:1]	9639.1	7107.8	9644.8	7704.5	119.2	84.2	68.8	83.7
36	770.5703	H81 C43 N1 O8 P1	-0.33ppm	PE [19:1 / 19:1]	9639.1	7107.8	9644.8	7704.5	22	22.5	8.1	17.6

Ecoli_FAS_100000-out

Figure 14.12.11 Screenshot of the output of the identification of PE by bottom-up shotgun analysis according to Basic Protocol 4. Here, individual molecular species of PE were identified by matching the masses of intact precursor in MS survey spectra and acyl anions of fatty acid moieties in MS/MS spectra. On multiple occasions, this analysis identified several PE species sharing the same number of carbon atoms and double bonds, yet composed of different fatty acids. For the subsequent PE quantification, LipidXplorer reported isotopically corrected intensities of monoisotopic peaks of their molecular anions (PRECURINTENS) and the combined abundances of fatty acid fragments (FASINTENS). The naming convention and output format were defined in the MFQL query of Basic Protocol 4.

estimate, we now use the function "sumIntensity". This only sums up the abundances of two different acyl anions and takes no action if the species has the same fatty acid in both moieties.

7. Store the query under the new name using Save As and close the editor panel with Close.
8. Turn to the Run panel.
9. Click on the green Browse button to upload the MasterScan (`Ecoli_FAS_100000.sc`), produced according to Basic Protocol 2. By default, it is located in the folder `Ecoli_FAS_100000`.
10. Make sure that both the "Isotopic correction MS" and "Isotopic correction MS/MS" boxes are checked and click on the Run LipidXplorer button.
11. Once the identification is completed, turn to the View panel and examine the results (Fig. 14.12.11).

COMPOSING MFQL QUERIES FOR LIPID IDENTIFICATION BY NEUTRAL LOSS FRAGMENTS

BASIC PROTOCOL 5

This protocol makes use of several advanced features of MFQL that add extra functionality to compositional constraints within the SUCHTHAT section. As a case study, here we will consider the identification of phosphatidylcholine (PC) molecular species in bovine heart total lipid extract in negative ion mode (Herzog et al., 2011).

Necessary Materials

Hardware

PC computer under Microsoft Windows XP, Vista, or 7

Software

LipidXplorer, downloaded and installed as described in Basic Protocol 1

Files

From the `spectra_article.zip` file (see Basic Protocol 2) choose the folder: `spectra_article\spectra_bovine_heart_negative`

Import the spectra

1. Open LipidXplorer and import the spectra as described in Basic Protocol 2. Click on Browse and choose the folder `<your chosen site>\spectra_bovine_heart_negative\Heart_100000`. Select `*.mzXML` as a file format (here spectra were already converted).
2. Turn to the Import Settings panel and load the same settings as in Basic Protocol 2: Click on Browse and load `lpdxImportSettings_tutorial.ini` file from your LipidXplorer folder. Select the configuration file `FAS_LTQ_100000MS1`.
3. Click the "Start import" button.

Compose the MFQL query

4. Add to the DEFINE section:

```
QUERYNAME = Phosphatidylcholine;
DEFINE PR = 'C[38..54] H[30..130] O[10] N[1]
    P[1]' WITH DBR = (2.5,9.5), CHG = −1;
DEFINE FA1 = 'C[14..22] H[20..50] O[2]' WITH DBR =
    (1.5,7.5), CHG = −1;
```

```
DEFINE FA1 = 'C[14..22] H[20..50] O[2]' WITH DBR =
    (1.5,7.5), CHG = −1;
```

The defined variables are almost the same as those employed for the PE identification in Basic Protocols 3 and 4 because, being structurally different, PC and PE molecules share the same elemental composition constraints. However, in negative mode, PC are detected as adducts with acetate or formate anions, whereas PE are detected as molecular (deprotonated) anions. Therefore, anions of PC precursors contain two more oxygen atoms (10 atoms in total).

5. Add another variable to the DEFINE section:

```
DEFINE headPC = 'C[3] H[6] O[2]' WITH CHG = 0;
```

Upon collisional fragmentation, acetate adducts of PC undergo neutral loss of the acetate and, additionally, of the methyl group: we defined the lost part here as the variable `headPC` having zero charge (`CHG=0`). The negative charge remains at the PC backbone, yielding abundant anion fragment. Its *m/z* and composition variy because, in different PC species, it comprises different fatty acid moieties.

6. Modify the IDENTIFY section as follows:

```
IDENTIFY
    PR IN MS1- AND
    FA1 IN MS2- AND
    FA2 IN MS2- AND
    headPC IN MS2-
```

7. Modify the `SUCHTHAT` section:

```
SUCHTHAT
    isEven(PR.chemsc[C]) AND
    isEven(FA1.chemsc[C]) AND
    isEven(FA2.chemsc[C]) AND
    FA1 + FA2 + headPC + 'C7 H15 P1 O4 N1'== PR
```

Terms in the SUCHTHAT section are bundled with *Boolean* operators. Here we only use the operator AND, to request that all four terms be simultaneously met. Another allowed operator is OR, which requests that at least one of several terms is true. The first three terms additionally request that the number of carbon atoms in the precursor, and both fatty acid moieties, be even. It is generally assumed that mammals do not produce fatty acids with an odd number of carbon atoms and, optionally, we can exclude these PC species from this search. The last term requires that the sum compositions of the fatty acid moieties, head group, and neutral loss complement each other to the expected composition of the intact anion of acetate adduct of PC.

8. Modify the REPORT section:

```
REPORT
  MASS = "%4.4f" % (PR.mass);
  CHEMSC = PR.chemsc;
  ERROR = "%2.2fppm" % (PR.errppm);
  SPECIE = "PC [%d:%d / %d:%d]" %
    (FA1.chemsc[C], FA1.chemsc[db] - 1.5,
    FA2.chemsc[C], FA2.chemsc[db] - 1.5);
  PRECURINTENS = PR.intensity;
  NLINTENS = headPC.intensity;
```

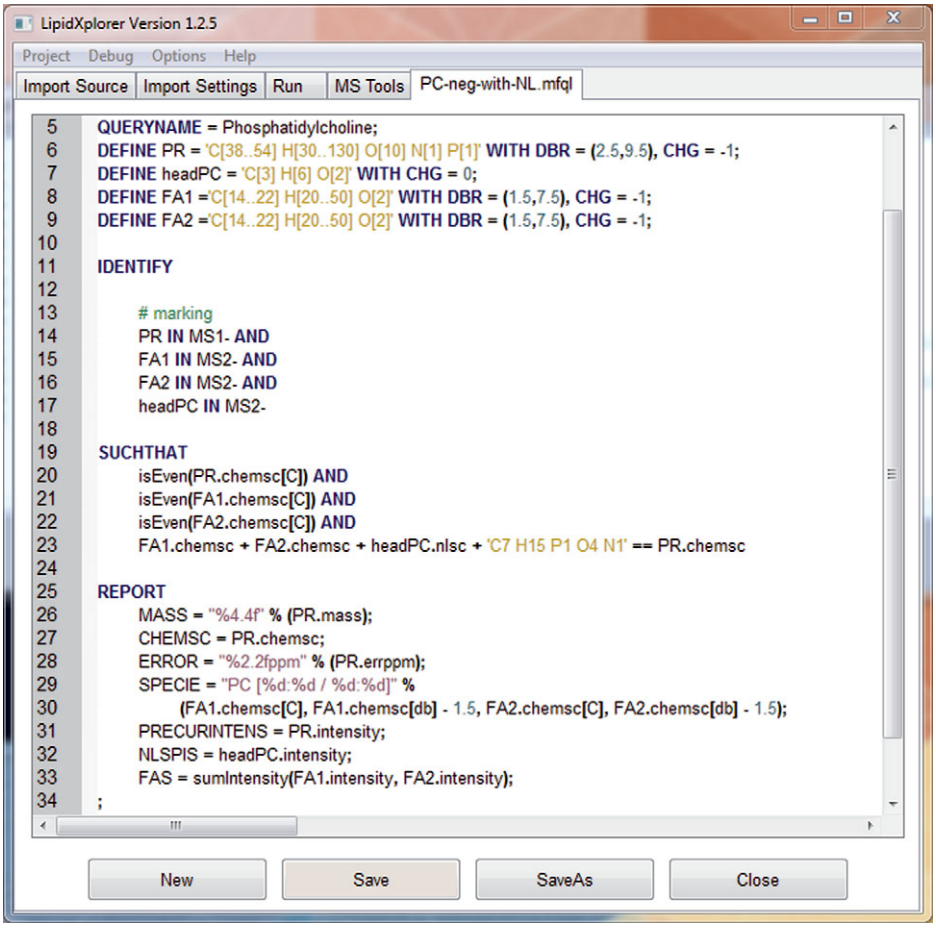
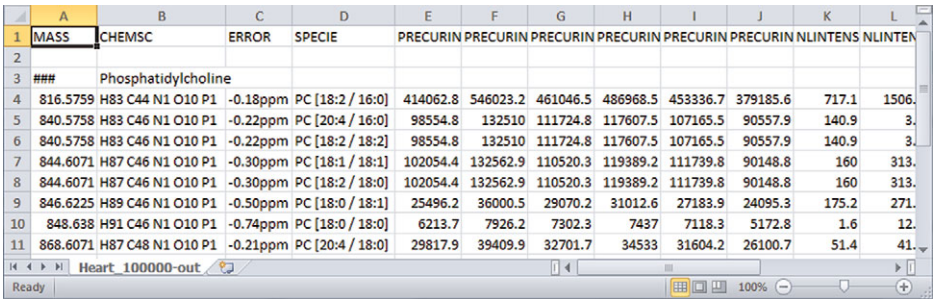


Figure 14.12.12 Screenshot of the MFQL editor panel with the query described in Basic Protocol 5 for bottom-up shotgun identification of PC species in negative ion mode.



	A	B	C	D	E	F	G	H	I	J	K	L
1	MASS	CHEMSC	ERROR	SPECIE	PRECURIN	PRECURIN	PRECURIN	PRECURIN	PRECURIN	PRECURIN	NLINTENS	NLINTEN
2												
3	###	PhosphatidylCholine										
4	816.5759	H83 C44 N1 O10 P1	-0.18ppm	PC [18:2 / 16:0]	414062.8	546023.2	461046.5	486968.5	453336.7	379185.6	717.1	1506.
5	840.5758	H83 C46 N1 O10 P1	-0.22ppm	PC [20:4 / 16:0]	98554.8	132510	111724.8	117607.5	107165.5	90557.9	140.9	3.
6	840.5758	H83 C46 N1 O10 P1	-0.22ppm	PC [18:2 / 18:2]	98554.8	132510	111724.8	117607.5	107165.5	90557.9	140.9	3.
7	844.6071	H87 C46 N1 O10 P1	-0.30ppm	PC [18:1 / 18:1]	102054.4	132562.9	110520.3	119389.2	111739.8	90148.8	160	313.
8	844.6071	H87 C46 N1 O10 P1	-0.30ppm	PC [18:2 / 18:0]	102054.4	132562.9	110520.3	119389.2	111739.8	90148.8	160	313.
9	846.6225	H89 C46 N1 O10 P1	-0.50ppm	PC [18:0 / 18:1]	25496.2	36000.5	29070.2	31012.6	27183.9	24095.3	175.2	271.
10	848.638	H91 C46 N1 O10 P1	-0.74ppm	PC [18:0 / 18:0]	6213.7	7926.2	7302.3	7437	7118.3	5172.8	1.6	12.
11	868.6071	H87 C48 N1 O10 P1	-0.21ppm	PC [20:4 / 18:0]	29817.9	39409.9	32701.7	34533	31604.2	26100.7	51.4	41.

Heart_100000-out
Ready
100%

Figure 14.12.13 Screenshot of the results from Basic Protocol 5. Note that MS/MS identified molecular species of PC. Here, isotopically corrected intensities of monoisotopic precursor peaks from MS spectra and intensities of peaks of neutral loss products in MS/MS spectra are reported for lipid quantification.

```
  FAS = sumIntensity(FA1.intensity, FA2.intensity);
;
```

Note that the only difference compared with the REPORT section described in Basic Protocol 4 is the additional column NLINTENS, which reports the intensity of the neutral loss product.

9. Click on Save to store the query (Fig. 14.12.12).
10. Close the editor and turn to the Run panel. Make sure that the Isotopic Corrections are selected and start the identification by Run LipidXplorer button.
11. Once the identification is completed, turn to the View panel and examine the output (Fig. 14.12.13).

BASIC PROTOCOL 6

TROUBLESHOOTING LIPID IDENTIFICATIONS

The LipidXplorer software helps the users to identify errors in importing mass spectra or compiling MFQLs. This protocol describes a basic troubleshooting procedure using the *dump* and *MS Tools* functions.

Necessary Resources

Hardware

PC computer under Microsoft Windows XP, Vista, or 7

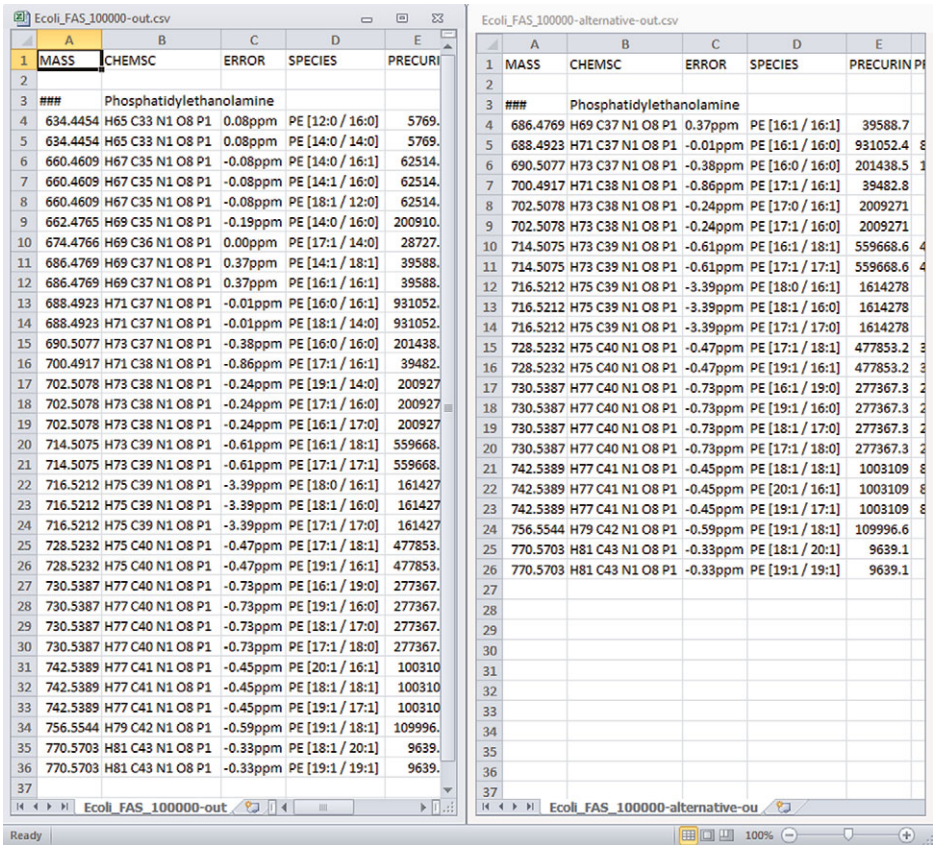
Software

LipidXplorer, downloaded and installed as described in Basic Protocol 1

Files

The PE identification query from Basic Protocol 4; the spectra import settings file from *https://wiki.mpi-cbg.de/lipidx/File:LpdxImportSettings_alternative.ini*; the spectra are the same as in Basic Protocol 2.

1. Download the alternative import settings file and store it under the LipidXplorer directory. Use this file for training purpose only!
2. Import the spectra as described in Basic Protocol 1, however use the alternative import settings file: in the Import Panel, click on Browse and open the file `lpdxImportSettings_alternative.ini`. Choose the configuration `FAS_LTQ_100000MS1_alternative`.
3. Turn to the Import Source panel and rename the MasterScan file to keep the already existing file. Change its name from `Ecoli_FAS_100000.sc` to `Ecoli_FAS_100000-alternative.sc`.
4. Turn back to the Import Settings and click on “Start import”.
5. Once spectra import is completed, turn to the Run panel. Use the `PE.mfql` query from the LipidXplorer `mfql` directory or the query composed according Basic Protocol 4. To identify PE species, load the query into LipidXplorer by clicking on “Add MFQL file” or by dragging and dropping it into the text list window at the left. Make sure that both isotopic corrections (Isotopic Correction MS and Isotopic Correction MS/MS) are selected.
6. Rename the current output file to keep the file previously created according to Basic Protocol 4. Click on the text field under “Specify output file” and change the file name at the end of the line from `Ecoli_FAS_100000-out.csv` to `Ecoli_FAS_100000-alternative-out.csv`. Note that by default all files generated by LipidXplorer are stored in the same folder with the spectra. Filenames are composed of the folder name and a suffix: `.sc` for the MasterScan or `-out.csv` for the output file.
7. Start the lipid identification by clicking on Run LipidXplorer.



Ecoli_FAS_100000-out.csv

	A	B	C	D	E
1	MASS	CHEMSC	ERROR	SPECIES	PRECURI
2					
3	###	Phosphatidylethanolamine			
4	634.4454	H65 C33 N1 O8 P1	0.08ppm	PE [12:0 / 16:0]	5769.
5	634.4454	H65 C33 N1 O8 P1	0.08ppm	PE [14:0 / 14:0]	5769.
6	660.4609	H67 C35 N1 O8 P1	-0.08ppm	PE [14:0 / 16:1]	62514.
7	660.4609	H67 C35 N1 O8 P1	-0.08ppm	PE [14:1 / 16:0]	62514.
8	660.4609	H67 C35 N1 O8 P1	-0.08ppm	PE [18:1 / 12:0]	62514.
9	662.4765	H69 C35 N1 O8 P1	-0.19ppm	PE [14:0 / 16:0]	200910.
10	674.4766	H69 C36 N1 O8 P1	0.00ppm	PE [17:1 / 14:0]	28727.
11	686.4769	H69 C37 N1 O8 P1	0.37ppm	PE [14:1 / 18:1]	39588.
12	686.4769	H69 C37 N1 O8 P1	0.37ppm	PE [16:1 / 16:1]	39588.
13	688.4923	H71 C37 N1 O8 P1	-0.01ppm	PE [16:0 / 16:1]	931052.
14	688.4923	H71 C37 N1 O8 P1	-0.01ppm	PE [18:1 / 14:0]	931052.
15	690.5077	H73 C37 N1 O8 P1	-0.38ppm	PE [16:0 / 16:0]	201438.
16	700.4917	H71 C38 N1 O8 P1	-0.86ppm	PE [17:1 / 16:1]	39482.
17	702.5078	H73 C38 N1 O8 P1	-0.24ppm	PE [19:1 / 14:0]	200927
18	702.5078	H73 C38 N1 O8 P1	-0.24ppm	PE [17:1 / 16:0]	200927
19	702.5078	H73 C38 N1 O8 P1	-0.24ppm	PE [16:1 / 17:0]	200927
20	714.5075	H73 C39 N1 O8 P1	-0.61ppm	PE [16:1 / 18:1]	559668.
21	714.5075	H73 C39 N1 O8 P1	-0.61ppm	PE [17:1 / 17:1]	559668.
22	716.5212	H75 C39 N1 O8 P1	-3.39ppm	PE [18:0 / 16:1]	161427
23	716.5212	H75 C39 N1 O8 P1	-3.39ppm	PE [18:1 / 16:0]	161427
24	716.5212	H75 C39 N1 O8 P1	-3.39ppm	PE [17:1 / 17:0]	161427
25	728.5232	H75 C40 N1 O8 P1	-0.47ppm	PE [17:1 / 18:1]	477853.
26	728.5232	H75 C40 N1 O8 P1	-0.47ppm	PE [19:1 / 16:1]	477853.
27	730.5387	H77 C40 N1 O8 P1	-0.73ppm	PE [16:1 / 19:0]	277367.
28	730.5387	H77 C40 N1 O8 P1	-0.73ppm	PE [19:1 / 16:0]	277367.
29	730.5387	H77 C40 N1 O8 P1	-0.73ppm	PE [18:1 / 17:0]	277367.
30	730.5387	H77 C40 N1 O8 P1	-0.73ppm	PE [17:1 / 18:0]	277367.
31	742.5389	H77 C41 N1 O8 P1	-0.45ppm	PE [20:1 / 16:1]	100310
32	742.5389	H77 C41 N1 O8 P1	-0.45ppm	PE [18:1 / 18:1]	100310
33	742.5389	H77 C41 N1 O8 P1	-0.45ppm	PE [19:1 / 17:1]	100310
34	756.5544	H79 C42 N1 O8 P1	-0.59ppm	PE [19:1 / 18:1]	109996.
35	770.5703	H81 C43 N1 O8 P1	-0.33ppm	PE [18:1 / 20:1]	9639.
36	770.5703	H81 C43 N1 O8 P1	-0.33ppm	PE [19:1 / 19:1]	9639.
37					

Ecoli_FAS_100000-out

Ecoli_FAS_100000-alternative-out.csv

	A	B	C	D	E
1	MASS	CHEMSC	ERROR	SPECIES	PRECURIN
2					
3	###	Phosphatidylethanolamine			
4	686.4769	H69 C37 N1 O8 P1	0.37ppm	PE [16:1 / 16:1]	39588.7
5	688.4923	H71 C37 N1 O8 P1	-0.01ppm	PE [16:1 / 16:0]	931052.4
6	690.5077	H73 C37 N1 O8 P1	-0.38ppm	PE [16:0 / 16:0]	201438.5
7	700.4917	H71 C38 N1 O8 P1	-0.86ppm	PE [17:1 / 16:1]	39482.8
8	702.5078	H73 C38 N1 O8 P1	-0.24ppm	PE [17:0 / 16:1]	2009271
9	702.5078	H73 C38 N1 O8 P1	-0.24ppm	PE [17:1 / 16:0]	2009271
10	714.5075	H73 C39 N1 O8 P1	-0.61ppm	PE [16:1 / 18:1]	559668.6
11	714.5075	H73 C39 N1 O8 P1	-0.61ppm	PE [17:1 / 17:1]	559668.6
12	716.5212	H75 C39 N1 O8 P1	-3.39ppm	PE [18:0 / 16:1]	1614278
13	716.5212	H75 C39 N1 O8 P1	-3.39ppm	PE [18:1 / 16:0]	1614278
14	716.5212	H75 C39 N1 O8 P1	-3.39ppm	PE [17:1 / 17:0]	1614278
15	728.5232	H75 C40 N1 O8 P1	-0.47ppm	PE [17:1 / 18:1]	477853.2
16	728.5232	H75 C40 N1 O8 P1	-0.47ppm	PE [19:1 / 16:1]	477853.2
17	730.5387	H77 C40 N1 O8 P1	-0.73ppm	PE [16:1 / 19:0]	277367.3
18	730.5387	H77 C40 N1 O8 P1	-0.73ppm	PE [19:1 / 16:0]	277367.3
19	730.5387	H77 C40 N1 O8 P1	-0.73ppm	PE [18:1 / 17:0]	277367.3
20	730.5387	H77 C40 N1 O8 P1	-0.73ppm	PE [17:1 / 18:0]	277367.3
21	742.5389	H77 C41 N1 O8 P1	-0.45ppm	PE [18:1 / 18:1]	1003109
22	742.5389	H77 C41 N1 O8 P1	-0.45ppm	PE [20:1 / 16:1]	1003109
23	742.5389	H77 C41 N1 O8 P1	-0.45ppm	PE [19:1 / 17:1]	1003109
24	756.5544	H79 C42 N1 O8 P1	-0.59ppm	PE [19:1 / 18:1]	109996.6
25	770.5703	H81 C43 N1 O8 P1	-0.33ppm	PE [18:1 / 20:1]	9639.1
26	770.5703	H81 C43 N1 O8 P1	-0.33ppm	PE [19:1 / 19:1]	9639.1
27					
28					
29					
30					
31					
32					
33					
34					
35					
36					
37					

Ecoli_FAS_100000-alternative-ou

Ready

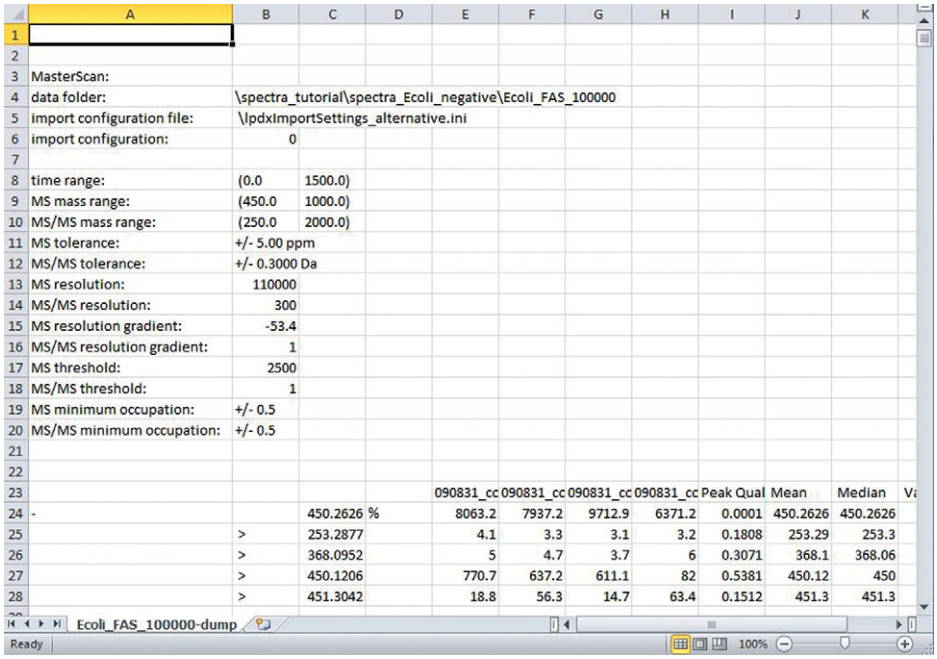
100%

Figure 14.12.14 Comparison of the output files produced according to Basic Protocol 4 (panel at the left) and Basic Protocol 6 (panel at the right): note that ten lipids are now missing. To locate and fix the problem, follow Basic Protocol 6.

8. Open the result file (Fig. 14.12.14) and compare it with the result obtained using Basic Protocol 3. Even a brief examination shows that a few lipids (for example, all species with *m/z* 660.4609) are now missing!

9. For troubleshooting, LipidXplorer offers *Dump MasterScan*, a basic function for examining the MasterScan content (note that it is a binary file and is unreadable by other software). In the Run panel, check the box Dump MasterScan. Now, together with identifying lipids, LipidXplorer will dump the MasterScan content into the file `<folder name>-dump.csv` located in the same folder with the imported spectra. Click Run LipidXplorer to proceed with identifying lipids and dumping the MasterScan. Note that you can also use this function without selecting a MFQL query, i.e., without making identifications.

10. Open the dump file by clicking on "View dump file". It can be also opened with any spreadsheet software, e.g., Microsoft Excel or Open Office. Its header reports the import settings (Fig. 14.12.15), which is followed by rows showing the MasterScan content (Fig. 14.12.16A), i.e., peaks imported by LipidXplorer from raw spectra. The first column defines the acquisition polarity. It remains void if spectra were acquired in positive mode, while the "-" symbol indicates negative mode. The second column indicates if the peak was imported from the MS spectrum (void) or MS/MS spectrum (">" symbol). The precursor peak from which the MS/MS spectrum was acquired is always the MS peak above the first MS/MS entry (Fig. 14.12.16A). Other columns contain peak abundances and a few useful statistical estimates. The last column



	A	B	C	D	E	F	G	H	I	J	K
1											
2											
3	MasterScan:										
4	data folder:	\spectra_tutorial\spectra_Ecoli_negative\Ecoli_FAS_100000									
5	import configuration file:	\lpdxImportSettings_alternative.ini									
6	import configuration:	0									
7											
8	time range:	(0.0	1500.0)								
9	MS mass range:	(450.0	1000.0)								
10	MS/MS mass range:	(250.0	2000.0)								
11	MS tolerance:	+/- 5.00 ppm									
12	MS/MS tolerance:	+/- 0.3000 Da									
13	MS resolution:	110000									
14	MS/MS resolution:	300									
15	MS resolution gradient:	-53.4									
16	MS/MS resolution gradient:	1									
17	MS threshold:	2500									
18	MS/MS threshold:	1									
19	MS minimum occupation:	+/- 0.5									
20	MS/MS minimum occupation:	+/- 0.5									
21											
22											
23					090831_cc	090831_cc	090831_cc	090831_cc	Peak Qual	Mean	Median
24	-		450.2626	%	8063.2	7937.2	9712.9	6371.2	0.0001	450.2626	450.2626
25		>	253.2877		4.1	3.3	3.1	3.2	0.1808	253.29	253.3
26		>	368.0952		5	4.7	3.7	6	0.3071	368.1	368.06
27		>	450.1206		770.7	637.2	611.1	82	0.5381	450.12	450
28		>	451.3042		18.8	56.3	14.7	63.4	0.1512	451.3	451.3

Ecoli_FAS_100000-dump

Ready 100%

Figure 14.12.15 Screenshot of a MasterScan dump: its header contains the import settings. The content of the MasterScan is shown in Figure 14.12.16.

contains labels placed by the MFQL interpreter that associate the identified lipid with variables defined in the query. A label consists of the query name, the variable name, an integer *m/z* value, type of spectrum (MS or MS/MS), and the sum composition computed for the matched peak. We further show how this information can be used for debugging lipid identifications.

11. Now, check the peak with *m/z* 660.4609, which is missing in the output (Fig. 14.12.16A). The intact precursor ion was correctly tagged as PE (highlighted in blue), i.e., the *m/z* 660.4609 is consistent with the *sc*-constraints defined for prPE and would stand for PE 30:1. Further, three fatty acids (FA 16:0, 16:1 and 18:1) were correctly tagged (highlighted in orange) in the corresponding fragment spectrum by the variables FA1 and FA2, yet they did not lead to PE identifications. Indeed, alternative fatty acid compositions compatible with PE 30:1, such as PE 14:0/16:1, PE 14:1/16:0, or PE 18:1/12:0, were not reported.

12. We now turn to the panel MS Tools and input the elemental composition of the fatty acid 14:1 (`H27 C14 O2`) into the text field under "sc-constraint or sum composition" and hit "Sum-composition-to-m/z" (Fig. 14.12.17). The MS Tools panel offers a few assisting functions, which use the same algorithms as the internal functions of LipidXplorer for calculating the *m/z*, elemental composition, or isotopic distribution of precursor and fragment ions.

13. Taking another look into the dump file reveals that the MS/MS spectrum of *m/z* 660.4609 does not contain peaks below *m/z* 253.3 and which may correspond to shorter fatty acid moieties, like FA 14:1 or 12:0. Why were these peaks not imported into the MasterScan?

14. Turn to the Import panel and check the import settings. A closer look reveals that "*m/z* range" of the MS/MS spectra was set to *m/z* `250` to `2000`, and therefore acyl anions of shorter fatty acids were disregarded at the import stage. Let us adjust the

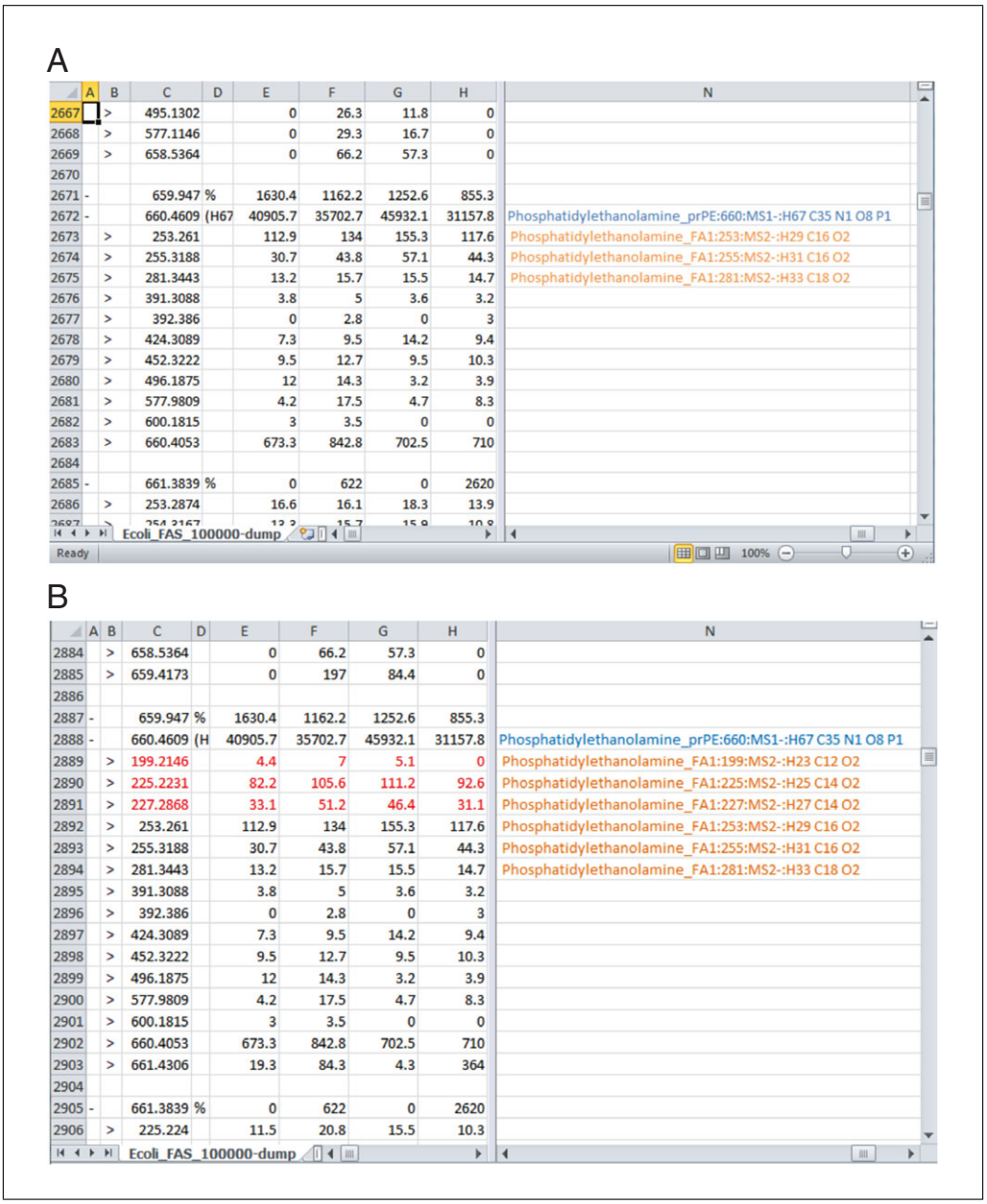


Figure 14.12.16 Screenshot of the main part MasterScan dump (screenshots **A** and **B**). Column A reports the acquisition polarity, which is negative in this case. Column B indicates if the peak is taken from MS survey spectrum (and is a precursor) or from MS/MS spectrum (and is a fragment), in which case the column contains a > symbol. For MS/MS, the precursor peak is always placed in the row above the first fragment peak. For example, the peak with *m*/*z* 659.947 was not subjected to MS/MS; the next peak with *m*/*z* 660.4609 was fragmented and produced ions with *m*/*z* 253.261, 255.3188, etc. (in **A**). Column D shows the computed sum composition of the precursor, while further columns (E to H) show the intensities of the corresponding precursors and fragment peaks in each acquired sample. The last columns (N and the following) contain labels produced by the MFQL identification process. A label consists of the query name, the variable name, an integer *m*/*z* value, the level of the spectra, and the sum composition assigned for this peak. The upper screenshot (**A**) shows the precursor with *m*/*z* 660.46 of the alternative import. Its fragment spectrum starts with the peak of *m*/*z* 253.261 and ends with the peak of *m*/*z* 660.405, because the fatty acids with *m*/*z* smaller than 253.3 are missing. The bottom screenshot (**B**) shows the MasterScan imported with the correct settings. The MS/MS spectrum starts with the peak *m*/*z* 199.215 which represents the missing fatty acid fragments (see Basic Protocol 6).

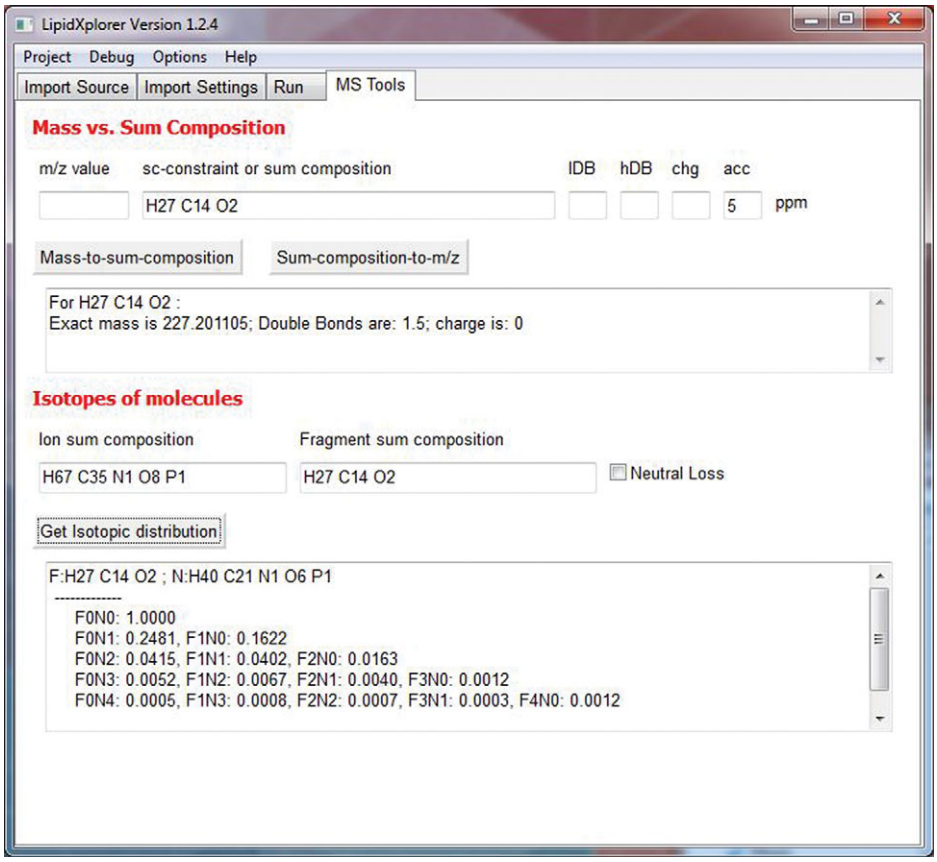


Figure 14.12.17 Screenshot of the MS Tools panel. It offers several supporting functions for calculating *m/z*, sum compositions and isotopic distributions for precursors and fragment ions (see Ejsing et al., 2006; Herzog et al., 2011, for the algorithm employed). The lower panel shows the example of an isotopic distribution.

settings by changing the start of "*m/z* range" to *m/z* `50` and re-import the spectra by selecting "Start import".

15. Turn to the "Run" panel and click on "Run LipidXplorer" to redo the identification. Turn on the option "Dump MasterScan".

16. Once the identification has been completed, open the results file by clicking on View. The PE lipid species with the *m/z* 660.4609 are now back in the results.

17. Open the dump file by clicking on "View dump file" and scroll down to the peak *m/z* 660.4609. The missed fragment peaks are now in the MS/MS spectrum and are also correctly tagged by LipidXplorer (Fig. 14.12.16B).

GUIDELINES FOR UNDERSTANDING RESULTS

As exemplified in Basic Protocol 3, LipidXplorer outputs the results file in `*.csv` format, and its content is user defined. It may include common descriptors such as names of lipid species, their sum compositions, *m/z* errors, etc., that are computed for each identified species and formatted in column layout. For the lipid quantification, users may output intensities of peaks of lipid precursors or their fragments, such as `PR.intensity` or `headPC.intensity` (see Basic Protocol 5 and Figs. 14.12.9, 14.12.11, and 14.12.13 for details). They are reported individually for each species in each sample and formatted as columns under a common filename header. LipidXplorer only reports the isotopically

corrected abundances of precursors and fragment ions, but not lipid concentrations: in lipidomics, quantification routines are diverse and project dependent. They make use of normalization factors and internal standards that might be chosen individually for each experiment. In high-throughput screens, intensities of precursor ions are directly output into a multivariate analysis software, bypassing the calculation of individual species abundances.

In the course of shotgun analysis, the composition of a sprayed analyte does not change with time. Therefore, the most common way to calculate the concentrations of lipid species is to compare the intensities of reporter peaks from quantified lipids and from the internal standard(s). At least one internal standard per analyzed lipid class should be spiked into the sample prior to lipid extraction. This method, however, requires several quality checks (e.g., linearity, dynamic range, and peak pattern consistency) to establish that the abundance of the internal standard adequately reflects the abundances of other species of the same lipid class, and this holds true for the entire range of concentrations of quantified lipids and total lipid content of analyzed samples. Also, blank spectra should be acquired prior to analyses to identify background peaks that might jeopardize the quantification by overlapping with the peaks of quantified lipids.

Is it possible to check if the quantification method is correct? LipidXplorer offers several means for independent validation of the internal consistency of quantitative lipid profiles. Usually, lipids can be quantified in multiple ways: for example, by considering the abundances of intact precursors in MS spectra or the abundances of different fragments or their combinations in MS/MS spectra, especially if spectra were acquired in both positive and negative modes (Schuhmann et al., 2012). For example, in Basic Protocol 5 PC species could be quantified in negative mode by considering:

```
PRECURINTENS = PR.intensity;
NLINTENS = headPC.intensity;
```

and

```
FAS = sumIntensity(FA1.intensity, FA2.intensity).
```

Although the quantification relies on the abundances of different reporter ions, if the analysis was carried out correctly the three interpretations should yield coherent PC species profiles. We underscore that there is no need to re-acquire spectra: all interpretations use the same spectra dataset organized in the same MasterScan file that is probed by the same MFQL query.

Reproducible technical and biological replicates help to ensure the sound interpretation of spectra dataset. No software (including LipidXplorer) will circumvent the wisdom of the familiar “garbage in–garbage out” rule.

The above refers to the internal consistency of profiles depicting relative abundances of individual lipid species of the same class. To reach consistency between independent experiments, lipid abundances should be adjusted for differences in the total amount of analyzed biological material. Total protein content, number of cells, DNA content, and organic phosphate content are among the most commonly used normalization factors. If the analysis provides absolute lipid quantities (in moles) and covers all major classes of the lipidome, it is usually convenient to normalize the abundances of lipid species or lipid classes to the total lipid content and report them in molar percentage (%mol).

To identify compositional differences between complex lipidomes, the data are further processed by the methods of multivariate statistics, such as Principal Component Analysis (PCA), Partial Least Squares Discriminant Analysis (PLS-DA), Cluster Analysis, and ANOVA, to mention just a few.

COMMENTARY

Background Information

Utility, scope, and limitations of LipidXplorer

LipidXplorer supports the identification and quantitation of species of any lipid class by any method of shotgun analysis on any mass spectrometer (Herzog et al., 2011). Although it was not discussed in this unit, LipidXplorer can also interpret precursor and neutral loss scanning spectra acquired on a triple quadrupole mass spectrometers in batch mode (Herzog et al., 2012), which is a basic methodology of shotgun lipidomics analysis (reviewed in Han and Gross, 2005; Han et al., 2012). LipidXplorer has already been applied to a variety of projects—from high-throughput clinical lipidomics screens (Graessler et al., 2009; Fernandez et al., 2011) to full-lipidome characterization of eukaryotic cells (Sampaio et al., 2011) or the entire model organism (Carvalho et al., 2012). A regularly updated application bibliography together with corresponding MFQL queries are provided at the LipidXplorer wiki site (*https://wiki.mpi-cbg.de/lipidx/*). The important feature of LipidXplorer is that its functionality is not limited to identifying lipids: by processing the entire dataset of shotgun spectra, it also reports data for lipid quantification.

What are the known limitations of LipidXplorer? In its current realization, it only supports shotgun lipidomics. Although some workaround solutions are possible, its core functionality does not support LC-MS or LC-MS/MS experiments. Also, LipidXplorer is sometime criticized for not providing lipid quantities directly in some convenient units, like moles or moles per liter. This, however, was done on purpose: we aimed at maintaining its ultimate flexibility, especially since software solutions for data processing and visualization are now becoming available (Klose et al., 2012).

However, in our experience, the most common problems in shotgun analyses are not directly related to the software. Let us briefly discuss them below.

Critical Parameters and Troubleshooting

Common problems with shotgun lipidomics can be conveniently sorted into two large categories: first, problems related to mass spectrometric analyses and processing of spectra; second, problems related to lipid recovery from biological matrices.

In our experience, most common errors in lipid identification arise from inaccurate selection of the import settings of LipidXplorer. Specifically, the mass resolution and intensity thresholds may strongly differ even between mass spectrometers of the same type, while inexperienced users often tend to select some arbitrary settings. Educated choice of processing settings requires hands-on analytical experience in shotgun profiling. In the appendix, we provide a detailed description of individual import settings. We further invite interested users to join the LipidXplorer users discussion group at: *http://groups.google.com/group/lipidxplorer-discussion-group* and seek advice of colleagues experienced with running a particular type of mass spectrometer. LipidXplorer wiki site also provides bibliography that might help in overcoming technical hurdles.

Note that the identification of lipids (in contrast to the identification of proteins by fragmenting peptides) usually relies upon matching of very limited number of fragment ions, and some of them (like, acyl anions of fatty acids) might originate from totally unrelated lipid species or even from the chemical background. Therefore, the most confident identification of molecular species relies upon a combination of species-specific and lipid class–specific fragments and greatly benefits from the high mass resolution and accuracy of modern instruments (reviewed in Ejsing et al., 2006; Schwudke et al., 2006, 2007b; Herzog et al., 2011). Note that LipidXplorer was specifically designed for taking advantage of high-resolution spectra.

Hybrid tandem mass spectrometers might offer several independent means of collisional activation of lipid precursors. For example, in hybrid linear ion trap (Orbitrap) instruments, CID (collision-induced dissociation) and HCD (higher energy collisional dissociation) at the ion trap and at the separate linear collision cell, respectively, yield complementary fragments (Schuhmann et al., 2011). LipidXplorer takes full advantage of the flexible experiment design: it does not rely on reference spectra databases, and the open format of MFQL helps to accommodate any fragmentation pathway in a MasterScan search query, hence supporting truly multifaceted lipid identification.

Another class of common lipid identification problems is related to lipid recovery: indeed, a mass spectrometer can only ionize and detect, and a software application (including LipidXplorer) can only

identify molecules that were recovered from a biological material in sufficient quantities and with low chemical noise. Lipids are usually recovered by a seemingly straightforward liquid-liquid extraction (Folch et al., 1957; Bligh and Dyer, 1959; Matyash et al., 2008). However, in our experience the quality of chemicals and glass- and plasticware, the pretreatment of samples for removing buffer components and salts, and the selection of appropriate controls and normalization factors determine the success of the entire analysis. Note that certain lipids (notably glycolipids) are poorly recovered by classical Folch or Bligh and Dyer methods and require a more sophisticated multistage extraction (Ejsing et al., 2009).

Literature Cited

Bligh, E.G. and Dyer, W.J. 1959. A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* 37:911-917.

Carvalho, M., Sampaio, J.L., Palm, W., Brankatschk, M., Eaton, S., and Shevchenko, A. 2012. Effects of diet and development on the Drosophila lipidome. *Mol. Syst. Biol.* 8: 600.

Dennis, E.A. 2009. Lipidomics joins the omics evolution. *Proc. Natl. Acad. Sci. U.S.A.* 106:2089-2090.

Deutsch, E. 2008. mzML: A single, unifying data format for mass spectrometer output. *Proteomics* 8:2776-2777.

Ejsing, C.S., Duchoslav, E., Sampaio, J., Simons, K., Bonner, R., Thiele, C., Ekroos, K., and Shevchenko, A. 2006. Automated identification and quantification of glycerophospholipid molecular species by multiple precursor ion scanning. *Anal. Chem.* 78:6202-6214.

Ejsing, C.S., Sampaio, J.L., Surendranath, V., Duchoslav, E., Ekroos, K., Klemm, R.W., Simons, K., and Shevchenko, A. 2009. Global analysis of the yeast lipidome by quantitative shotgun mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* 106:2136-2141.

Fernandez, C., Schuhmann, K., Herzog, R., Fielding, B., Frayn, K., Shevchenko, A., James, P., Holm, C., and Strom, K. 2011. Altered desaturation and elongation of fatty acids in hormone-sensitive lipase null mice. *PLoS One* 6: e21603.

Folch, J.M., Lees, M., and Sloane-Stanley, G.H. 1957. A simple method for the isolation and purification of total lipids from animal tissue. *J. Biol. Chem.* 226:497-509.

Graessler, J., Schwudke, D., Schwarz, P.E., Herzog, R., Shevchenko, A., and Bornstein, S.R. 2009. Top-down lipidomics reveals ether lipid deficiency in blood plasma of hypertensive patients. *PLoS One* 4:e6261.

Gross, R.W. and Han, X. 2011. Lipidomics at the interface of structure and function in systems biology. *Chem. Biol.* 18:284-291.

Han, X. and Gross, R.W. 2005. Shotgun lipidomics: Electrospray ionization mass spectrometric analysis and quantitation of cellular lipidomes directly from crude extracts of biological samples. *Mass Spectrom. Rev.* 24:367-412.

Han, X., Yang, K., and Gross, R.W. 2012. Multidimensional mass spectrometry-based shotgun lipidomics and novel strategies for lipidomic analyses. *Mass Spectrom. Rev.* 31:134-178.

Harkewicz, R. and Dennis, E.A. 2011. Applications of mass spectrometry to lipids and membranes. *Annu. Rev. Biochem.* 80:301-325.

Herzog, R., Schwudke, D., Schuhmann, K., Sampaio, J.L., Bornstein, S.R., Schroeder, M., and Shevchenko, A. 2011. A novel informatics concept for high-throughput shotgun lipidomics based on the molecular fragmentation query language. *Genome Biol.* 12:R8.

Herzog, R., Schuhmann, K., Schwudke, D., Sampaio, J.L., Bornstein, S.R., Schroeder, M., and Shevchenko, A. 2012. LipidXplorer: A software for consensual cross-platform lipidomics. *PLoS One* 7:e29851.

Klose, C., Surma, M.A., Gerl, M.J., Meyenhofer, F., Shevchenko, A., and Simons, K. 2012. Flexibility of a eukaryotic lipidome—insights from yeast lipidomics. *PloS One* 7:e35063.

Matyash, V., Liebisch, G., Kurzchalia, T.V., Shevchenko, A., and Schwudke, D. 2008. Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics. *J. Lipid Res.* 49:1137-1146.

Sampaio, J.L., Gerl, M.J., Klose, C., Ejsing, C.S., Beug, H., Simons, K., and Shevchenko, A. 2011. Membrane lipidome of an epithelial cell line. *Proc. Natl. Acad. Sci. U.S.A.* 108:1903-1907.

Schuhmann, K., Herzog, R., Schwudke, D., Metelmann-Strupat, W., Bornstein, S.R., and Shevchenko, A. 2011. Bottom-up shotgun lipidomics by higher energy collisional dissociation on LTQ Orbitrap mass spectrometers. *Anal. Chem.* 83:5480-5487.

Schuhmann, K., Almeida, R., Baumert, M., Herzog, R., Bornstein, S.R., and Shevchenko, A. 2012. Shotgun lipidomics on a LTQ Orbitrap mass spectrometer by successive switching between acquisition polarity modes. *J. Mass Spectrom.* 47:96-104.

Schwudke, D., Oegema, J., Burton, L., Entchev, E., Hannich, J.T., Ejsing, C.S., Kurzchalia, T., and Shevchenko, A. 2006. Lipid profiling by multiple precursor and neutral loss scanning driven by the data-dependent acquisition. *Anal. Chem.* 78:585-595.

Schwudke, D., Hannich, J.T., Surendranath, V., Grimard, V., Moehring, T., Burton, L., Kurzchalia, T., and Shevchenko, A. 2007a. Top-down lipidomic screens by multivariate analysis of high-resolution survey mass spectra. *Anal. Chem.* 79:4083-4093.

Schwudke, D., Liebisch, G., Herzog, R., Schmitz, G., and Shevchenko, A. 2007b. Shotgun lipidomics by tandem mass spectrometry under data-dependent acquisition control. *Methods Enzymol.* 433:175-191.

Shevchenko, A. and Simons, K. 2010. Lipidomics: Coming to grips with lipid diversity. *Nat. Rev. Mol. Cell. Biol.* 11:593-598.

Sparkman, O.D. 2006. Mass spectrometry desk reference. 2nd ed. Global View Publishing, Pittsburgh.

Wenk, M.R. 2010. Lipidomics: New tools and applications. *Cell* 143:888-895.

APPENDIX: SETTINGS FOR IMPORTING MASS SPECTRA

selection window: Defines the width of the isolation window, which is used by a tandem mass spectrometer for selecting precursor ions for fragmentation. The width is expressed in Daltons (Da). For example, a typical isolation window for LTQ Orbitrap mass spectrometers is within 1 to 2 Da. The precise definition of the selection window w for a peak p is:

$$[p - w, p + w]$$

This is a mandatory setting.

timerange: Only scans acquired within the time window will be imported and processed. It is specified by the start and end times expressed in seconds. This is a mandatory setting.

calibration masses: This is a list of m/z values of reference ions which are used to offset the spectra mass calibration. LipidXplorer will search for their m/z in the spectra and calculate the m/z deviations and use them for adjusting masses of other peaks and improving the mass accuracy. If several reference masses are provided, the mass deviation will be calculated by a linear approximation. This setting is optional.

massrange: Defines the imported m/z range in the spectra. Reducing the mass range decreases the processing time and the number of peaks imported into a MasterScan. This speeds up the identification. The range is defined by specifying the flanking masses. Also, the lower mass is the starting m/z value for the resolution gradient which is discussed below. This setting is mandatory.

resolution: The actual mass resolution in the spectra, defined as:

$$R = \frac{M}{\Delta M}$$

This value is used for averaging and aligning the spectra. The algorithms will, respectively, average and align the peaks if their m/z are closer than allowed by the actual mass resolution. This is one of the most important settings that enable LipidXplorer to correctly import spectra acquired at different mass spectrometers. This setting is mandatory.

tolerance: LipidXplorer calculates accurate m/z of lipids at run time and compares them to the experimental m/z values from the spectra. The tolerance value is the maximum allowed deviation between experimental and calculated m/z set as parts per million (ppm) or Daltons (Da). It is not applied while importing spectra into a MasterScan, but becomes important later for the lipid identification. It is used as follows: a theoretical mass m measured with a given tolerance a fits to a peak p if $\in[p - a, p + a]$. This setting is mandatory.

threshold: Is the required minimum intensity for a peak to be imported in the MasterScan. LipidXplorer offers to input the threshold value as *relative* or *absolute*. The *relative* threshold is the percentage of the base peak intensity minus the highest peak in the spectrum. The base peak for the relative threshold used for the aligned spectra is the highest base peak of all imported spectra. LipidXplorer corrects the threshold value by dividing it by the square root of the number of scans in the given mass spectra time segment. This setting is mandatory.

min occupation: This states the minimal fraction of all acquisitions in which a particular ion should be recognized in order to be included into a MasterScan. For example: a min occupation of 0.5 indicates that each ion in the MasterScan was observed in at least 50%

of all samples. Effectively, this is a user-defined threshold of peak pattern reproducibility between independent replicates. The user specifies which samples are replicates using the Group Samples button in the Import Source panel. This opens a window in which the replicates are defined. Once the replicates are defined, the min occupation should be set to `1.0`. Now, only peaks that are present in all replicates will be imported. The min occupation is a mandatory option.

resolution gradient: Instrument mass resolution changes along the *m*/*z* scale. For accurate averaging and alignment, LipidXplorer simulates these changes. The gradient is approximated by a linear function with a slope and a starting value. The starting value is the resolution provided in the resolution setting at the starting *m*/*z* of the mass range; the slope is defined here as a "resolution gradient." For example: a value of -78.5 means that the resolution decreases by about 78.5 arbitrary units when the mass increases by one a.m.u. On some Orbitrap machines, the resolution may drop by 50,000 within the *m*/*z* range of 300 to 1200. This setting is mandatory.

MS1 offset: Masses could be shifted by the same offset in all acquisitions because of imprecise calibration of the mass spectrometer, but could be adjusted by applying an offset (expressed in Daltons). This setting is optional.

PMO: The Precursor Mass Offset (PMO) is a correction function mainly used for spectra acquired on LTQ Orbitrap and LTQ Orbitrap XL instruments. It corrects a precursor mass offset that is often applied in MS/MS experiments (Schuhmann et al., 2011) to improve the isolation efficiency. Otherwise, LipidXplorer will fail to correctly associate precursors and their MS/MS spectra. The direction of the shift is given by a positive or negative prefix, which is opposite to the offset selected at the LTQ Orbitraps. Note that this is not required on LTQ Orbitrap Velos and later instruments.