LipidCreator (v1.1.0.x)

Manual (v17), Jan. 28th, 2020

Contributors: Bing Peng, Dominik Kopzcyinski, Nils Hoffmann

This manual provides general information on the usage of LipidCreator. It explains all functionalities covered by LipidCreator's four interactive tutorials and provides additional information for more advanced use-cases. To start a tutorial, the user has to click on one of the four tutorial buttons in the "Home" tab as illustrated in Figure T1(B). The following tutorials are currently provided: i) setting up a PRM or ii) SRM analysis, iii) creating heavy isotope labeled lipids and iv) adding optimized collision energies for mediator analysis. These tutorials guide the user through the user interface of LipidCreator and explain step by step all important information relevant for the understanding of the four workflows.

T1. Generate lipid molecules
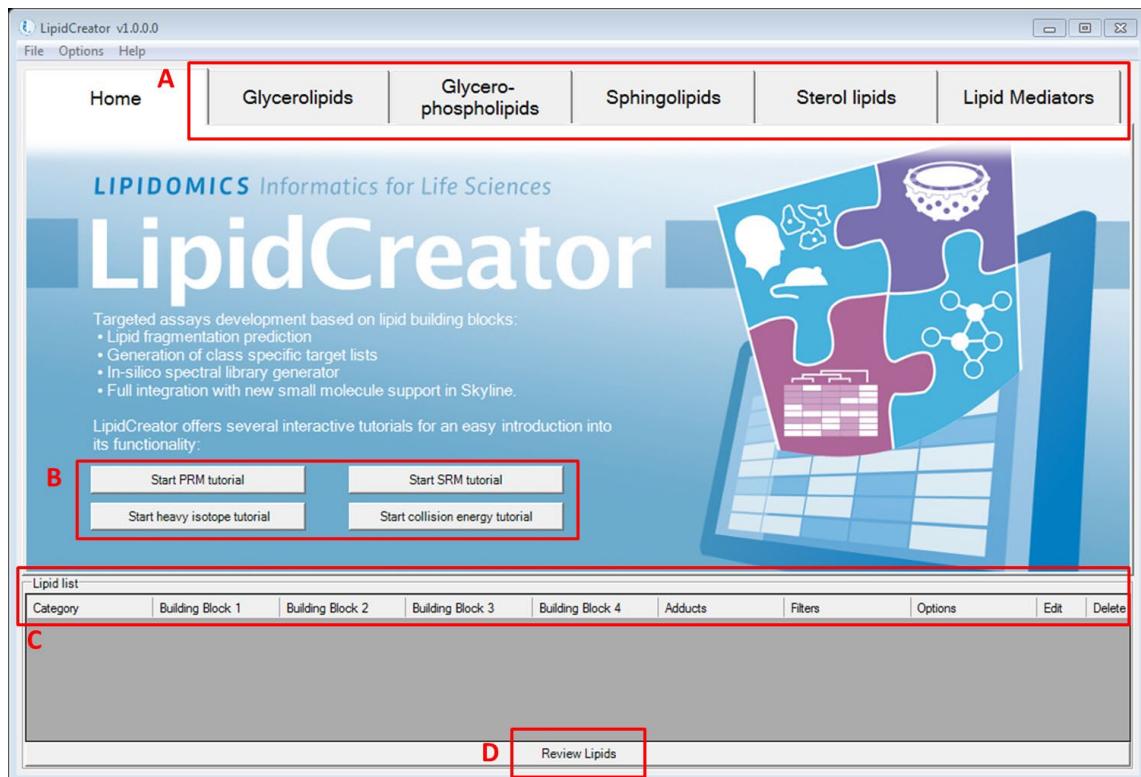
Figure T1. Home tab of LipidCreator.

A: Tab pages for five lipid categories. Each tab provides an individual interface.

B: Interactive tutorials that guide the user step-by-step through LipidCreator.

C: 'Lipid list' serves as a "shopping basket" to collect lipid assemblies.

D: Launching the computation of a transition list according to the 'Lipid list' selection.
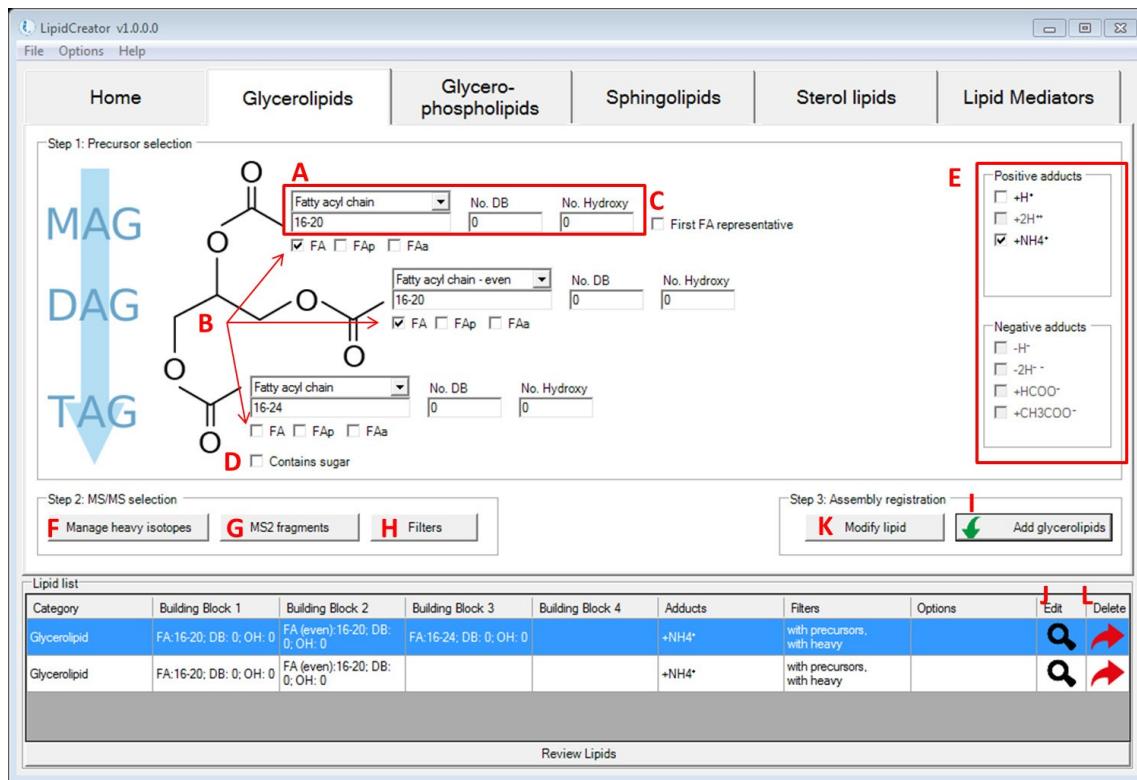
Figure T2. Interface for the definition of transitions for glycerolipids.

A: These fields let the user define one fatty acyl chain. From the drop-down list, the calculation for fatty acyl (FA) length can be chosen for even, odd or all potentially possible FA lengths. Input of either FA length, number of double bonds (DB), or number of hydroxy groups on this FA may be a range of numbers or individual values, e.g. '8, 9, 16-20, 23'. In LipidCreator, the allowed range for one FA length is 2-30, No. DB is 0-6, Hydroxy No. is 0-10 (except for sphingolipids). For sphingolipids, Hydroxy No. for long chain base (LCB) is 2 and 3, for FA, it is 0-3.

B: The check boxes for (de)selecting different types of FAs. For glycerolipids, the number of checked FAs defines the lipid class. TAG has three FAs; DAG has two FAs, whereas MAG has only one FA. FA here refers to an ester-linked fatty acid. FAp and FAa are ether-linked fatty acids. FAp has an ether bond to an alkenyl group, and FAa has an ether bond to an alkyl group, respectively. The minimum number of DBs for FAp is 1.

C: The "First FA representative" checkbox is to quickly apply FA information from Figure T2A to all other FAs in current interface.

D: This checkbox is to replace one FA with a glucose head group (Figure T3).

E: All supported adducts in LipidCreator. For positive mode, [M+H]1+, [M+2H]2+ and [M+NH4]1+ are valid. For negative mode, [M-H]1-, [M-2H]2-, [M+HCOO]1- and [M+CH3COO]1- are valid. In LipidCreator, recommendation on adduct selection can be reviewed, when hovering the mouse cursor over different head groups (HG).

F: To manage heavy isotopes for glycerolipids, please go to [Section S2.6](#) for details.

G: To select MS^2 fragments for glycerolipids, please go to [Section S2.4](#) for details.

H: To apply filters for glycerolipids, please go to [Section S2.7](#) for details.

I: Adds the complete lipid assembly into the 'Lipid list" basket.

J-K: To modify a lipid assembly, double click on the 🔍 icon to retrieve information according to Step 1 window. After making changes (including HG selection, FA profile, adducts selection, management of heavy isotopes, MS^2 fragments selection and filters selection), click on 'Modify lipid' from Step 2 window to update the assembly.

L: To delete a lipid assembly from the 'Lipid list', double click on the ➦ Icon.

Figure T3. Interface for the definition of transitions for glycerolipids with glucose head group.

A: Multiple selection is possible when choosing head groups in LipidCreator.

B: The selected head groups are displayed in 'Lipid list'.

Figure T4. Interface for the definition of transitions for glycerophospholipids.

A: Multiple selection of head groups is possible when choosing them from the head group menu.

B: Radio buttons switch between different types of glycerophospholipids.

C: Ester- or ether-linked fatty acyls (fatty acid, plasmenyl or plasmanyl) can be selected.

Figure T5. Interface for the definition of transitions for lyso-glycerophospholipids.

Figure T6. Interface for the definition of transitions for cardiolipins.

Figure T7. Interface for the generation of transitions for sphingolipids.

A: Head group selection for sphingolipids in LipidCreator.

B: Radio buttons switch between different backbones of sphingolipids.

Figure T8. Interface for the definition of transitions for lyso-sphingolipids.
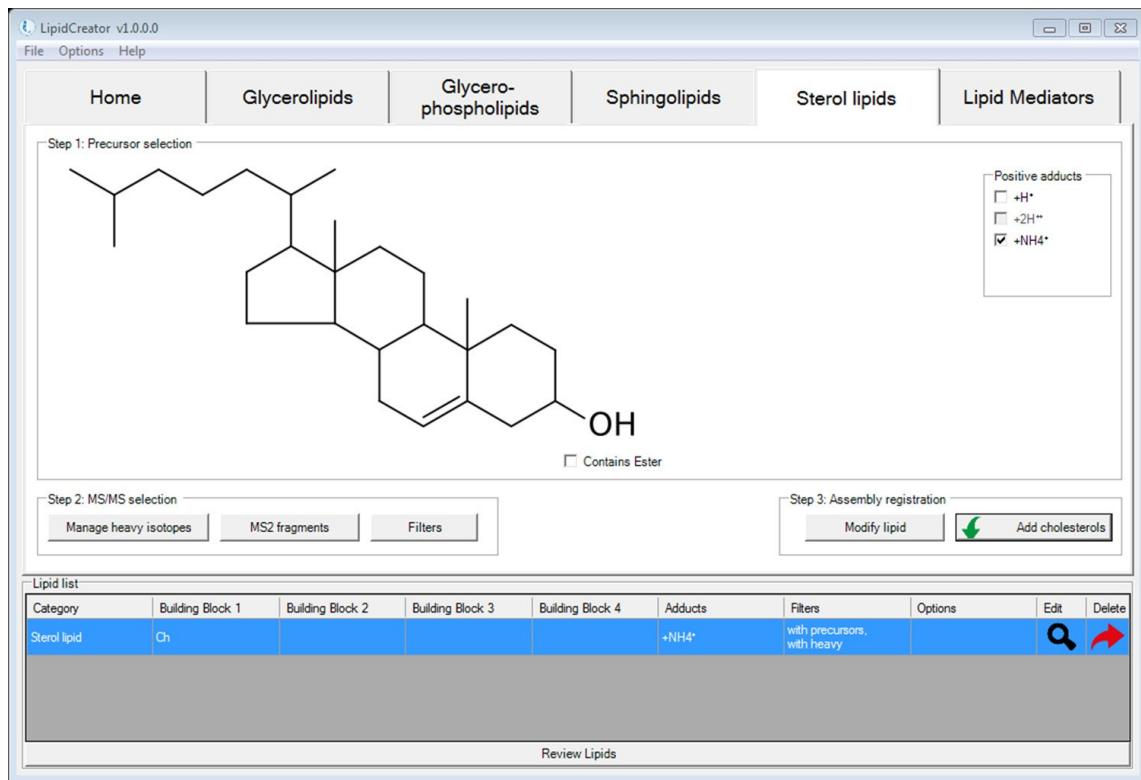
Figure T9. Interface for the definition of transitions for cholesterol.
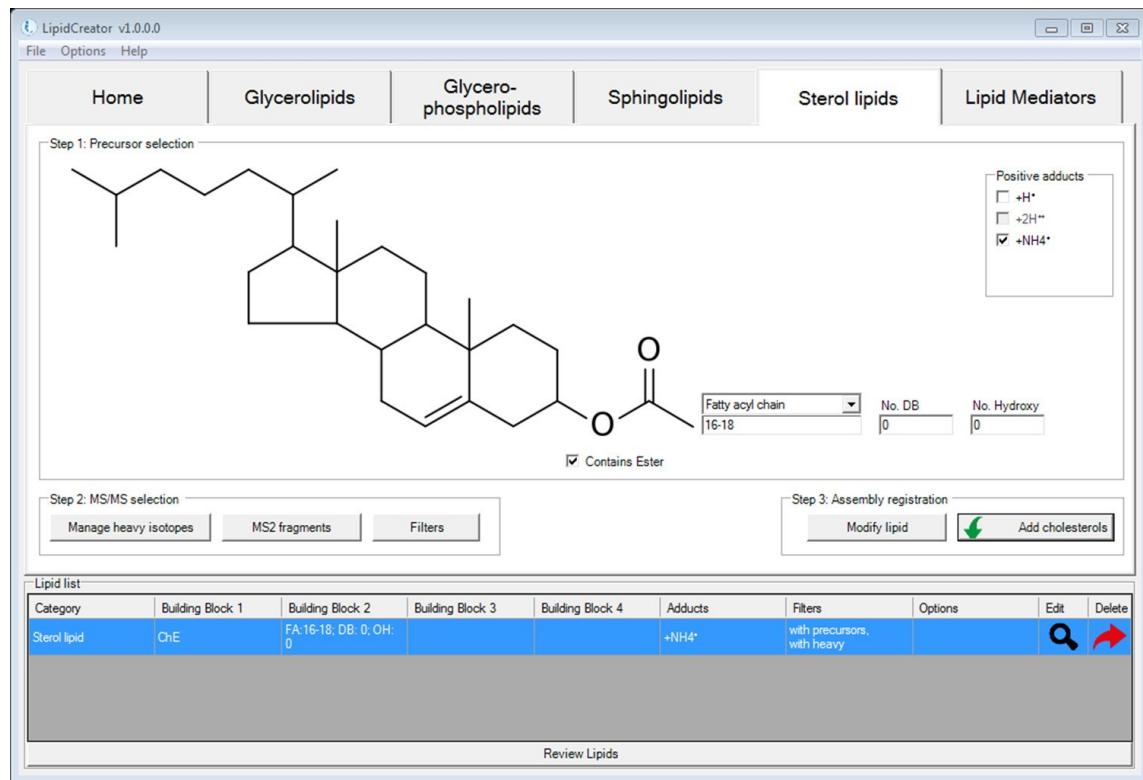
Figure T10. Interface for the definition of transitions for cholesteryl esters.

Figure T11. Interface for the definition of transitions for lipid mediators.

A: Head group selection for lipid mediators in LipidCreator.

B: This area displays the chemical structure of each lipid mediator when hovering the mouse cursor over different mediator names.

T2. Import/export lipid list/setting/project
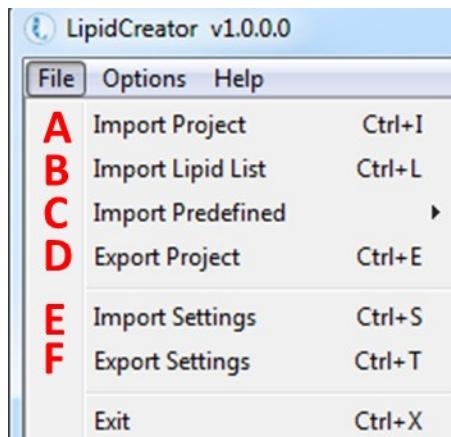
Figure T12. File Menu of LipidCreator.

A: Import project into LipidCreator. A project includes the lipid list, user defined MS^2 fragments, user-defined heavy labelled isotopes, and a selection of optimal collision energies.

B: Import lipid list from *.csv file. The lipid list should follow the nomenclature described in Table S2. Otherwise, please use the lipid name translator (Section S2.3) for import. In the *.csv file, lipid names (including adduct name) should be given one per line.

C: Import predefined lipid lists from previous work.

PNAS yeast, DOI: 10.1073/pnas.0811700106

Mouse brain, DOI: 10.1007/s13361-014-1013-x and 10.1038/s41592-018-0010-6

Mouse heart, DOI: 10.1016/j.celrep.2018.08.017

Mouse platelet, doi: 10.1182/blood-2017-12-822890

Human platelet, doi: 10.1182/blood-2017-12-822890

D: Export current project in *.lcXML format from LipidCreator for storage.

It is possible to add user-defined project into 'Predefined' when use LipidCreator standalone. Copy *.lcXML file into the folder (or create new folder) at…/LipidCreator/data/predefined.

E: Import settings in LipidCreator. Settings include user defined MS^2 fragments, user defined heavy labelled isotopes, and a selection of optimal collision energies.

F: Export the current settings in *.lcXML format.

T3. Lipid name translator

Figure T13. Interface for the lipid name translator (Options→Lipid name translator).

A: List of old lipid names. Not all isotope formats can be recognized directly from the name. The isotope labels need to be defined additionally in LipidCreator.

B: List of translated lipid name according to the nomenclature in LipidCreator. When the old lipid name has no adduct defined, a default adduct will be appended to the translated name automatically. After translation, the list will be imported into LipidCreator by clicking on 'Insert'.

T4. Select MS2 fragments
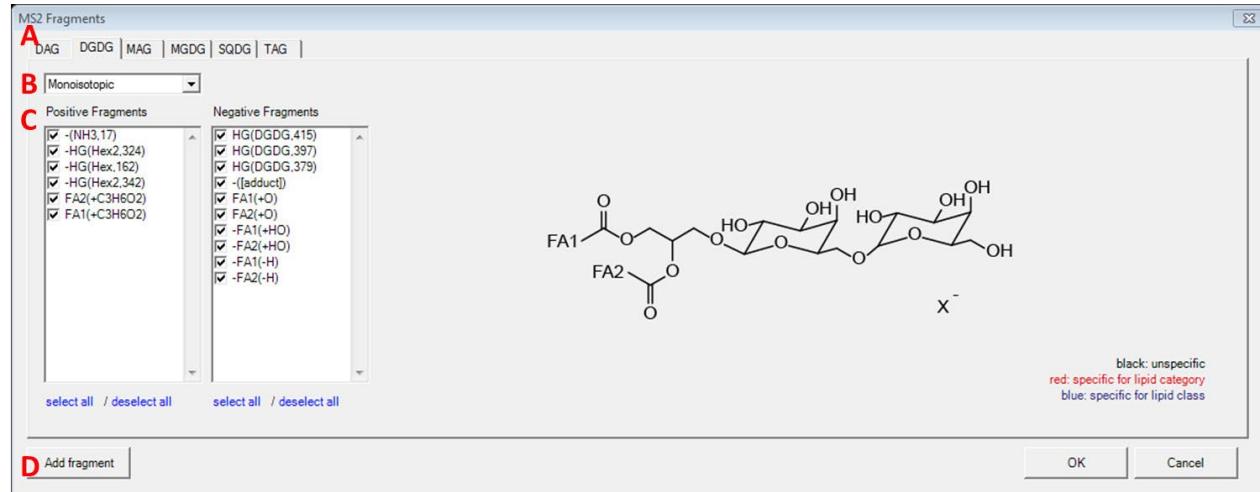
Figure T14. Interface for MS^2 fragments for glycerolipids.

A: Tabs for selecting lipid classes.

B: Drop-down list for selecting either monoisotopic or isotopic species. Isotopic coded species will only show up after being defined (Section S2.6).

C: List of MS^2 fragment types of each lipid class for positive and negative mode. Each predefined fragment type offers a chemical structure preview on the right side.

D: Add user-defined fragments, please see Section S2.5, Figure T20 for details.
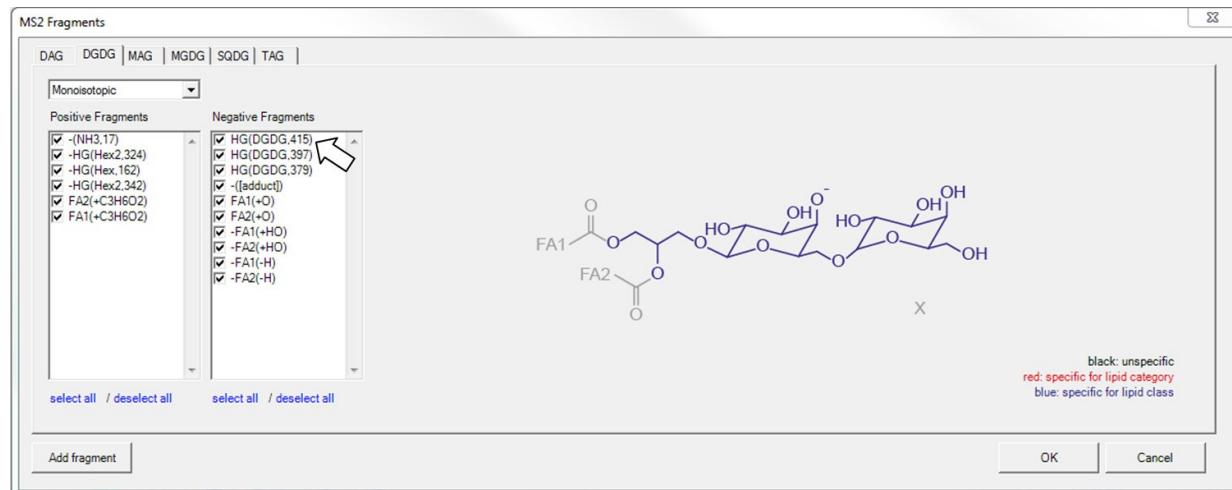
Figure T15. Preview of MS^2 fragments for glycerolipids when hovering over fragment names.
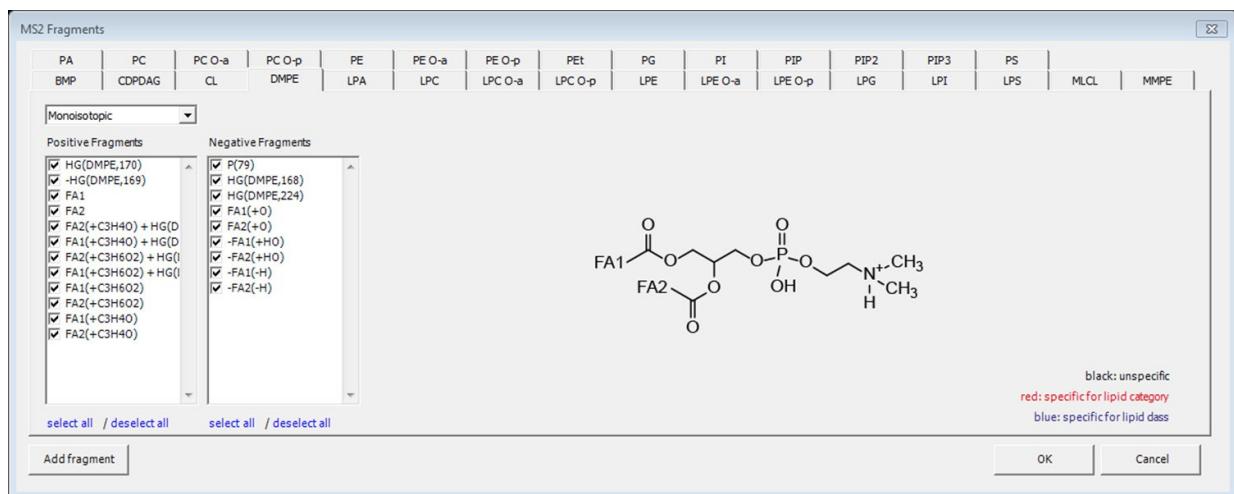
Figure T16. Interface for MS^2 fragments for glycerophospholipids.
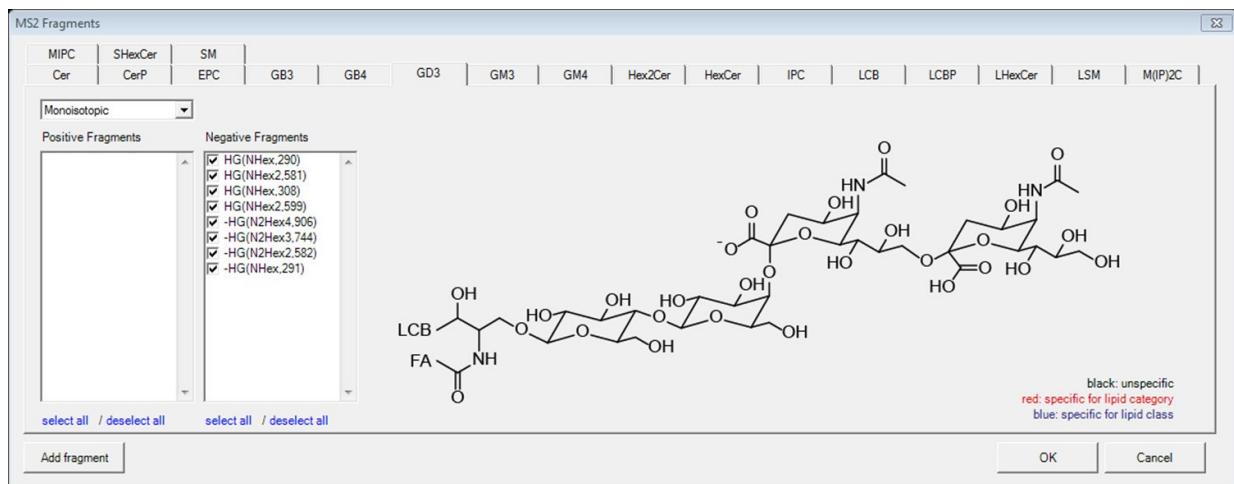
Figure T17. Interface for MS^2 fragments for sphingolipids.
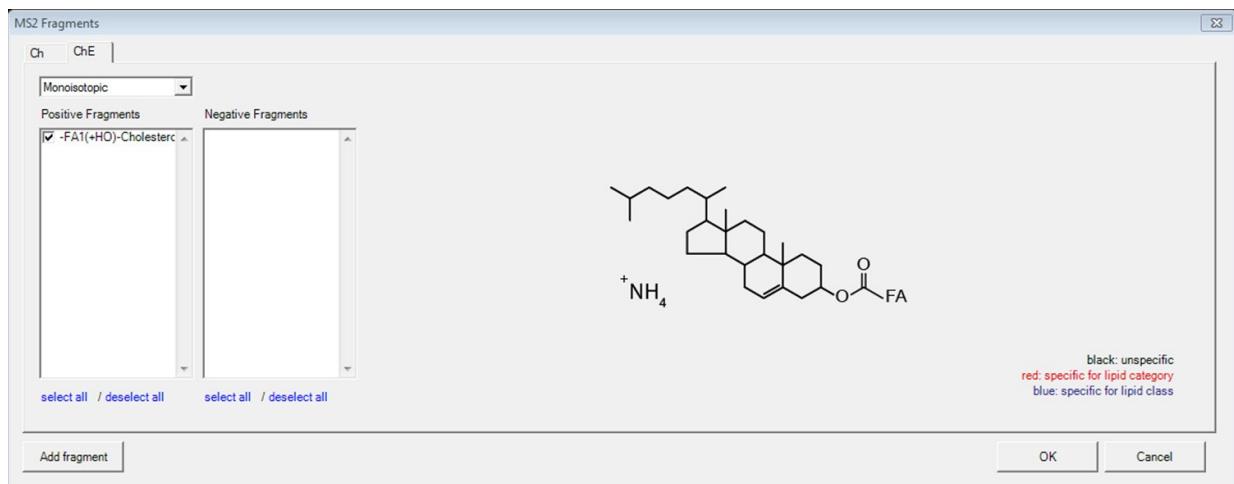
Figure T18. Interface for MS^2 fragments for cholesterol and cholesteryl esters.
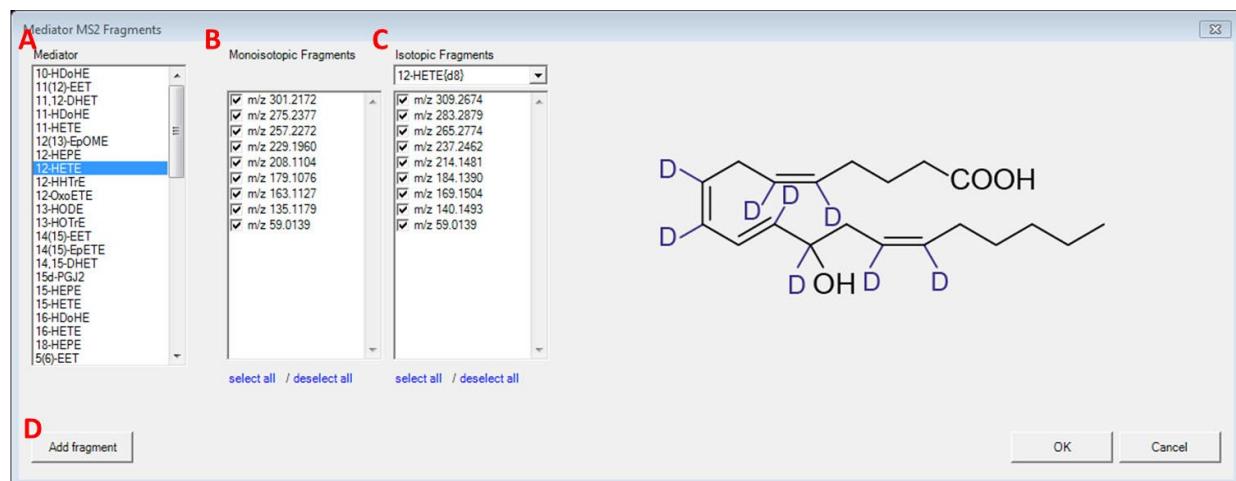
Figure T19. Interface for MS^2 fragment masses for lipid mediators.

A: List of individual lipid mediators.

B: List of MS^2 fragment masses for selected lipid mediator.

C: List of MS^2 fragment masses for isotope labelled versions of the selected lipid mediator.

D: Add user-defined fragments, please see Section S2.5, Figure T24 for details.

T5.   Generate user-defined fragments
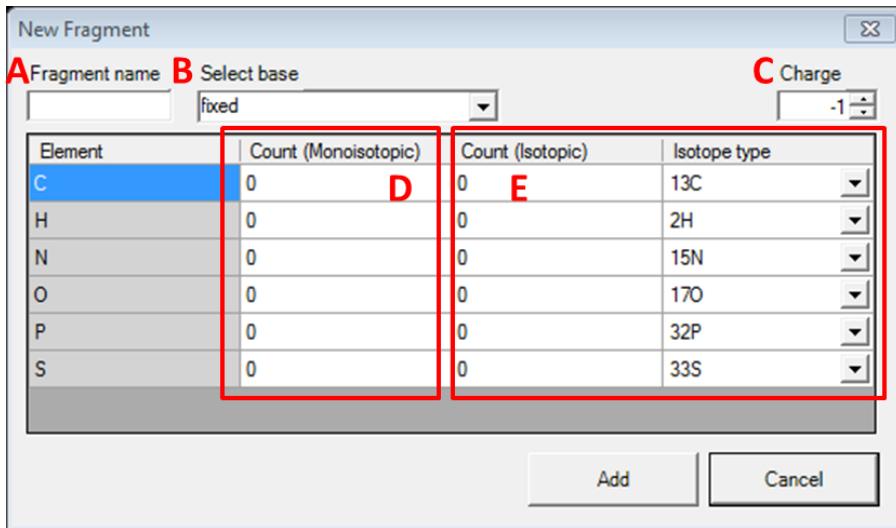
Figure T20. Interface for adding new fragments.

A: User-defined fragment name.

B: Select base of the user-defined fragment type. Depending on the chosen lipid class, either the base can be fixed or it can contain building blocks, e.g. HG, HG + FA1, FA1 + FA2, etc.

C: Fragment charge. A positive value indicates that this fragment originated from positive ionization mode, while a negative value indicates that this fragment originated from negative ionization mode.

D: A constant set of elements can be defined which will be added to the fragment. When 'fixed' base is selected, element numbers can only be positive; otherwise negative counts are also allowed.

E: A field to input constant set of isotopic elements and drop-down list to select the type of isotopic elements.

Here is an example of how to add new types of fragments for PG.
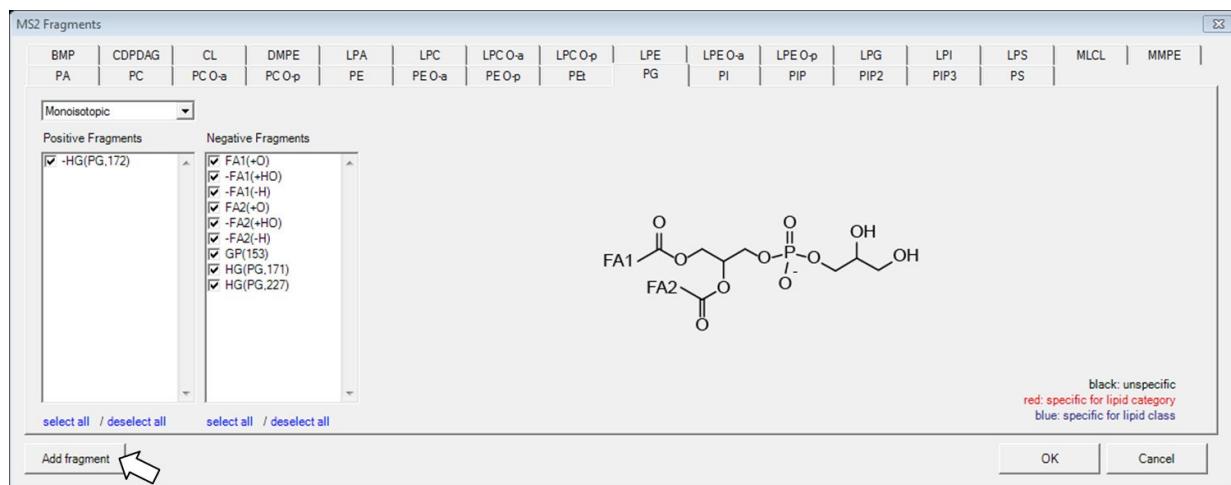
Figure T21. Click on 'Add fragment' from the PG tab of glycerophospholipids.
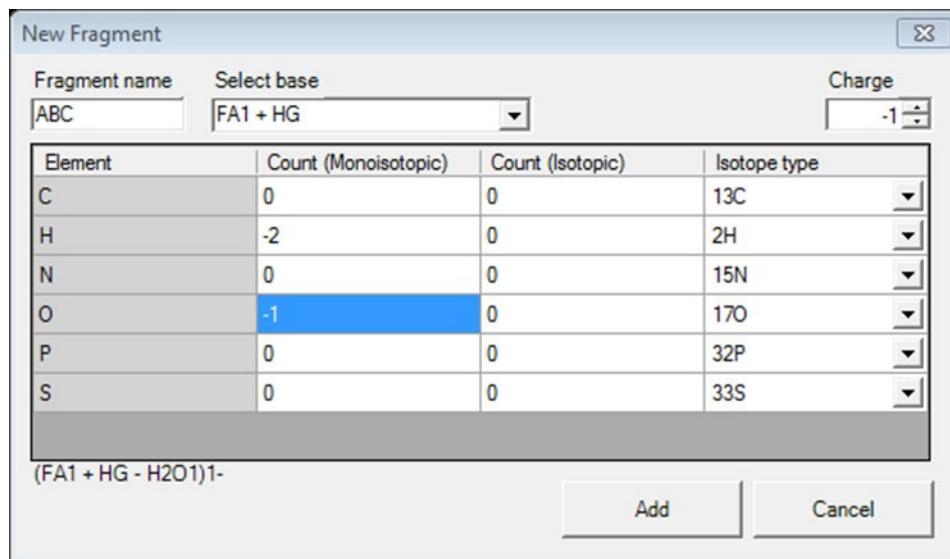
Figure T22. Generate a fragment 'ABC' for PG class.
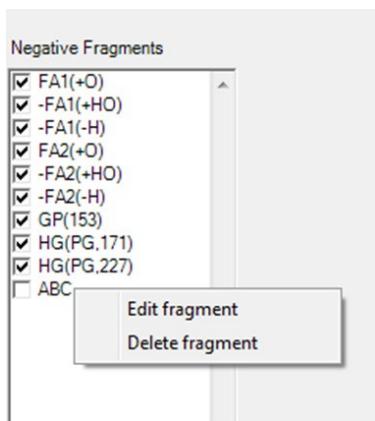
Figure T23. The 'ABC' fragment has been added to the fragment list. User-defined fragments can be edited or deleted by right clicking on the fragment name.
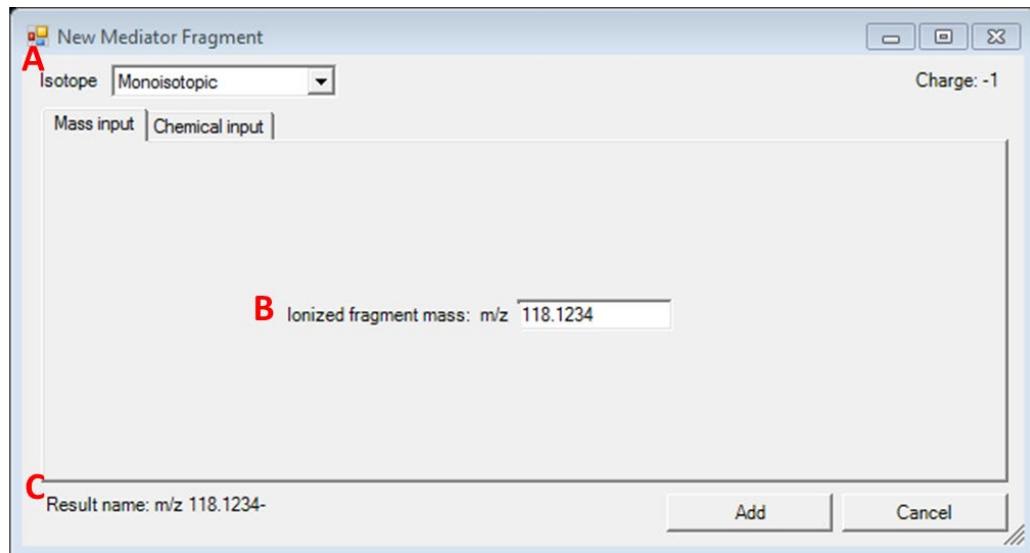
Figure T24. Adding a monoisotopic mediator fragment with direct mass input for Mediator.

A: In the 'Isotope' drop-down list, either monoisotopic or other isotopes can be selected to add a fragment, depending on the mediator selection in Figure T19.

B: In the 'Mass Input' tab, please directly type in the ionized fragment mass. The fragment charge is set to -1 by default.

C: The preview for display name of the fragment.

The mass will appear in the transition list with exactly this value and will be identical in the current list. Identified replicates will be denied when adding.
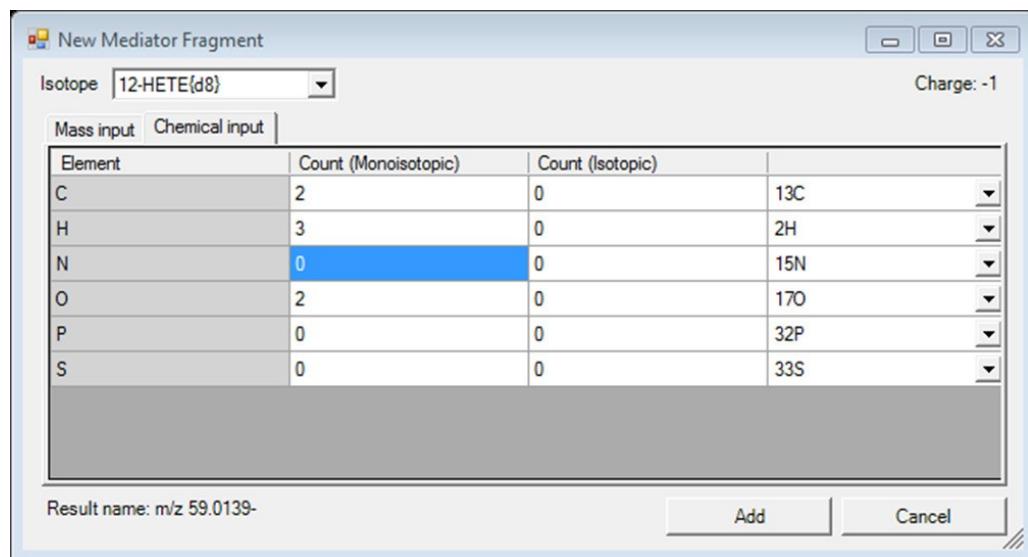
Figure T25. Adding a fragment using the chemical formula input for Mediator.

To add a fragment for a monoisotopic species, use only the 'Count (Monoisotopic)' column to enter the counts of elements. For heavy isotope species, both 'Count (Monoisotopic)' and 'Count (Isotopic)' are valid. The result name will be displayed as the (m/z) mass.

T6. Manage heavy isotopes
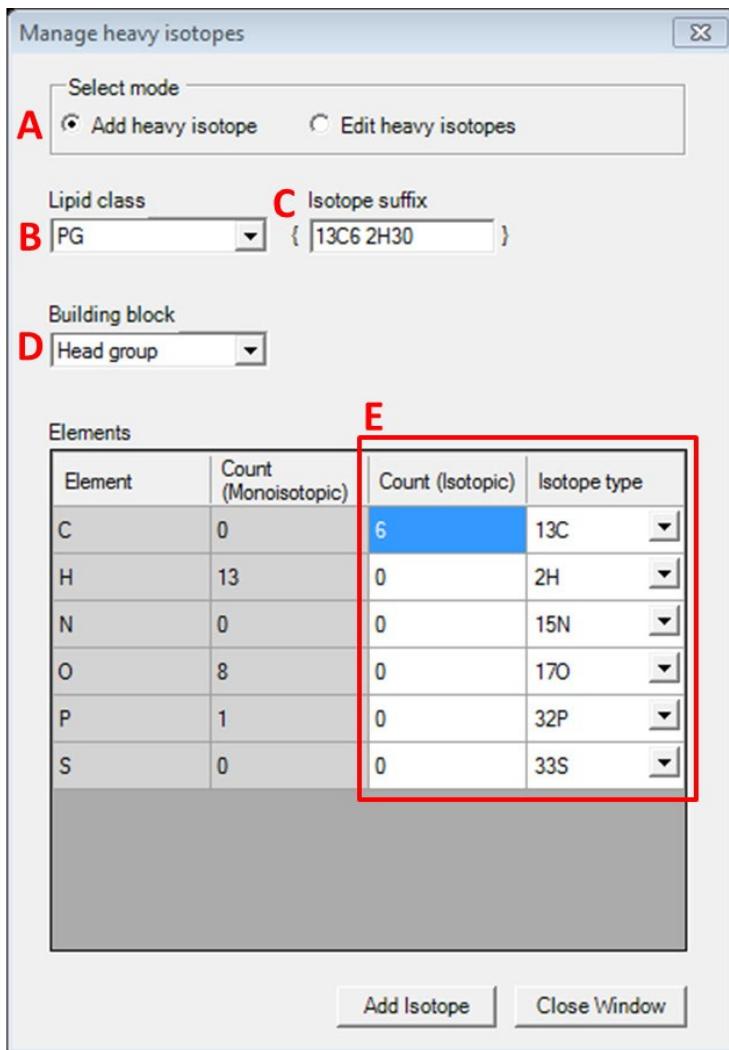
Figure T26. Interface for managing heavy isotopes for glycerophospholipids.

This interface is automatically adapted to the category where user is currently working.

A: Radio buttons for either adding new isotopes or editing existing user-defined isotopes.

B: Drop-down list for selecting the lipid class (depending on the lipid category).

C: User-defined name as suffix for the lipid class.

D: Drop-down list for selecting building blocks.

E: A field to input constant set of isotopic elements. The count for monoisotopic elements will be automatically changed according to the typed count of isotopic elements.
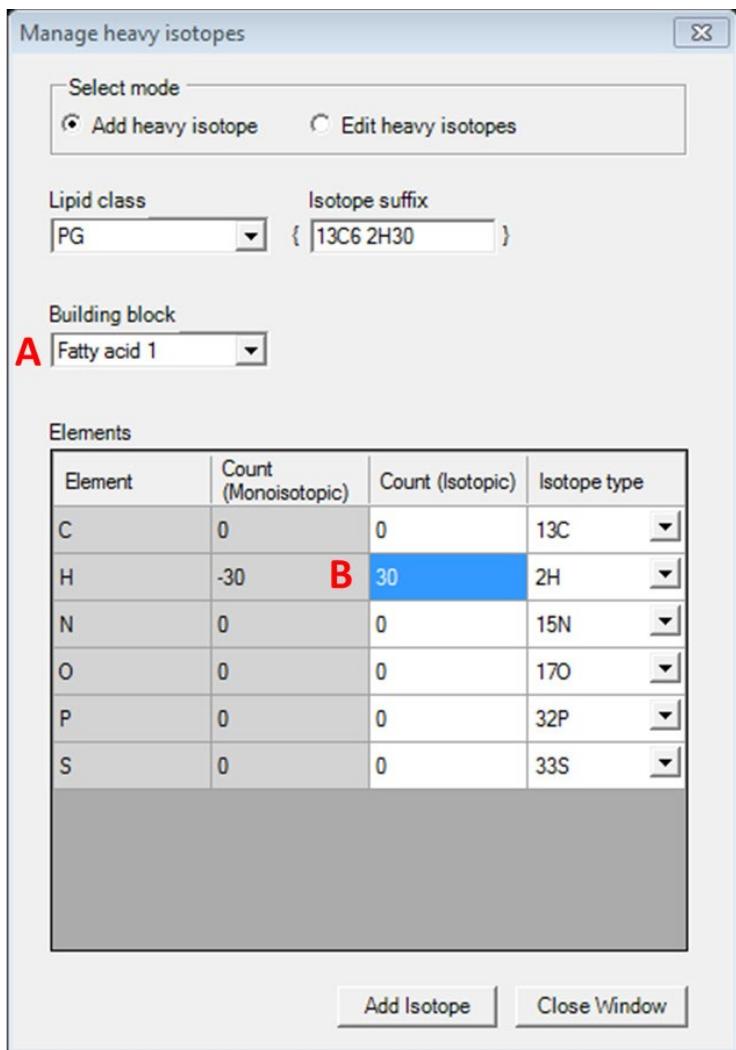
Figure T27. Interface for managing heavy isotopes for glycerophospholipids.

A: Select 'Fatty acyl 1' to add isotope elements.

B: The heavy-labelled element numbers act as an upper limit for the element, since the fatty acyl building block has a variable number of elements depending e.g. on the carbon chain length.
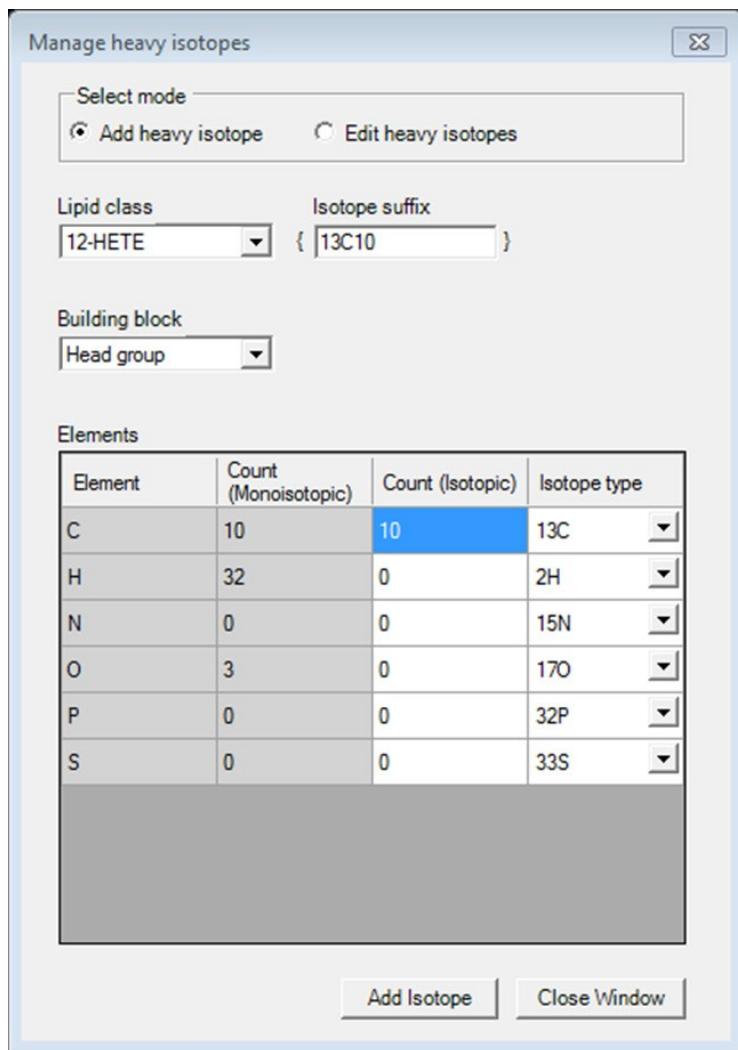
Figure T28. Interface for managing heavy isotopes for mediators.

The chemical formula (numbers of elements in 'Count (Monoisotopic)' column) for all mediators are provided when adding heavy isotopes.

T7.   Filters for the transition list
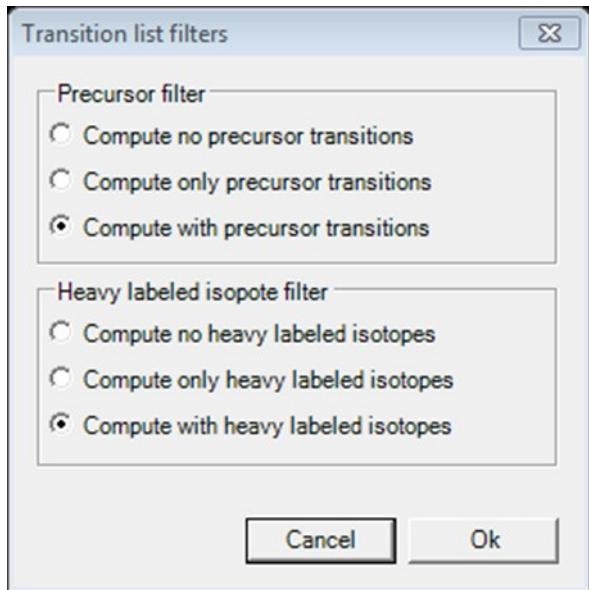
Figure T29. Filters are applied before the transition lists are calculated.

After changing the filter settings, please click on 'Modify lipid' to update the 'Lipid list'.

T8. Collision energy optimization function

Collision energy function is valid for lipid mediators. To activate this function, first click on Options→ 'Collision Energy Computation'→ 'Thermo Scientific Q Exactive HF' or 'Agilent 6545 Q-TOF LC/MS'.

Afterwards, open the dialog from Options→ 'Collision Energy optimization'.
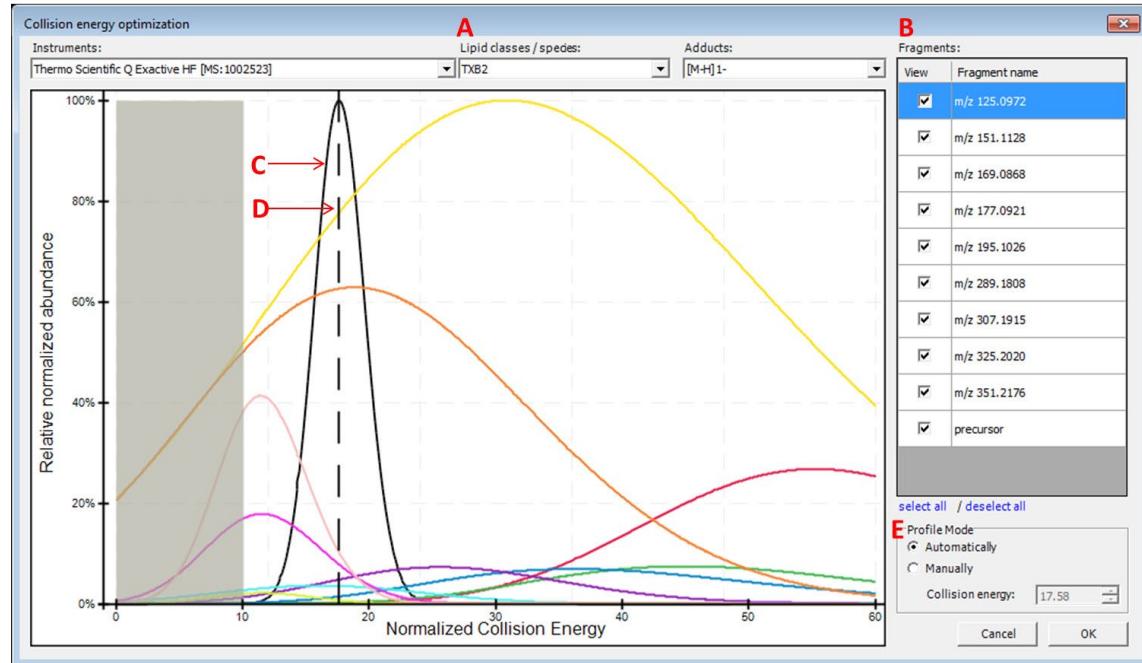
Figure T30. Collision energy optimization interface.

A: Choose a lipid mediator species from the drop-down list. Isotopes appear as separate entries to their monoisotopic counterparts.

B: Fragment list for the chosen lipid species. Optimal collision energy depends on fragment selection.

C: The black curve is the automatically calculated product distribution over all selected fragment distributions from the list. Its mode indicates the optimal collision energy over all selected fragments.

D: The dashed line indicates the chosen collision energy. This line is moveable in manual profile mode.

E: Automatic or manual profile mode can be selected. When the manual profile mode is activated, either move the dashed line with the mouse or type in a collision energy value of your choice.

Activating the computation of collision energy is independent from added lipid species in the 'Lipid list'. The collision energy can be defined before or after lipid assembly, these two information will combine after click on 'Review Lipids'.

By applying collision energy computation, the CE value can be used either to generate an MS method from the transition list or to provide corresponding relative intensities for fragments in the generated *in-silico* spectral library.

T9. Reviewing the lipid transition list

After click on Review Lipids, the calculated lipid precursor names (Figure T31) will be listed for further selecting/deselect before calculate transitions (Figure T32, 33).

Figure T31. Interface for reviewing the name of calculated lipid precursors.

A: The checkbox allows you to select/deselect the generated lipids for further calculation of transitions.

B: Click on the name of each column to activate sort ascending/descending functions.

C: Display options between subspecies and species level of lipid name. e.g. DAG 16:0-16:0 in subspecies level, DAG 32:0 in species level.

Lipid Transitions Review

Molecule List Name	Precursor Name	Precursor Molecule Formula	Precursor Adduct	Precursor Ion m/z	Precursor Charge	Product Name	Product Molecule Formula	Product Adduct	Product Ion m/z	Product Charge	Note
DAG	DAG 16:0-17:0	C36H70O5	[M+NH4]1+	600.5562	+1	-(H2O+NH3,35)	C36H68O4	[M+H]1+	565.5190	+1	
DAG	DAG 16:0-17:0	C36H70O5	[M+NH4]1+	600.5562	+1	-FA 16:0(+H)-...	C20H38O3	[M+H]1+	327.2894	+1	
DAG	DAG 16:0-17:0	C36H70O5	[M+NH4]1+	600.5562	+1	-FA 17:0(+H)-...	C19H36O3	[M+H]1+	313.2737	+1	
DAG	DAG 16:0-18:0	C37H72O5	[M+NH4]1+	614.5718	+1	precursor	C37H72O5	[M+NH4]1+	614.5718	+1	
DAG	DAG 16:0-18:0	C37H72O5	[M+NH4]1+	614.5718	+1	-(H2O+NH3,35)	C37H70O4	[M+H]1+	579.5347	+1	
DAG	DAG 16:0-18:0	C37H72O5	[M+NH4]1+	614.5718	+1	-FA 16:0(+H)-...	C21H40O3	[M+H]1+	341.3050	+1	
DAG	DAG 16:0-18:0	C37H72O5	[M+NH4]1+	614.5718	+1	-FA 18:0(+H)-...	C19H36O3	[M+H]1+	313.2737	+1	
DAG	DAG 16:0-19:0	C38H74O5	[M+NH4]1+	628.5875	+1	precursor	C38H74O5	[M+NH4]1+	628.5875	+1	Interference ...
DAG	DAG 16:0-19:0	C38H74O5	[M+NH4]1+	628.5875	+1	-(H2O+NH3,35)	C38H72O4	[M+H]1+	593.5503	+1	Interference ...
DAG	DAG 16:0-19:0	C38H74O5	[M+NH4]1+	628.5875	+1	-FA 16:0(+H)-...	C22H42O3	[M+H]1+	355.3207	+1	
DAG	DAG 16:0-19:0	C38H74O5	[M+NH4]1+	628.5875	+1	-FA 19:0(+H)-...	C19H36O3	[M+H]1+	313.2737	+1	

Number of transitions: 45

Options: A ☐ Edit mode   B ☐ Only show unique transtions   ☐ Send spectral library to Skyline   E Check transition List

[Back] [Store transition list] [Send to Skyline] [Store spectral library]

Figure T32. Interface for reviewing calculated lipid transitions.

A: The checkbox allows you to activate user's edit mode for manually editing, adding and deleting the transition list.

B: The checkbox allows you to discard the non-unique transitions from the review panel and further actions.

C & D: The chemical formulas in the transition list represent neutral lipids/fragments. The masses represent ionized precursors/fragments after adding the corresponding adducts.

E: This button is to check whether current transition list is compatible with Skyline.

F: In the review list, the non-unique transitions (only the precursor and product masses are considered) will be highlighted and noted when being repeated. The note will be imported to Skyline when the transition list is being sent.

Figure T33. Interface for reviewing calculated lipid transitions with collision energy.

A: The 'Explicit Collision Energy' column will appear after collision energy computation is activated.

B: The 'Store spectral library' button is valid after collision energy computation is activated.

T10. Storing a transition list

Following the steps described in Section S2.9, the reviewed transition list can be exported as a *.csv or *.xls file by clicking on 'Store transition list'. The list can be stored in either one full list or it can be split into two lists separated by polarity mode.

T11. Storing a spectral library

The creation of a spectral library is possible after collision energy computation is activated according to Section S2.8 (Figure T33). Spectral libraries are written in Skyline *.blib format, which is a SQLIte database file.

T12. Integration with Skyline

12-1. Installation of Skyline

Please install Skyline through https://skyline.ms/project/home/software/Skyline/begin.view

12-2. Install LipidCreator to Skyline

Please install LipidCreator from Skyline Tools→Tool store.

Or

Please add the downloaded .zip file to Skyline through Tools→ External Tools→ Add→ From file…→ Choose LipidCreator.zip→ Wait until LipidCreator shows in the 'Menu contents' (this step may take some seconds)→ OK.

12-3. Create transition list for Skyline

Start LipidCreator from Tools and use it as described from Section S2.1-9. After the created transition list has been reviewed, click on 'Send to Skyline'. The lipid list will appear in the 'Targets' window in Skyline. Then either export the project (see Section S2.2) for further editing or close LipidCreator directly.

12-4. Create *In-silico* spectral library

After activating the collision energy computation (see [Section S2.8](#)), the created transition list and spectral library can be sent to Skyline at once (Figure T34).

Or

Add *.blib file manually through Settings→Peptide Settings→Library→Edit list→Add

Figure T34. Interface for reviewing calculated lipid transition with collision energy on Skyline platform.

A. This check box allows you to send spectral library to Skyline without additionally save *.blib file at local drive.

B. By clicking on this button, transitions (and spectral library) will be sent to Skyline.

12-5. Export MS method from Skyline

The final MS method can be generated from Skyline by selecting favored MS vendor/type.

12-6. Viewing LipidCreator log messages in Skyline

LipidCreator displays warnings and errors directly to the user as dialogs. However, detailed error messages are also logged to the Tools→Immediate Window in Skyline and into a log file "lipidcreator.log" below the Tools\LipidCreator\data directory of the Skyline installation.

12-7. Review PRM/DIA results with spectral matching

To review transitions including precursor, please set Transition setting→Filter→Ion types: f,p.

Select *in-silico* spectral library through Settings→Peptide Settings→Library. Then click on Library Match from View menu. After import data, extracted MS2 spectral through click on chromatographic peak.

T13. Statistics for LipidCreator launching

When initially launching LipidCreator, a dialog will appear with the following message:
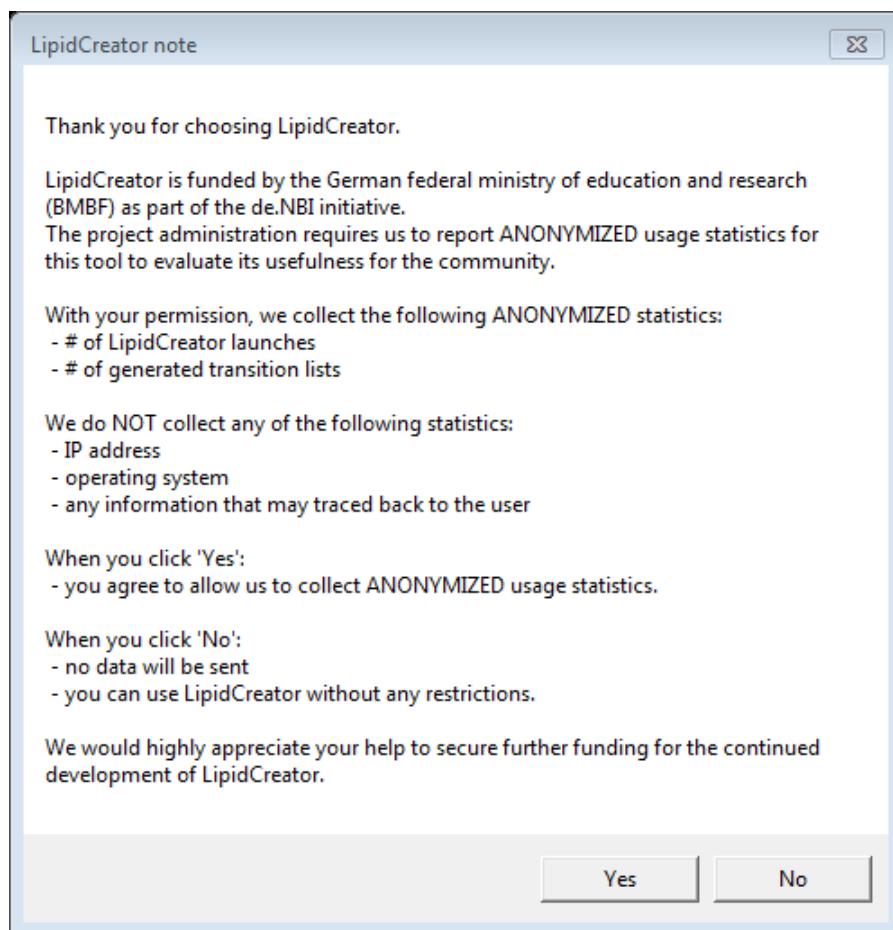
Figure T35. Note when launching LipidCreator for the first time.

If the user changes his/her mind after clicked on 'Yes' or 'No', this function can be turned off/on by (de)selecting Options→ 'Send anonymous statistics'.

T14. Command line usage

LipidCreator has a comprehensive command line interface allowing the user to run several tasks without a graphical user interface. Therefore, it is suitable for an integration into automated pipeline frameworks capable of adding customized processing nodes[1,2] or running command line executions[3]. Launching LipidCreator from the command line with the argument "help" prints the following output:

```
> .\LipidCreator.exe  help
usage: LipidCreator.exe (option)

options are:
  dev:                launching LipidCreator as  developer
  transitionlist:     creating transition list from lipid list
  translate:          translating a list with old lipid names into
                      current nomenclature
  library:            creating a spectral library in *.blib format
                      from a lipid list
  random:             generating a random lipid name (not
                      necessarily reasonable in terms of chemistry)
  agentmode:          secret agent mode
```

The options 'transitionlist', 'translate' and 'library' are the most common modes.

---

[1] Berthold, Michael R., et al. "KNIME-the Konstanz information miner: version 2.0 and beyond." AcM SIGKDD explorations Newsletter 11.1 (2009): 26-31.

[2] Goecks, Jeremy, Anton Nekrutenko, and James Taylor. "Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences." Genome biology 11.8 (2010): R86.

[3] Köster, Johannes, and Sven Rahmann. "Snakemake—a scalable bioinformatics workflow engine." Bioinformatics 28.19 (2012): 2520-2522.

Creating a transition list

A transition list contains all necessary information about the precursor name / mass / charge / adduct in combination with its fragment name / mass / charge / adduct / (collision energy). To create a transition list, a lipid list is necessary. An 'example-lipid-list.csv' can be found within the LipidCreator root directory in the folder 'data\examples'. The 'transitionlist' mode has the following options:

```
> .\LipidCreator.exe transitionlist help
Creating a transition list from a lipid list

usage: LipidCreator.exe transitionlist input_csv output_csv [opts
[opts ...]]
  opts are:
    -p 0:               Compute no precursor transitions
    -p 1:               Compute only precursor transitions
    -p 2:               Compute with precursor transitions
    -h 0:               Compute no heavy labeled isotopes
    -h 1:               Compute only heavy labeled isotopes
    -h 2:               Compute with heavy labeled isotopes
    -s:                 Split in positive and negative list
    -x:                 Developer or Xpert mode
    -l:                 Create LipidCreator project file instead
                        of transition list
    -d:                 Delete replicate transitions (equal
                        precursor and fragment mass)
    -c instrument mode: Compute with optimal collision energy (not
                        available for all lipid classes)
```

Optionally, several options (e.g. precursor or heavy labeled fragments) can be set or collision energy optimization can be enabled. A command for creating a transition list without replicates and heavy labeled isotope fragments (if by default included in LipidCreator) looks like this:

```
> .\LipidCreator.exe transitionlist data\examples\example-lipid-
list.csv output-transition-list.csv -h 0 -d
```

The resulting transition list file can be further imported in follow-up analysis tools such as Skyline.

Translating lipid names

Several different structural nomenclatures for lipid names exist. Especially, many legacy names exist for the same lipid species. This results in problems when parsing lipid lists in tools which only have one nomenclature implemented. To solve this issue, a lipid name translation engine was developed to recognize old and deprecated lipid names (e.g. as still used by LIPID MAPS)

to translate them into lipid species names according to the current nomenclature[4]. To translate lipid lists into the current nomenclature, the user has to type in:

```
> .\LipidCreator.exe translate data\examples\old-lipid-name-
list.csv new-nomenclature-lipid-list.csv
```

The old lipid names will be preserved in the new list. The recognition system is far from being complete, so we encourage all users who encounter issues with the translator to send us[5] their lipid lists. Such feedback allows us to extend the translation engine further.

Creating a spectral library

It is possible to create an *in-silico* spectral library in *.blib format with the command line interface. To estimate the relative fragment abundances, true MS^2 measurements were used to train parameterized statistical models. Although we measured different lipid species with different adducts on different platforms, we do not guarantee correctness of the estimated spectra or generalizability to arbitrary platforms. To create a spectral library for lipids, a lipid list (*.csv) or LipidCreator project file (*.lcXML) as well as the PSI-MS controlled vocabulary term[6] of the measurement platform is required. For example, the user has to type:

```
> .\LipidCreator.exe library data/examples/example-lipid-list.csv
output-library.blib MS:1002523
```

Here, 'MS:1002523' is the CV term for the instrument 'Thermo Scientific Q Exactive HF'. The created spectral library can be now easily imported in follow-up analysis tools such as Skyline[7].

---

[4] Pauling, Josch K., et al. "Proposal for a common nomenclature for fragment ions in mass spectra of lipids." PloS one 12.11 (2017): e0188394.

[5] https://lifs.isas.de/

[6] https://www.ebi.ac.uk/ols/ontologies/ms

[7] MacLean, Brendan, et al. "Skyline: an open source document editor for creating and analyzing targeted proteomics experiments." Bioinformatics 26.7 (2010): 966-968.

T15 Support for additional platforms for CE Optimization

Due to slight differences in the way that vendor data is transformed, and specifically, how activation/collision dissociation details are reported in *.mzML, an adaptation of the existing codebase may be necessary.

Please request access to the necessary tools and support at https://lifs.isas.de/support.html if you want to add data for a specific new platform or if you wish to use your own reference data for an existing platform. The following sections describe the tables that are required for a) LipidCreator to read/use a custom model, and b) the tables used by flipR as input and those produced by it for import in LipidCreator. The required projects are available at https://github.com/lifs-tools/flipr.

T15.1 MS Instrument Table

The comma-separated MS instrument table for LipidCreator (data\ms-instruments.csv) contains PSI-MS controlled vocabulary terms for instrument model to identify individual MS platforms on each row. The following columns are required in this file:

- **CV_term**: the PSI-MS CV term id identifying the instrument.

- **model**: the PSI-MS CV term's description, e.g. the name of the instrument (term MS:1002523, 'Q Exactive HF').

- **min_CE**: the minimum collision energy covered by the instrument. Collision energies calculated by LipidCreator can not be lower than this threshold. The model calculated by flipR may exceed the range set by min_CE and max_CE.

- **max_CE**: the maximum collision energy covered by the instrument. Collision energies calculated by LipidCreator can not be higher than this threshold. The model calculated by flipR may exceed the range set by min_CE and max_CE.

- **x axis label**: the label for the collision energy axis, e.g. 'Collision Energy [eV]' or 'Relative Collision Energy'.

- **modes**: the modes to enable for the instrument platform, e.g. 'PRM' for PRM-only mode.

Example ('…' represent skipped lines in the file):

```
CV_term,model,min_CE,max_CE,x                       axis                        label,modes
...
MS:1002523,Thermo Scientific Q Exactive HF,10,60,Normalized Collision Energy,PRM
```

To feed data into the fragment intensity prediction (FIP) pipeline, MS/MS data needs to be available in *.mzML format. We recommend msConvert for the conversion from native vendor format into *.mzML.

msConvert may not always report the most precise term for the instrument, but rather a more generic term that identifies the instrument family.

A list of the available instruments is available via the OntologyLookupService[8].

T15.2 Transition Table

The transition table used for the CE model calculation can be generated using LipidCreator, similarly to how a transition list is created for Skyline. However, LipidCreator uses specific internal IDs for heavy labeled species, which requires the "developer" mode to create modified transition lists that allow a reimport of the parameter files after the parameter estimation step.

You can start LipidCreator from the command line in developer mode as follows:

```
LipidCreator.exe dev
```

This will ensure that precursor names for heavy-labeled instances are written with a specific placeholder, e.g. "{d8}" for 8 Deuterium atoms replacing 8 hydrogens. This allows the correct mapping of the corresponding model parameters in LipidCreator. Adduct names for heavy labeled instances are also exported unaltered in developer mode, not with the Skyline specific nomenclature, e.g. [M8H2-H]1-, to ensure correct mapping to the originating monoisotopic precursor. Precursor and product masses are correctly calculated for heavy labeled instances. The following columns are required in this file (following the format for Skyline transition lists[9]):

- **MoleculeGroup**

- **PrecursorName**

- **PrecursorFormula**

- **PrecursorAdduct**

- **PrecursorMz**

- **PrecursorCharge**

- **ProductName**

---

[8]https://www.ebi.ac.uk/ols/ontologies/ms/terms?iri=http%3A%2F%2Fpurl.obolibrary.org%2Fobo%2FMS_1000031

[9]https://skyline.ms/_webdav/home/software/Skyline/%40files/tutorials/SmallMolecule-3_6.pdf

- **ProductFormula**

- **ProductAdduct**

- **ProductMz**

- **ProductCharge**

- **Note**

Example ('…' represent skipped lines in the file):

```
MoleculeGroup PrecursorName PrecursorFormula    PrecursorAdduct     PrecursorMz
       PrecursorCharge      ProductName   ProductFormula      ProductAdduct ProductMz
       ProductCharge Note
…
PIP2   PIP2 17:0-20:4      C46H83O19P3   [M-H]1-      1031.46686541991    -1
       precursor     C46H83O19P3   [M-H]1-      1031.46686541991    -1
```

Please follow the transition list tutorial for LipidCreator to create a transition list for your target molecules (see). When you reach the step "LipidsReview", select "Store transition list", choose "tsv files (*.tsv)" as your output file format and select "No" when asked whether to split the output by polarity.

T15.3 Mapping Table (Transitions to measurements)

This tab-separated table defines the mapping from transitions to actual measurements and acts as the glue between LipidCreator transitions lists and the transition extraction and model training with flipR.

- **Instrument:** the PSI-MS CV term for the instrument, e.g. 'MS:1002523'.

- **MoleculeGroup:** the molecule group / class of the molecule, e.g. 'PIP2'.

- **PrecursorName:** the name of the precursor molecule, e.g. 'PIP2 17:0-20:4'.

- **PrecursorAdduct:** the precursor adduct, e.g. '[M-H]1-'.

- **PPMS:** a '|' (bar) separated list of ppms to use for transition / m/z matching, e.g. '5|10'.

- **File:** the source file for this instance, containing MS2 scans.

- **Group:** a group identifier to distinguish multiple measurements of the same molecule.

Example ('…' represent skipped lines in the file):

```
Instrument      MoleculeGroup PrecursorName  PrecursorAdduct      PPMS   File    Group
…
MS:1002523      PIP2    PIP2 17:0-20:4 [M-H]1- 5|10    measurements/QExHF03 NM 0001427.mzML 0001427
```

T15.4 Feature Table (Transitions applied to mzMLs)

The feature table (*-fip.tsv) is created by the transition extraction step and serves as the main data input for flipR. It holds one m/z feature per row, with the following columns, specifying its provenance, parameters and information that is used by downstream steps to maintain a mapping between input transition list and output parameter file for LipidCreator. The following columns are reported in the output file (some correspond to *.mzML elements / attributes, some come from LipidCreator / Skyline):

- **instrument**
- **localDateTimeCreated**
- **origin**
- **scanNumber**
- **polarity**
- **basePeakMz**
- **basePeakIntensity**
- **totalIonCurrent**
- **id**
- **scanDefinition**
- **msLevel**
- **isolationWindowTargetMz[0]**
- **isolationWindowLowerOffset[0]**
- **isolationWindowUpperOffset[0]**
- **precursorActivationType**
- **precursorCollisionEnergy**
- **precursorCollisionEnergyUnit**
- **ionInjectionTime[0]**
- **isolationMzMin[0]**
- **isolationMzMax[0]**
- **precursorCharge[0]**
- **precursorMz[0]**
- **msFunction**

- **retentionTime**

- **spectrumType**

- **rawTic**

- **group**

- **foundMass**

- **foundMassRange[ppm]**

- **foundMassLowerBound**

- **foundMassUpperBound**

- **foundMassError[ppm]**

- **foundIntensity**

- **scanRelativeIntensity**

- **calculatedMass**

- **species**

- **precursorAdduct**

- **fragment**

- **adduct**

Example (some columns omitted for brevity):

```
instrument    localDateTimeCreated      origin scanNumber    polarity      basePeakMz
    basePeakIntensity    totalIonCurrent      …       isolationWindowTargetMz[0]
    isolationWindowLowerOffset[0]    isolationWindowUpperOffset[0]
    precursorActivationType    precursorCollisionEnergy
    precursorCollisionEnergyUnit      …        rawTic group foundMass
    foundMassRange[ppm] foundMassLowerBound foundMassUpperBound foundMassError[ppm]
    foundIntensity       scanRelativeIntensity      calculatedMass      species
    precursorAdduct      fragment      adduct
MS:1002523   2018-11-21T08:02:11.851    QExHF03_NM_0001279.mzML   1      NEGATIVE
    317.2117523   1.8311354e07 2.737138e07   …      317.212188720703    0.25   0.25
    HCD    10.0    electronvolt        …        2.6918232E7   0001279
    299.2013854980469   5      299.19990399299996  299.20289600700005  -
0.04846886780363687 567272.8      0.021073926  299.2014     9-HEPE [M-H]1-
    299.201       [M-H]1-
```

S.2.15.5 LipidCreator Parameter Table (After model training and selection)

The comma-separated lipid creator parameter file (data\ce-parameters\MS_CVTERMID.csv,
e.g. **MS_1002523.csv** for the Thermo Scientific Q Exactive HF) contains collision energy

calculation parameters for each lipid class, as reported and concatenated by flipR. There are as many rows for each fragment, as there are parameters. The following columns are required in this file:

- **instrument:** the PSI-MS CV term id identifying the instrument, e.g. MS:**1002523 for** Thermo Scientific Q Exactive HF.
- **class:** the lipid class, e.g. 10-HDoHE, needs to be double quoted, when the name contains a comma.
- **adduct:** the precursor adduct for this lipid class, e.g. [M-H]1-.
- **fragment:** the fragment identifier. If no common name is available, use e.g. "m/z 121.0658". The precursor must be reported as "precursor".
- **ParKey:** the model parameter name, currently one of "model", "meanlog", "sdlog", "scale", and "shift".
- **ParValue:** the model parameter values, currently, for "model" only "dlnormPar" is recognized. Other parameters are expected to be reported as double numbers with a "." as the decimal separator.

Example ('…' represent skipped lines in the file):

```
instrument,class,adduct,fragment,ParKey,ParValue
…
MS:1002523,10-HDoHE,[M-H]1-,m/z 121.0658,meanlog,4.10313116901712
MS:1002523,10-HDoHE,[M-H]1-,m/z 121.0658,model,dlnormPar
MS:1002523,10-HDoHE,[M-H]1-,m/z 121.0658,scale,0.140234378276546
MS:1002523,10-HDoHE,[M-H]1-,m/z 121.0658,sdlog,0.512948788298359
MS:1002523,10-HDoHE,[M-H]1-,m/z 121.0658,shift,2.97212321271515
```