SUPPORTING INFORMATION

Mapping Lipid Fragmentation for Tailored Mass Spectral Libraries

Paul D. Hutchins[1,3], Jason D. Russell[2,3], and Joshua J. Coon[1,2,3,4,*]

[1]Department of Chemistry, University of Wisconsin–Madison, Madison, WI 53706, USA
[2]Morgridge Institute for Research, Madison, WI 53715, USA
[3]Genome Center of Wisconsin, Madison, WI 53706, USA
[4]Department of Biomolecular Chemistry, University of Wisconsin–Madison, Madison, WI 53706, USA

* Corresponding author: jcoon@chem.wisc.edu

SUPPORTING INFORMATION

Generation of Putative Identifications

For complex extract analysis, candidate lipid identifications are assigned to experimental MS/MS spectra using a modified dot-product spectral similarity score as described previously.[1] First, MS/MS spectra are converted to the Mascot Generic Format (MGF) and their intensities scaled relative to the most intense fragment in each spectrum. Each spectrum is then compared to all *in silico* lipid spectra whose precursor *m/*z values falls within a user-defined *m/*z tolerance of the experimental spectrum's parent ion *m/z*. To quantify spectral similarity between the *in silico* and experimental spectra, a modified dot-product score[2] is calculated which takes into account the correlation of fragment *m/z* values and fragment intensities between the spectra.

$$Dot\ Product = 1000 * \frac{\left(\sum A_{exp} * A_{lib}\right)^2}{\sum {A_{exp}}^2 * \sum {A_{lib}}^2}$$

$$A = (Relative\ Intensity)^m * (m/z)^n$$

As described previously,[1] the intensity (m = 1.2) and mass (n = 0.9) weighting factors used were optimized on a set of lipids from the NIST Tandem Mass Spectral Library to return the maximum number of correct identifications and the maximum score separation between the correct match and the top-scoring incorrect match. For reference standard analysis, identifications are provided by the user in the form of a table of lipid names and retention times.

Lipid MS/MS Processing

For each putative lipid identifications which rises above a user-supplied dot product threshold, the corresponding MS/MS scan is located and scaled to the most intense fragment ion. When analyzing reference standards, MS/MS spectra which fall within a user-supplied *m/*z and retention time tolerance are assigned to the provided lipid identification. All ions which fall below a relative

intensity threshold are removed to filter spectral noise. Additionally, the spectral signal-to-noise ratio is calculated using the ratio of the most intense fragment to the median intensity fragment. To remove low quality spectra which complicate fragmentation mapping, all spectra below a user-supplied S/N threshold are removed.

Spectral Clustering

Spectra which are assigned the same putative identification are subsequently clustered to produce consensus spectra. First, replicate spectra are ranked according to the S/N and the dot product is calculated between the top ranked (reference) and second ranked spectra. If this dot product score rises above a user-supplied threshold (see Supplementary Table 5) the spectra are clustered together. This process is repeated until the dot product between the reference spectrum and all others has been calculated. If any spectra remain unclustered, the unclustered spectrum with the highest S/N ratio becomes the new reference spectrum and the clustering process is repeated with the remaining spectra. If no spectra are clustered with the reference spectrum, the reference spectrum is assigned to a "cluster" with itself and the next unclustered spectrum becomes the new reference spectrum.. This clustering process is continued until no spectra remain. To remove MS/MS spectra which contain co-fragmented lipids or were incorrectly identified, the cluster with the largest number of spectra is retained for further analysis. If two clusters contain the same number of spectra, the cluster with the highest average identification dot product score (for complex sample analysis) or the lowest retention time deviation (for reference standard analysis) is retained. For the remaining cluster, a consensus spectrum is created from all fragments present in a user-defined percentage of the clustered spectra using the median *m/z* and relative intensity.

Generating Annotation Spectra

To reveal the underlying fragmentation pathways of each spectrum, Library Forge creates *m/z*-transformed spectra (annotation spectra) which make fragments which stem from the same fragmentation pathway isobaric. The type of transformation applied to each spectrum depends on the type and number of specific side chains present in the lipid as summarized in the table below:

Fragment Type	Transformation
Fragment	N/A
Neutral Loss	*m/z* - precursor *m/z*
Chain Fragment	*m/z* - chain *m/z*
Chain Neutral Loss	*m/z* - (precursor *m/z* - chain *m/z*)

Note that for lipids which contain two different side chains, two annotation spectra are created for the chain fragment and chain neutral loss types. To remove potentially spurious fragment annotations, each transformed *m/z* value must fall within the *m/z* ranges specified below:

Fragment Type	Rule(s)
Fragment	*m/z* < adduct + headgroup + backbone + H_2O_2
Neutral Loss	|*m/z*| < adduct + headgroup + backbone AND m/z < 0
Chain Fragment	|*m/z*| < (adduct + headgroup + backbone) / 2
Chain Neutral Loss	|*m/z*| < (adduct + headgroup + backbone) / 2

Next, all annotation spectra which stem from a unique lipid class/adduct combination are linked together. Lipid class/adduct combinations which do not contain sufficient unique spectra, as defined by the user, are removed and consensus annotation spectra are subsequently created for each lipid class/adduct combination.

Filtering Annotation Spectra

To remove spurious annotations, all transformed *m/z* values are filtered by the percent of annotation spectra in which they were observed. For chain-based annotation spectra, the

following equation is used to calculate this percentage as more than one chain-based spectra may be created if the lipid contains more than one of that type of chain.

$$Percent\ Identified = \frac{2 * \left(\frac{u_t}{c_t}\right) * f}{a}$$

$$u_t = unique\ seed\ chains\ of\ type\ t$$

$$c_t = number\ of\ chains\ of\ type\ t\ for\ lipid\ class$$

$$f = isobaric\ annotation\ fragments\ identified$$

$$a = total\ annotation\ spectra\ for\ given\ fragment\ type$$

If a given fragment's percent identified score rises above a user-supplied threshold it is retained and used for *in silico* library generation. Because of lipids modular construction, some artefactual fragment rules are generated via the creation of annotation spectra. A simple example of these artifacts is one which stems from the inherent relationship between chain fragments and chain neutral loss in diacyl lipids where a neutral loss of one fatty acid is equivalent to a fatty acid fragment of the other acyl chain plus the headgroup of the lipid. Accordingly, when a particular *m/z* from the original spectra maps to more than one fragment definition (collisions), the rule which was identified in a greater percentage of the annotation spectra and has a lower rule modifier *m/z* is retained and all other colliding fragment rules are removed. Once filtered, the median *m/z* and relative intensity are calculated and used for *in silico* library generation.

NIST 1950 Analysis and Validation

To validate the MS/MS identifications made using the *in silico* spectral libraries outlined in this manuscript, we generated an LC-MS/MS dataset consisting of triplicate analysis of the NIST 1950 SRM human plasma reference material. Extensively characterized in an interlaboratory study consisting of 31 laboratories, this reference material represents the best available method for

validating lipid identification pipelines. The LC-MS/MS data generated from analysis of the NIST 1950 SRM were processed using Compound Discoverer 2.0 (Thermo Scientific). First, all MS1 scans were extracted for feature detection (Select Spectra node) and the retention times were aligned (Align Retention Times node) across the LC-MS/MS experiments. Chromatographic features were then detected from the aligned scans (Detect Unknown Compounds node) and grouped according to common m/z and elution profile between experiments (Group Unknown Compounds node). Each LC-MS/MS experiment was then re-searched to reduce missing values in the dataset (Fill Gaps node). Feature groups whose area was less than three times the area in the blank injection were removed (Mark Background Compounds node). To generate an unaligned dataset for extraction of the unaligned retention times, MS1 scans were extracted (Select Spectra node) and features were detected directly from the unaligned scans (Detect Unknown Compounds node).

MS/MS Lipid identifications were generated using LipiDex to search spectra against the generated *in silico* libraries and map these identifications onto chromatographic features (see Supplementary Table 6 for parameters used). No additional manual filtering was applied to the dataset. When necessary, lipid identifications were converted to their sum identification (e.g. PC 16:0_18:1 to PC 34:1) to permit comparison to the reported consensus NIST 1950 lipid identifications. Lipid identifications were considered "validated" if the identification was reported by at least three laboratories in the interlaboratory study.

Lipid Data Analyzer 2 Dataset

To ensure that the successful generation of in silico lipid spectra in Library Forge did not stem from the original putative identifications having been generated by the same program we created a seed lipid identification set using a different lipid software: Lipid Data Analyzer 2 (LDA 2). Lipids from the Q Exactive HF HAP1 dataset were identified in LDA 2 using precursor mass matching parameters of 2 matching isotopes, 0.1% relative base peak cutoff, the Q Exactive positive and

negative mass lists from the Biological Experiment dataset (http://genome.tugraz.at/lda2/data/Biological/Bio_QExactive.zip), and the OrbiTrap_exactive_neg and OrbiTrap_exactive_MS/MS rule sets. All lipid identification report sheets were then reformatted into the LipiDex MS/MS identification result sheet format and imported into Library Forge for subsequent analysis. *In silico* libraries were generated in Library Forge using a target *m/z* tolerance of 15 ppm, minimum cluster dot product of 700, 3 minimum unique IDs, a fragment *m/z* tolerance of 0.02 Th, min. consensus spectra percent of 0%, minimum relative intensity of 15, minimum annotation spectra percent of 45% and a minimum MS/MS signal-to-noise ratio of 25.

Supplementary Table 1. Library Generation Standards. Show here is a list of all lipid reference standards used for library generation in Library Forge.

Lipid	Vendor	Mixture Name
CDP DG 18:1_18:1	Avanti	-
CE 16:0(d7)	Avanti	-
Cer[NS] d18:1_17:0	Avanti	-
Cer1P d18:1_16:0	Avanti	-
DG 12:0_12:0	Avanti	-
DG 18:1_12:0	Avanti	-
LysoSM 18:1	Cayman	-
PC 15:0_15:0	Avanti	-
PS 14:0_14:0	Avanti	-
S1P d18:1	Cayman	-
SM d18:1_17:0	Avanti	-
TG 18:1_18:1_18:1	Avanti	-
TG 20:0_18:1_16:0	Larodan	-
CL 15:0_15:0_15:0_16:1	Avanti	Cardiolipin Mix I
CL 22:1_22:1_22:1_14:1	Avanti	Cardiolipin Mix I
CL 24:1_24:1_24:1_14:1	Avanti	Cardiolipin Mix I
Cer[NS] d18:1_12:0	Avanti	Cer/Sph Mixture I
Cer[NS] d18:1_25:0	Avanti	Cer/Sph Mixture I
Cer1P d18:1_12:0	Avanti	Cer/Sph Mixture I
GlcCer[NS] d18:1_12:0	Avanti	Cer/Sph Mixture I
LacCer[NS] d18:1_12:0	Avanti	Cer/Sph Mixture I
S1P d17:0	Avanti	Cer/Sph Mixture I
S1P d17:1	Avanti	Cer/Sph Mixture I
SM d18:1_12:0	Avanti	Cer/Sph Mixture I
SP d17:0	Avanti	Cer/Sph Mixture I
SP d17:1	Avanti	Cer/Sph Mixture I
PS 14:1_14:1	Avanti	Ion Mobility Kit
PE 14:1_14:1	Avanti	Ion Mobility Kit
PG 14:1_14:1	Avanti	Ion Mobility Kit
PI 14:1_14:1	Avanti	Ion Mobility Kit
PA 14:1_14:1	Avanti	Ion Mobility Kit
PC 14:1_14:1	Avanti	Ion Mobility Kit
CL 14:1_14:1_14:1_15:1	Avanti	Ion Mobility Kit
DG 14:1_14:1	Avanti	Ion Mobility Kit
CE 19:0	Avanti	Ion Mobility Kit
LysoPC 18:1	Avanti	Ion Mobility Kit
Cer[NS] d18:1_18:1	Avanti	Ion Mobility Kit
SM d18:1_18:1	Avanti	Ion Mobility Kit
TG 18:0_18:0_18:0	Avanti	Ion Mobility Kit
CE 16:0	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
CE 16:1	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
CE 18:1	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
CE 18:2	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
CE 20:3	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
CE 20:4	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix

CE 20:5	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
CE 22:6	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
Cer[NS] d18:1_16:0	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
DG 16:0_16:0	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
DG 16:0_18:0	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
DG 16:0_18:1	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
DG 16:0_18:2	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
DG 16:0_18:3	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
DG 16:0_20:4	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
DG 16:0_20:5	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
DG 16:0_22:6	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
Cer[DS] d18:1_16:0	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
HexCer[NS] d18:1_16:0	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
LacCer[NS] d18:1_16:0	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
LysoPC 16:0	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
LysoPE 18:0	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
PC 16:0_16:1	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
PC 16:0_18:1	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
PC 16:0_18:2	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
PC 16:0_18:3	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
PC 16:0_20:3	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
PC 16:0_20:4	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
PC 16:0_20:5	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
PC 16:0_22:6	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
PE 18:0_18:1	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
PE 18:0_18:2	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
PE 18:0_18:3	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
PE 18:0_20:3	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
PE 18:0_20:4	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
PE 18:0_20:5	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
PE 18:0_22:5	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
PE 18:0_22:6	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
SM d18:1_16:0	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
SM d18:1_18:0	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
SM d18:1_24:0	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
SM d18:1_24:1	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
TG 50:1	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
TG 52:1	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
TG 52:2	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
TG 52:3	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
TG 52:4	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
TG 54:4	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
TG 54:5	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix
TG 56:7	AB Sciex	Lipidyzer System Suitability Lipid Classes Light Mix

Supplementary Table 2. Library Validation Standards. Show here is a list of all heavy isotope-labeled lipid reference standards used for library validation.

Lipid	Manufacturer	Mixture Name
CE 18:1(d7)	Avanti	SPLASH Standard
DG 15:0_18:1(d7)	Avanti	SPLASH Standard
Cholesterol(d7)	Avanti	SPLASH Standard
LysoPC 18:1(d7)	Avanti	SPLASH Standard
LysoPE 118:1(d7)	Avanti	SPLASH Standard
PA 15:0_18:1(d7)	Avanti	SPLASH Standard
PC 15:0_18:1(d7)	Avanti	SPLASH Standard
PE 15:0_18:1(d7)	Avanti	SPLASH Standard
PG 15:0_18:1(d7)	Avanti	SPLASH Standard
PI 15:0_18:1(d7)	Avanti	SPLASH Standard
PS 15:0_18:1(d7)	Avanti	SPLASH Standard
SM d18:1_18:1(d7)	Avanti	SPLASH Standard
TG 15:0_15:0_18:1(d7)	Avanti	SPLASH Standard
MG 18:1(d7)	Avanti	SPLASH Standard
dCer[NS] 16:0	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dCE 16:0	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dCE 16:1	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dCE 18:1	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dCE 18:2	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dCE 20:3	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dCE 20:4	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dCE 20:5	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dCE 22:6	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dDG 16:0/16:0	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dDG 16:0/18:0	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dDG 16:0/18:1	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dDG 16:0/18:2	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dDG 16:0/18:3	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dDG 16:0/20:4	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dDG 16:0/20:5	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dDG 16:0/22:6	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dFFA 16:0	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dFFA 17:1	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dLPC 16:0	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dLPE 18:0	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dPC 16:0/16:1	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dPC 16:0/18:1	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dPC 16:0/18:2	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dPC 16:0/18:3	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dPC 16:0/20:3	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dPC 16:0/20:4	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dPC 16:0/20:5	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dPC 16:0/22:4	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dPC 16:0/22:5	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dPC 16:0/22:6	AB Sciex	Internal Standards Kit for Lipidyzer Platform

dPE 18:0/18:1	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dPE 18:0/18:2	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dPE 18:0/18:3	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dPE 18:0/20:3	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dPE 18:0/20:4	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dPE 18:0/20:5	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dPE 18:0/22:5	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dPE 18:0/22:6	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dSM 16:0	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dSM 18:1	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dSM 24:0	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dSM 24:1	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dTG 50:1-FA16:0	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dTG 52:1-FA18:0	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dTG 52:2-FA18:1	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dTG 52:3-FA18:2	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dTG 52:4-FA18:3	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dTG 54:4-FA20:3	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dTG 54:5-FA20:4	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dTG 56:7-FA22:6	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dDCER 16:0	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dHCER 16:0	AB Sciex	Internal Standards Kit for Lipidyzer Platform
dLCER 16:0	AB Sciex	Internal Standards Kit for Lipidyzer Platform

Supplementary Table 3. LC-MS/MS Analysis Parameters for Q Exactive HF Analysis**.** Detailed data collection and parameters for the HAP1 and reference standard dataset collected on the Q Exactive HF platform (Thermo) and presented in the manuscript.

Q Exactive HF HAP1 Analysis

Positive Polarity

Parameter	Value
Polarity	positive
Full MS Resolution	15000
Full MS AGC Target	3.00E+06
Full MS Max. IT (ms)	50
Full MS Scan Range	*m/z* 500 - 1600
MS/MS Resolution	120,000
MS/MS AGC Target	1.00E+06
MS/MS Max. IT (ms)	250
MS/MS Loop Count	4
MS/MS Isolation Window	0.4
MS/MS Stepped NCE	20, 25
Dynamic Exclusion (s)	10.0

Negative Polarity

Parameter	Value
Polarity	negative
Full MS Resolution	15000
Full MS AGC Target	3.00E+06
Full MS Max. IT (ms)	50
Full MS Scan Range	*m/z* 500 - 1600
MS/MS Resolution	120,000
MS/MS AGC Target	1.00E+06
MS/MS Max. IT (ms)	250
MS/MS Loop Count	4
MS/MS Isolation Window	0.4
MS/MS Stepped NCE	20, 25
Dynamic Exclusion (s)	10.0

HAP1 Analysis

Gradient (400 uL/min)

Time	%B
0.000	2
2.000	2
5.000	30
6.000	50
20.00	85
21.00	99
28.00	99
28.25	2
30.00	2
-	-

Q Exactive HF Standard Analysis

Positive Polarity

Parameter	Value
Polarity	positive
Full MS Resolution	30000
Full MS AGC Target	3.00E+06
Full MS Max. IT (ms)	100
Full MS Scan Range	*m/z* 200 - 1400
MS/MS Resolution	15,000
MS/MS AGC Target	1.00E+06
MS/MS Max. IT (ms)	400
MS/MS Loop Count	5
MS/MS Isolation Window	1.4
MS/MS Stepped NCE	20, 25
Dynamic Exclusion (s)	10.0

Negative Polarity

Parameter	Value
Polarity	negative
Full MS Resolution	30000
Full MS AGC Target	3.00E+06
Full MS Max. IT (ms)	100
Full MS Scan Range	*m/z* 200 - 1400
MS/MS Resolution	15,000
MS/MS AGC Target	1.00E+06
MS/MS Max. IT (ms)	400
MS/MS Loop Count	5
MS/MS Isolation Window	1.4
MS/MS Stepped NCE	20, 25
Dynamic Exclusion (s)	10.0

Standard Analysis

Gradient (300 uL/min)

Time	%B
0.000	10
2.500	10
3.250	40
22.00	50
23.50	75
36.00	80
38.00	99
42.00	99
42.375	10
45.000	10

Supplementary Table 4. LC-MS/MS Analysis Parameters for Orbitrap Fusion Lumos Standard CAD analysis**.** Detailed data collection and parameters for the reference standard dataset collected on the Orbitrap Fusion Lumos platform (Thermo) and presented in the manuscript.

Orbitrap Fusion Lumos
Standard Analysis

Positive Polarity		**Negative Polarity**	
Parameter	Value	Parameter	Value
Polarity	positive	Polarity	positive
Full MS Resolution	30000	Full MS Resolution	30000
Full MS AGC Target	1.00E+06	Full MS AGC Target	1.00E+06
Full MS Max. IT (ms)	50	Full MS Max. IT (ms)	50
Full MS Scan Range	*m/z* 200 - 1400	Full MS Scan Range	*m/z* 200 - 1400
MS/MS Resolution	120,000	MS/MS Resolution	120,000
MS/MS AGC Target	5.00E+05	MS/MS AGC Target	5.00E+05
MS/MS Max. IT (ms)	400	MS/MS Max. IT (ms)	400
MS/MS Isolation Window	2.1	MS/MS Isolation Window	2.1
Collision Energy (%)	35	Collision Energy (%)	35
Activation Time (ms)	10	Activation Time (ms)	10
Activation Q	0.25	Activation Q	0.25
Dynamic Exclusion (s)	10.0	Dynamic Exclusion (s)	10.0

Standard Analysis
Gradient (300 uL/min)

Time	%B
0.000	10
2.500	10
3.250	40
22.00	50
23.50	75
36.00	80
38.00	99
42.00	99
42.375	10
45.000	10
0.000	10
2.500	10

Supplementary Table 5. Library Forge Processing Parameters**.** Detailed data processing parameters for the library generation from LipiDex identifications presented in the manuscript.

QE HF HCD HAP1 Analysis

Parameter	Value
Target m/z Tolerance (ppm)	15
Target Retention Tolerance (min)	1
Minimum ID Dot Product	200
Minimum Cluster Dot Product	700
Minimum Unique IDs	2
Fragment m/z Tolerance (Th)	0.01
Min. Consensus Spectra %	0
Minimum Relative Intensity	15
Min. Annotation Spectra %	65
Min. MS/MS Signal-to-Noise	50

QE HF HCD Standards Analysis

Parameter	Value
Target m/z Tolerance (ppm)	15
Target Retention Tolerance (min)	1
Minimum ID Dot Product	200
Minimum Cluster Dot Product	700
Minimum Unique IDs	*2*
Fragment m/z Tolerance (Th)	0.01
Min. Consensus Spectra %	0
Minimum Relative Intensity	5
Min. Annotation Spectra %	65
Min. MS/MS Signal-to-Noise	50

QE HF HCD HAP1 Analysis
LipidBlast Seed Library

Parameter	Value
Target m/z Tolerance (ppm)	15
Target Retention Tolerance (min)	1
Minimum ID Dot Product	50
Minimum Cluster Dot Product	700
Minimum Unique IDs	2
Fragment m/z Tolerance (Th)	0.01
Min. Consensus Spectra %	0
Minimum Relative Intensity	20
Min. Annotation Spectra %	65
Min. MS/MS Signal-to-Noise	100

Fusion Lumos CAD Standard Analysis

Parameter	Value
Target m/z Tolerance (ppm)	20
Target Retention Tolerance (min)	1
Minimum ID Dot Product	200
Minimum Cluster Dot Product	700
Minimum Unique IDs	*2*
Fragment m/z Tolerance (Th)	0.02
Min. Consensus Spectra %	0
Minimum Relative Intensity	20
Min. Annotation Spectra %	65
Min. MS/MS Signal-to-Noise	100

Supplementary Table 6.
Detailed Compound Discoverer and LipiDex processing parameters for the NIST 1950 validation dataset presented in the manuscript.

Select Spectra

Parameter	Value
Precursor Selection	Use MS(n - 1) Precursor
Lower RT Limit	1.4
Upper RT Limit	21
Scan Type	Is Full
Polarity Mode	Is +;-
S/N Threshold (FT-only)	1.5
Unrecognized Charge Replacement	1
Unrecognized Mass Analyzer Replacement	FTMS
Unrecognized MS Order Replacements	MS1
Unrecognized Activation Type Replacements	HCD
Unrecognized Polarity Replacements	+
Unrecognized MS Resolution Replacements	60000
Unrecognized MSn Resolution Replacements	30000

Align Retention Times

Parameter	Value
Alignment Model	Adaptive Curve
Alignment Fallback	Use Linear Model
Maximum Shift [min]	0.5
Shift Reference File	TRUE
Mass Tolerance	20 ppm
Remove Outlier	TRUE

Detect Unknown Compounds

Parameter	Value
Mass Tolerance (ppm)	10 ppm
Intensity Tolerance (%)	100
S/N Threshold	3
Min. Peak Intensity	100000
Ions	$[M+H]^{+}1$; $[M-H]^{-}1$
Min. Element Counts	C H
Max. Element Counts	C100 H190 N10 Na2 O15 P2 S2

Group Unknown Compounds

Parameter	Value
Mass Tolerance	10 ppm
RT Tolerance [min]	0.2
Rule #1	Unspecified
Rule #2	Unspecified
Preferred MS Order	MS1
Preferred Ion	$[M+H]^{+}1$

Fill Gaps

Parameter	Value
Mass Tolerance	10 ppm
RT Tolerance [min]	0.2
S/N Threshold	1.5

Mark Background Compounds

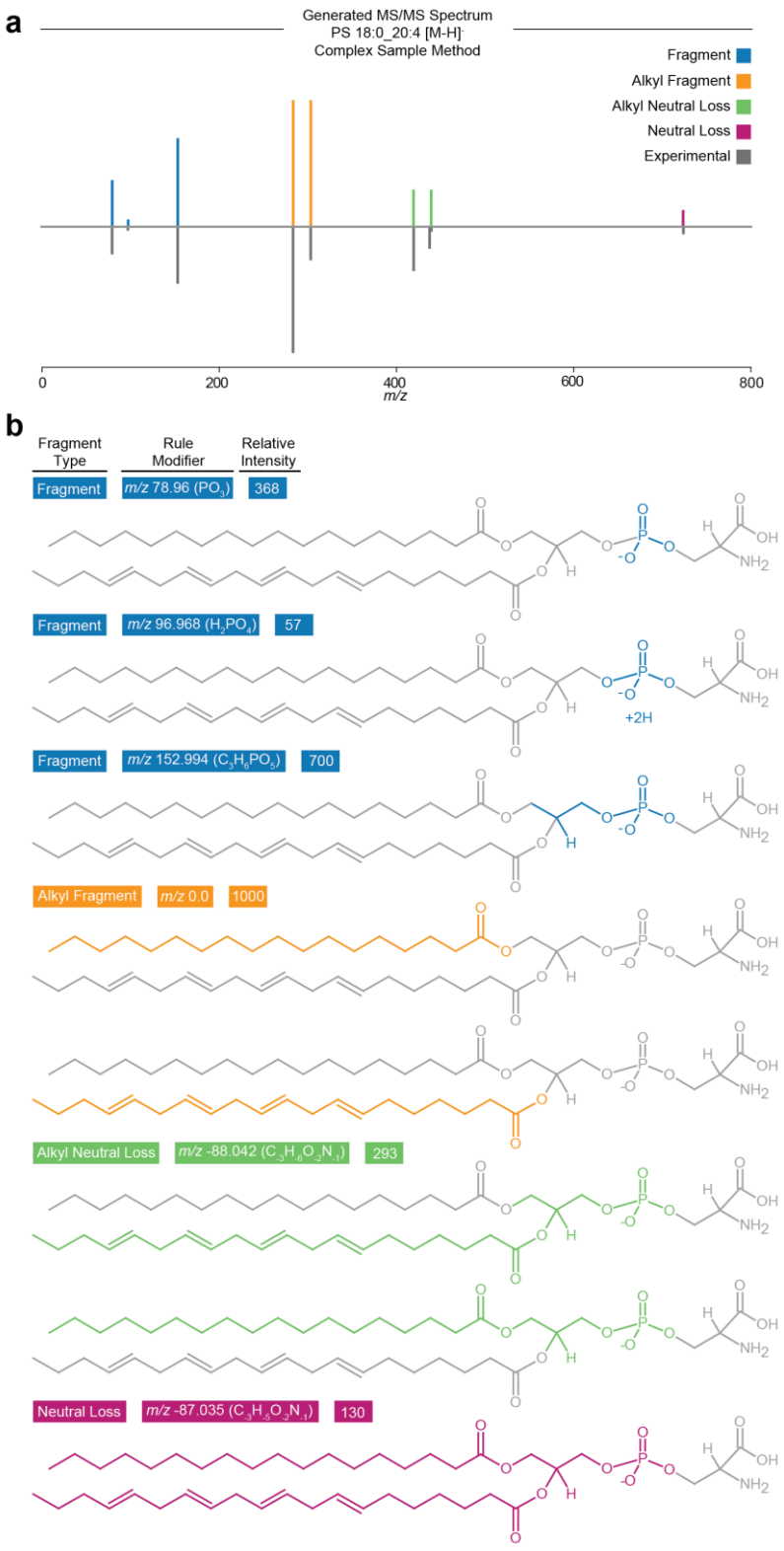
Parameter	Value
Max. Sample/Blank	3
Max. Blank/Sample	0
Hide Background	TRUE

Spectrum Searcher (LipiDex)

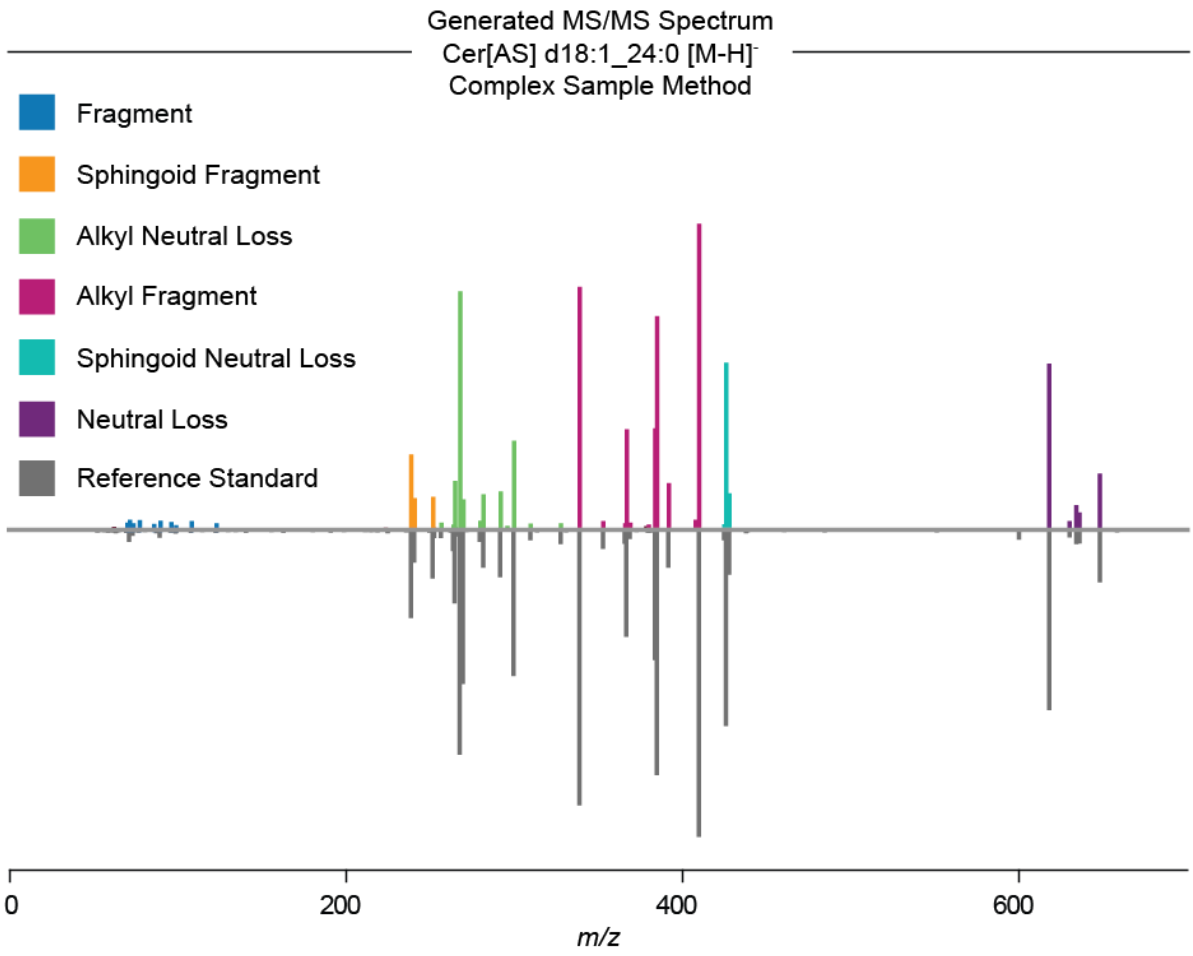
Parameter	Value
MS1 Search Tolerance	0.01 Th
MS2 Search Tolerance	0.02 Th
MS2 Low Mass Cutoff	61 Th

Peak Finder (LipiDex)

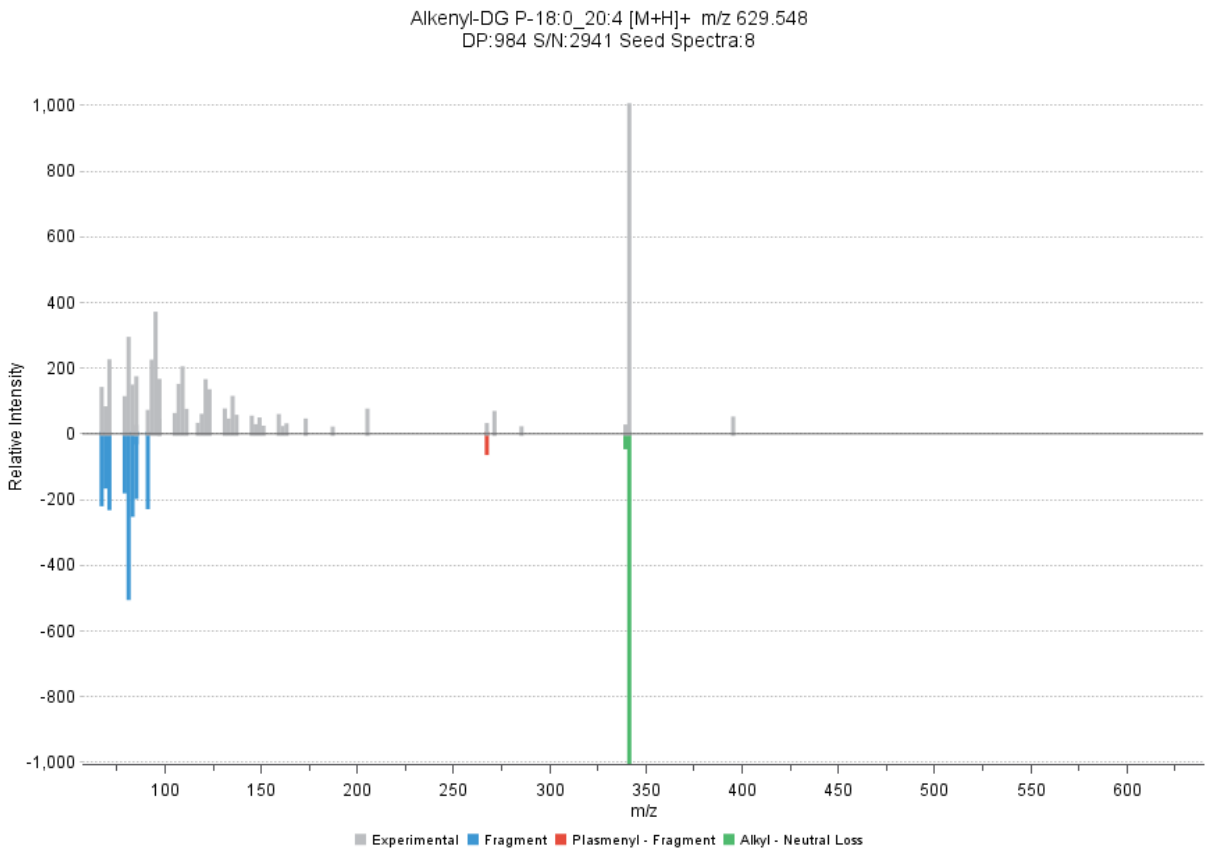
Parameter	Value
File Type	Compound Discoverer
Min. Lipid Spectral Purity	75%
Min. MS2 Search Dot Prod.	500
Min MS2 Search Rev. Dot Product	700
FWHM Window Multiplier	2
Max. Mass Difference	15 ppm
Max. RT M.A.D. Factor	3.5
Feature Found in n Files	1



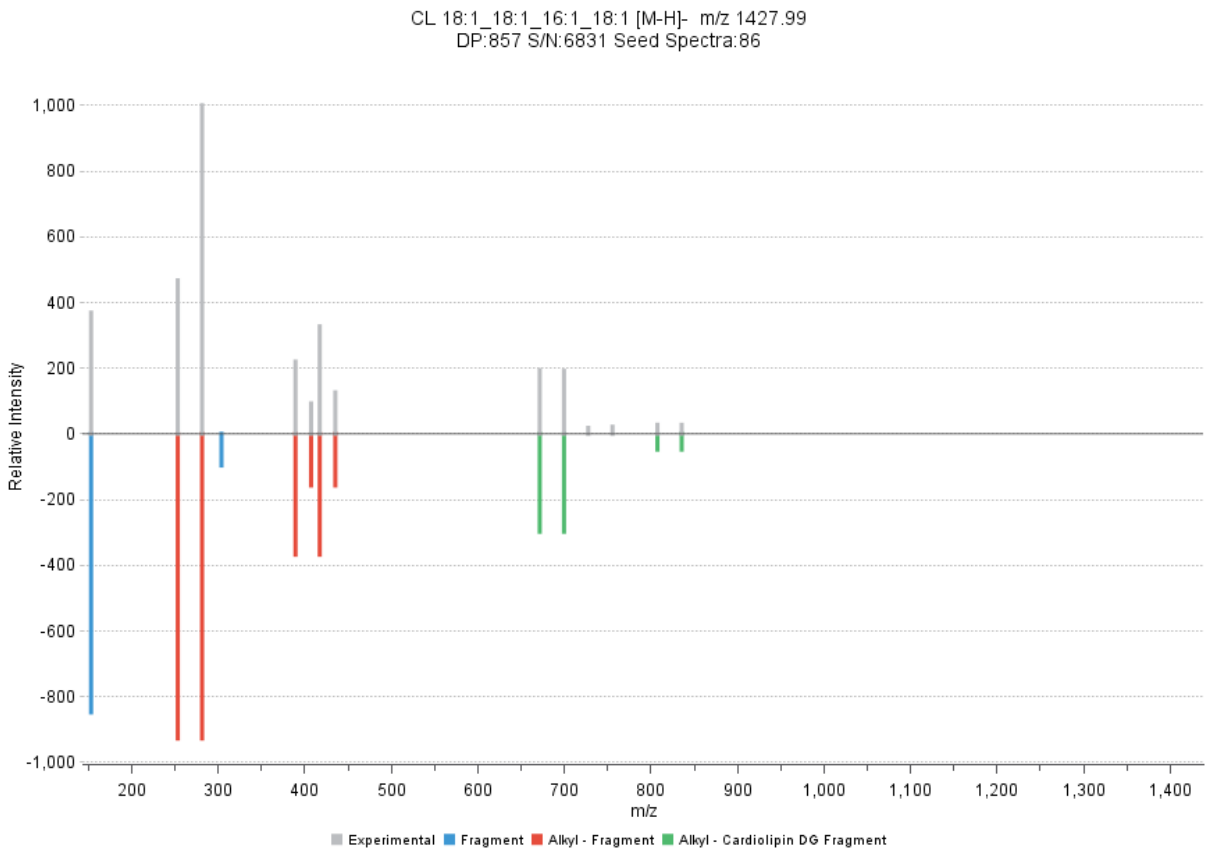
Supplementary Figure 1: (a) *In-silico* generated and experimental spectrum for PS 18:0_20:4 [M-H]$^-$. (b) Structural representation of PS 18:0_20:4 [M-H]$^-$ fragments (colored atoms). Note these fragment structures are used for depiction of fragment types and are not meant to reflect actual gas-phase structures.



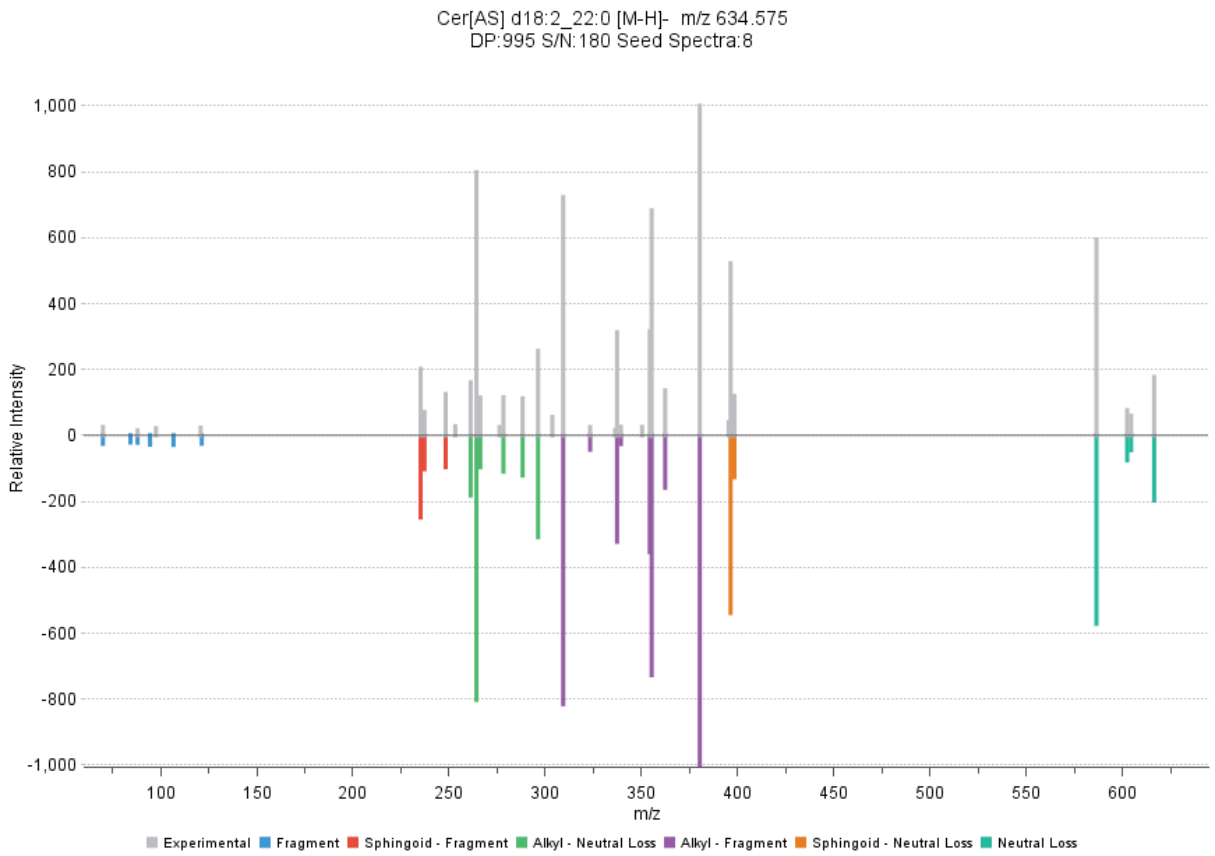
Supplementary Figure 2: Cer[AS] Reference Standard Spectrum. Reference standard MS/MS spectra (dark grey) and generated in-silico (colored) spectra for Cer[AS] d18:1_24:0 [M-H]-.



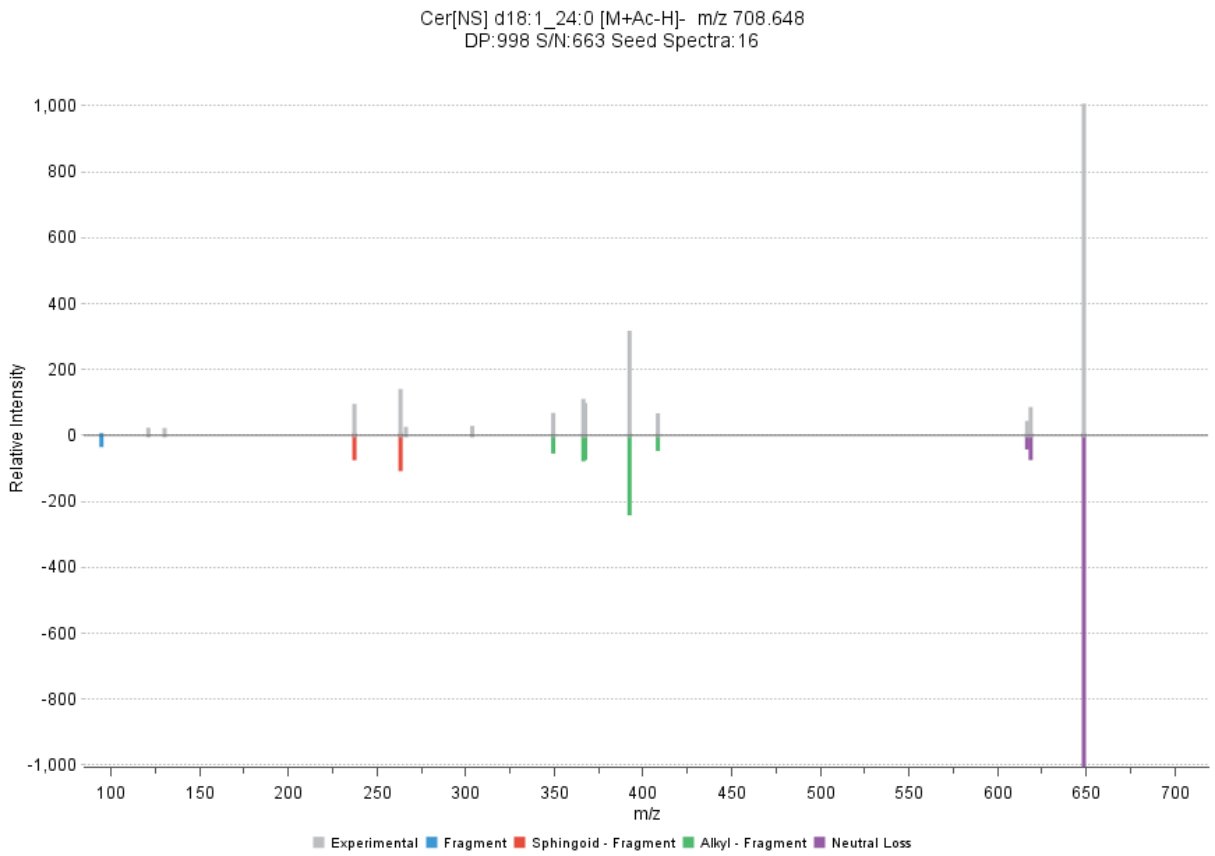
Supplementary Figure 3: Representative experimental (grey) and *in silico* (colored) Alkenyl-DG $[M+H]^+$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



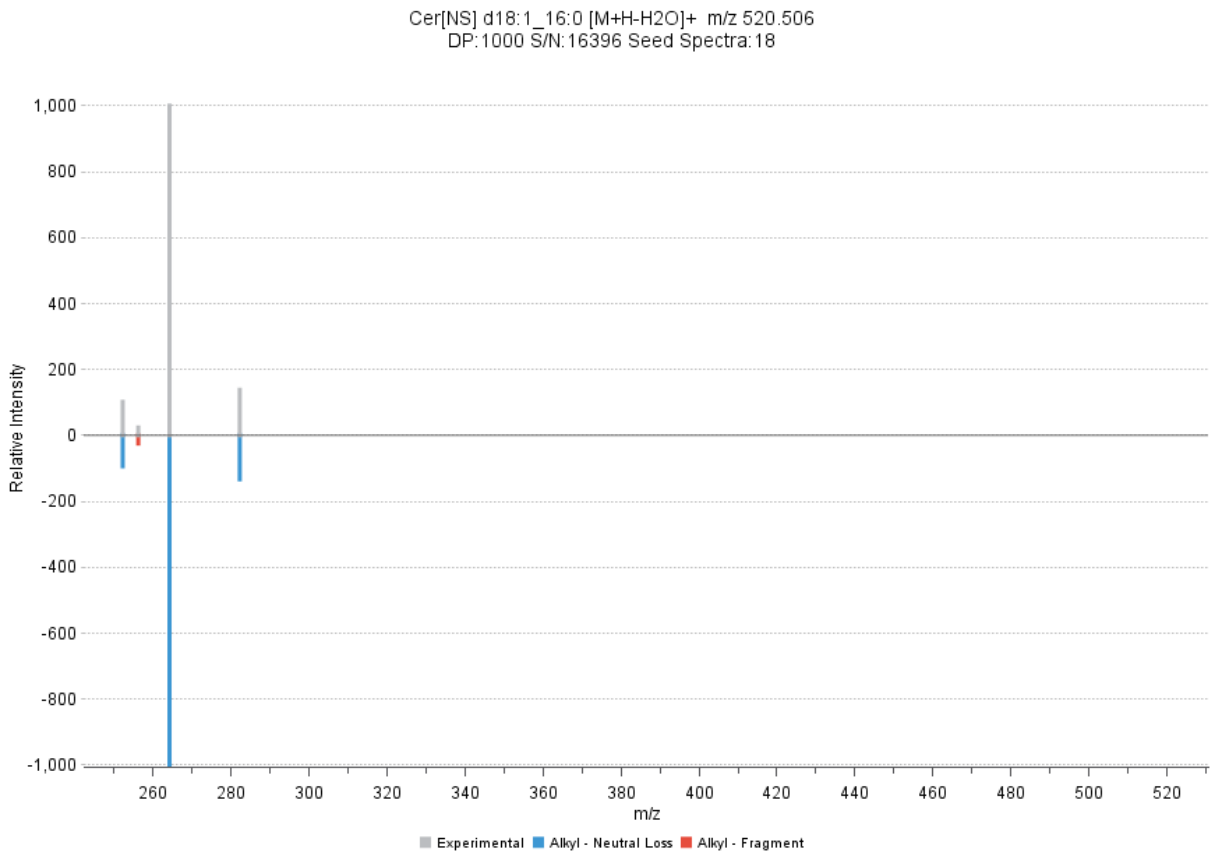
Supplementary Figure 4: Representative experimental (grey) and *in silico* (colored) CL $[M-H]^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



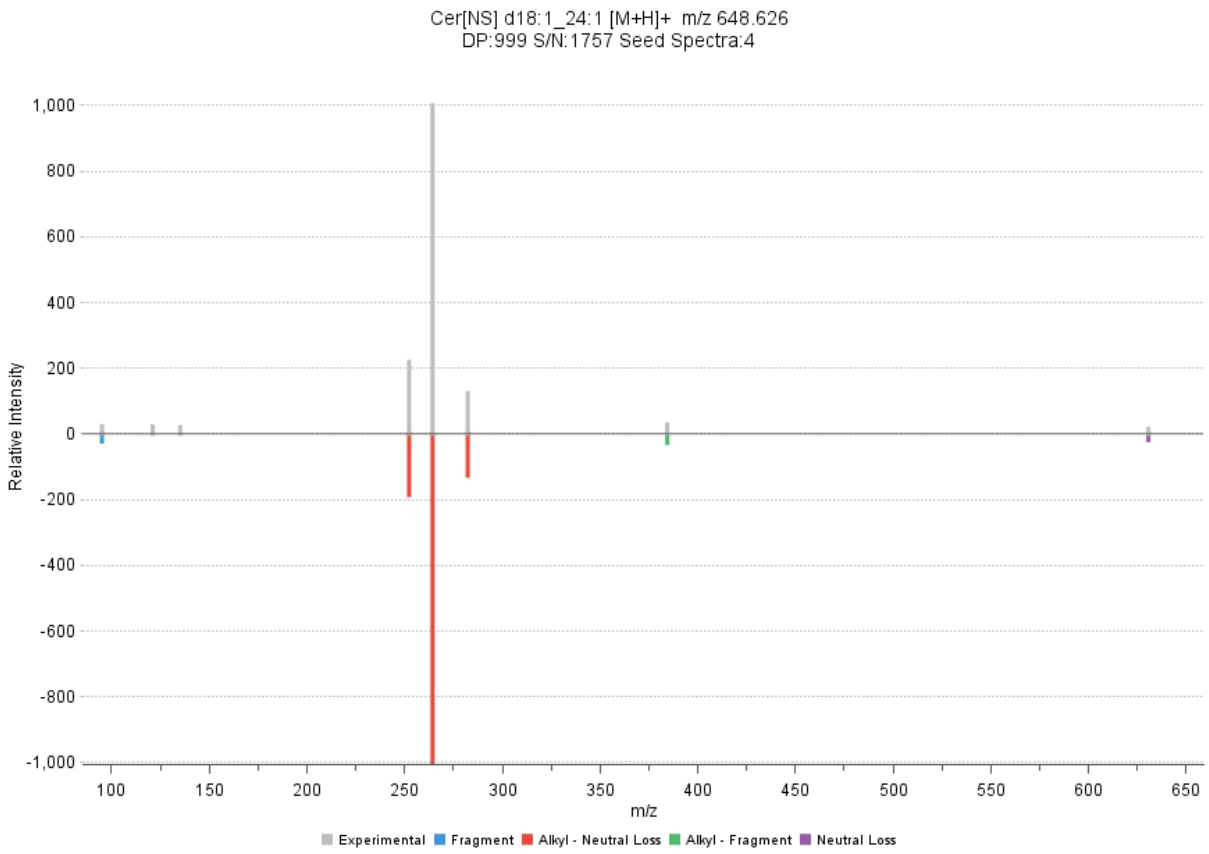
Supplementary Figure 5: Representative experimental (grey) and *in silico* (colored) Cer[AS] $[M-H]^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



Supplementary Figure 6: Representative experimental (grey) and *in silico* (colored) Cer[NS] $[M+Ac-H]^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.

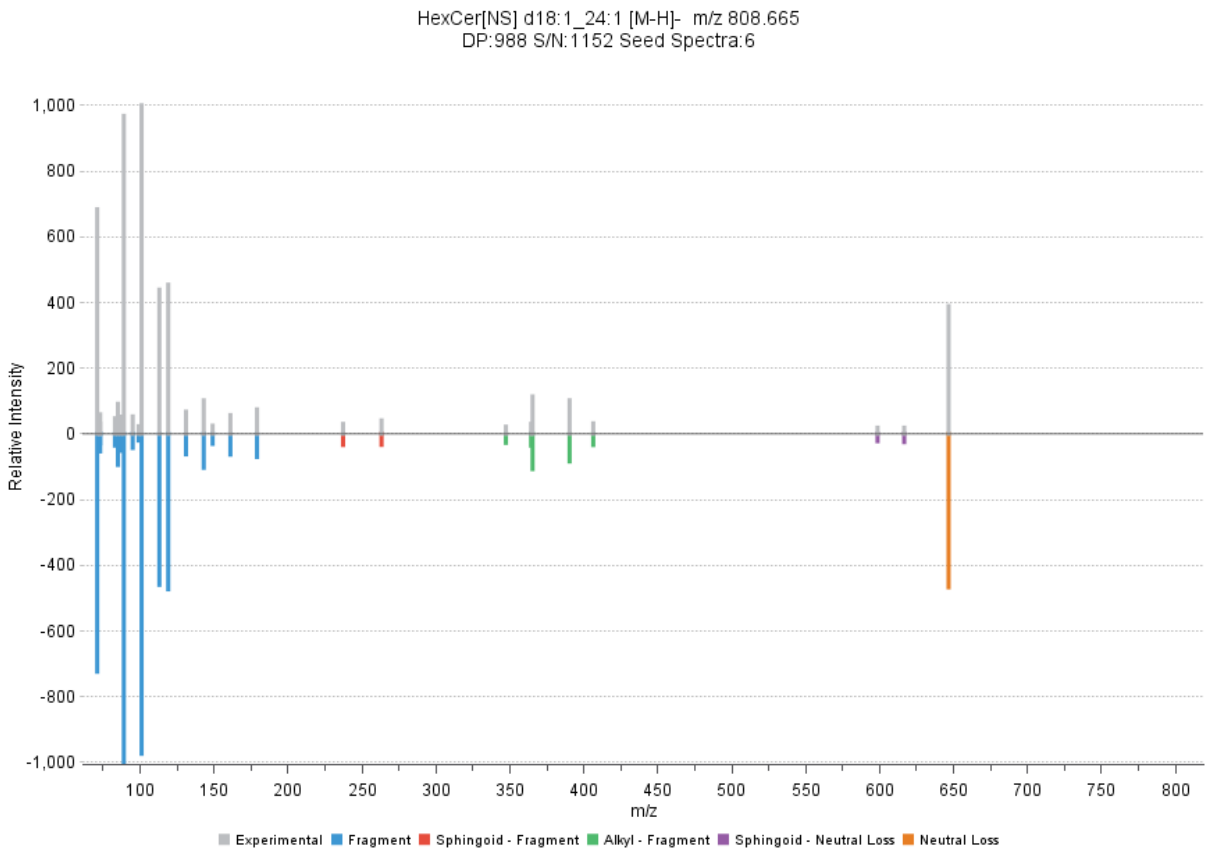


Supplementary Figure 7: Representative experimental (grey) and *in silico* (colored) Cer[NS] $[M+H-H_2O]^+$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.

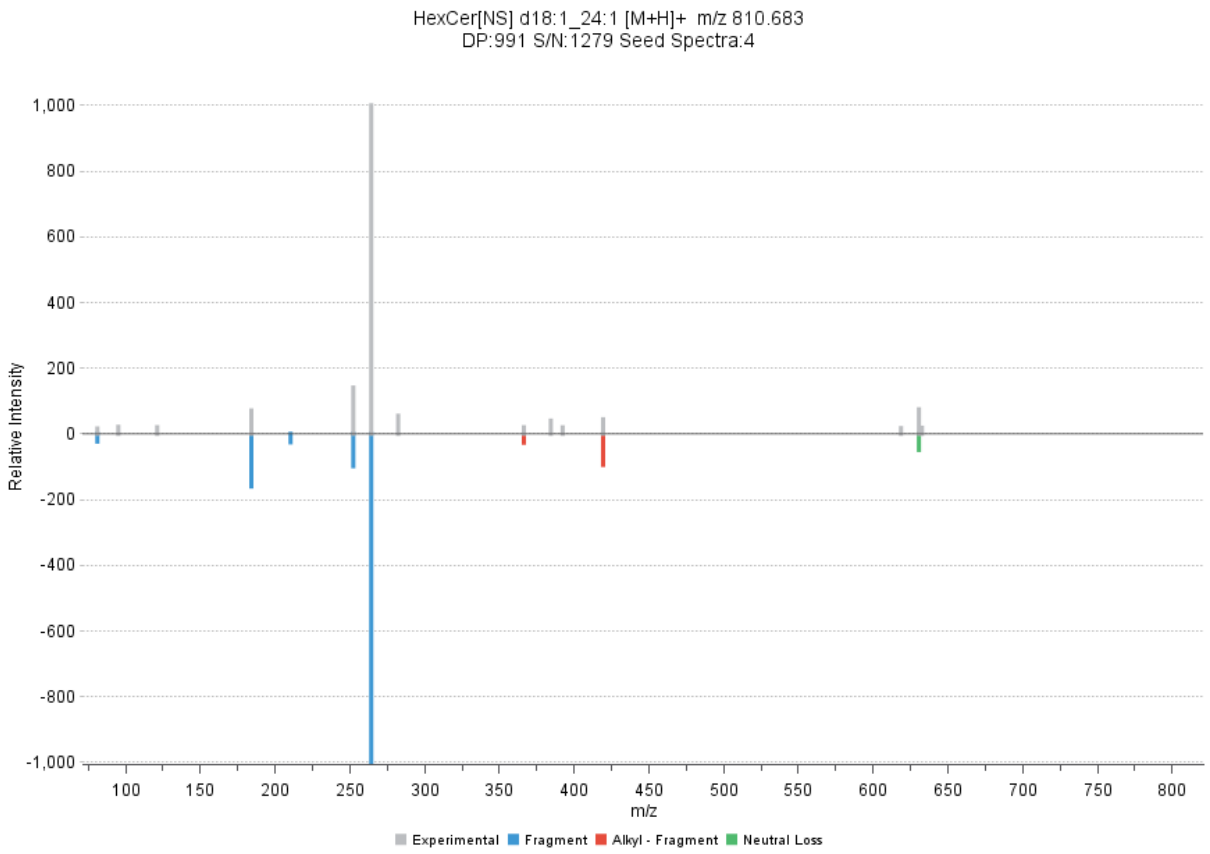


Supplementary Figure 8: Representative experimental (grey) and *in silico* (colored) Cer[NS] $[M+H]^+$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.

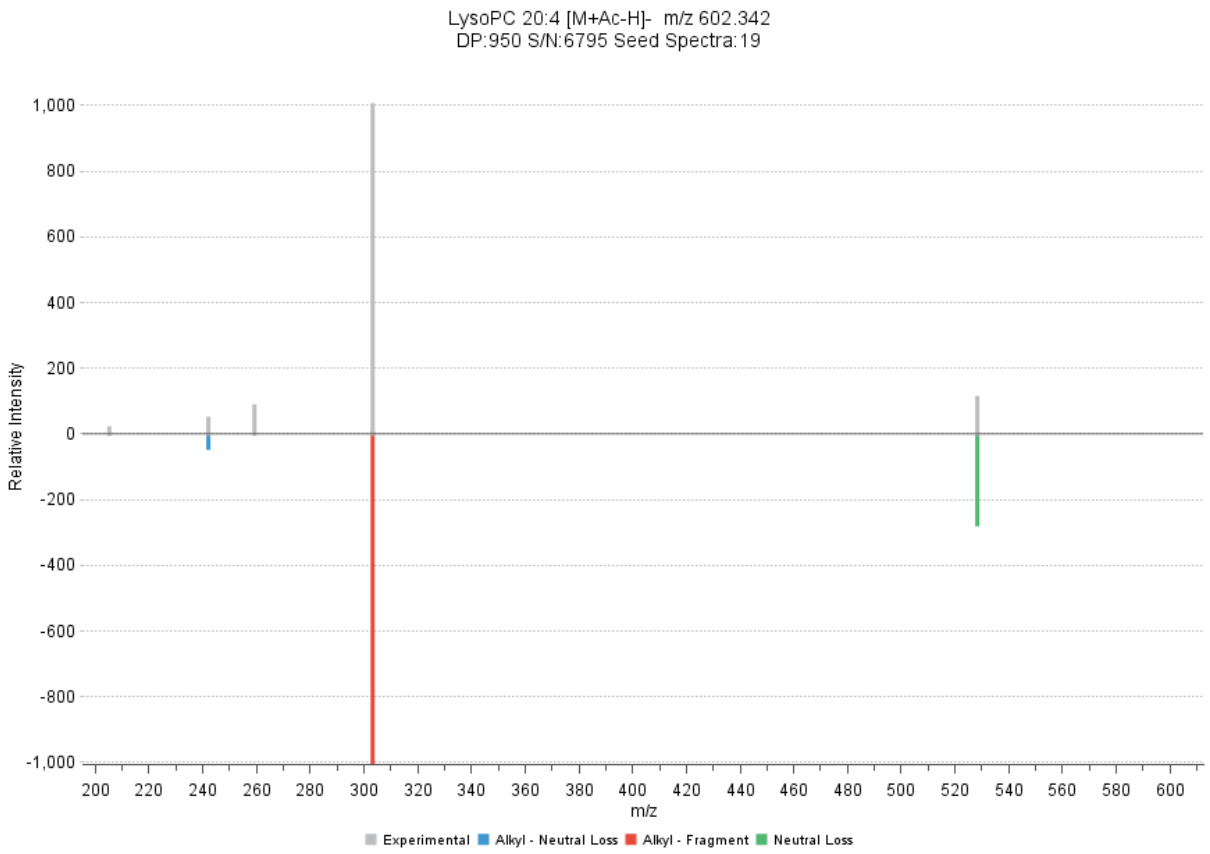
Supplementary Figure 9: Representative experimental (grey) and *in silico* (colored) Cer[NS] [M-H]$^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



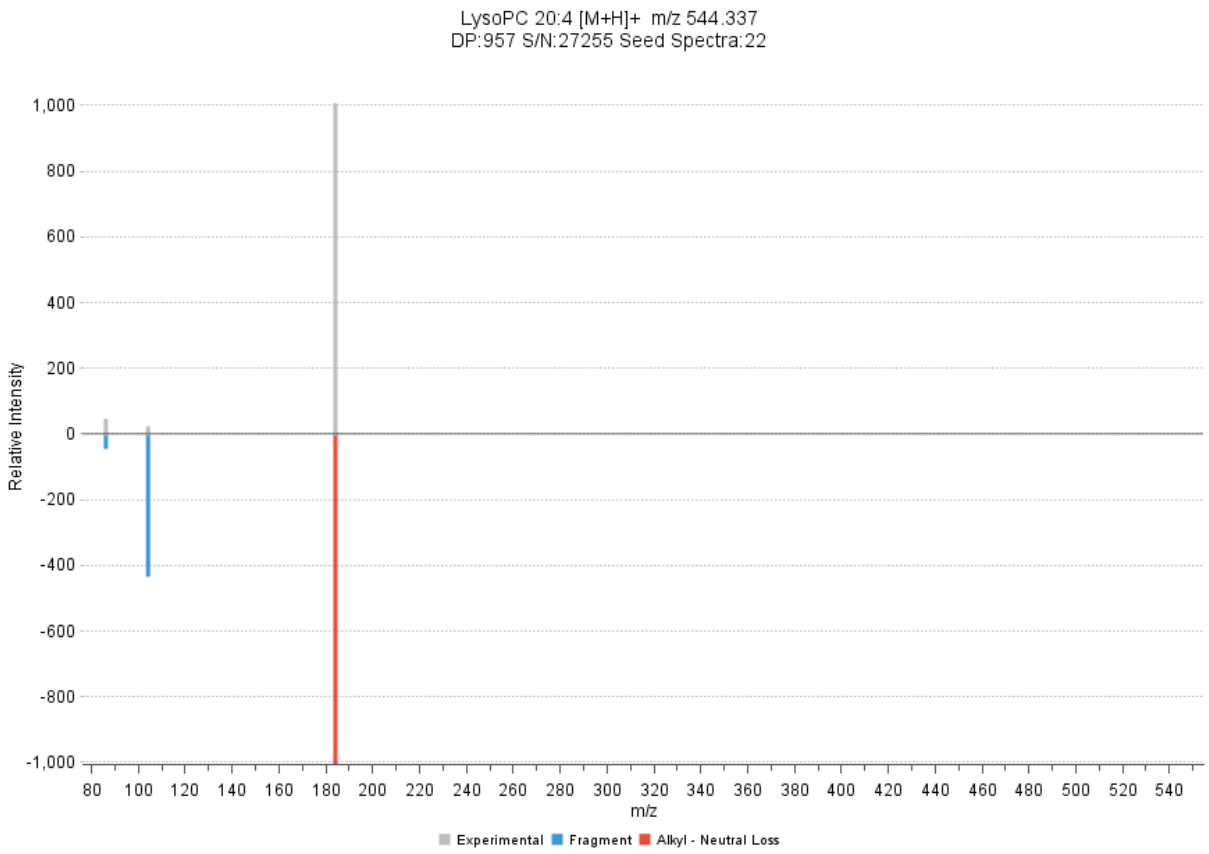
Supplementary Figure 10: Representative experimental (grey) and *in silico* (colored) HexCer[NS] $[M-H]^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



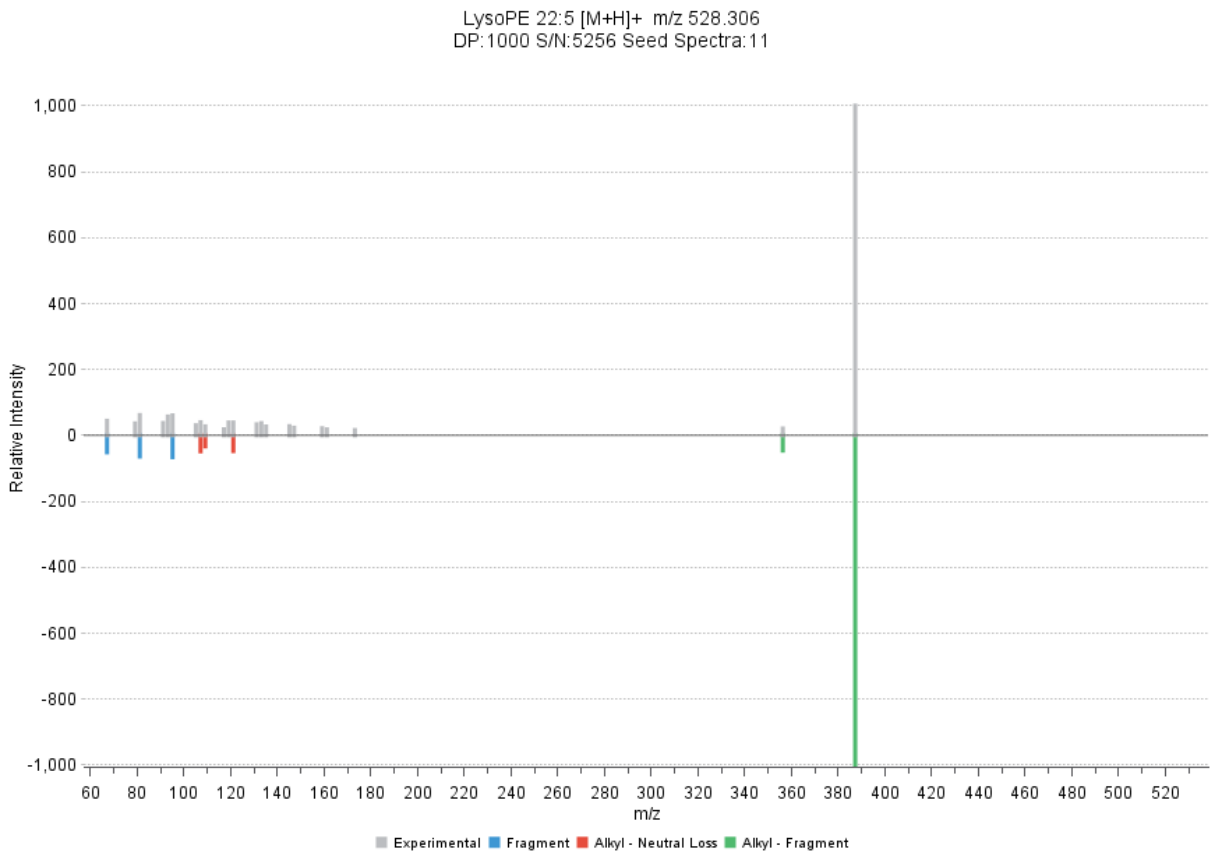
Supplementary Figure 11: Representative experimental (grey) and *in silico* (colored) HexCer[NS] $[M+H]^+$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



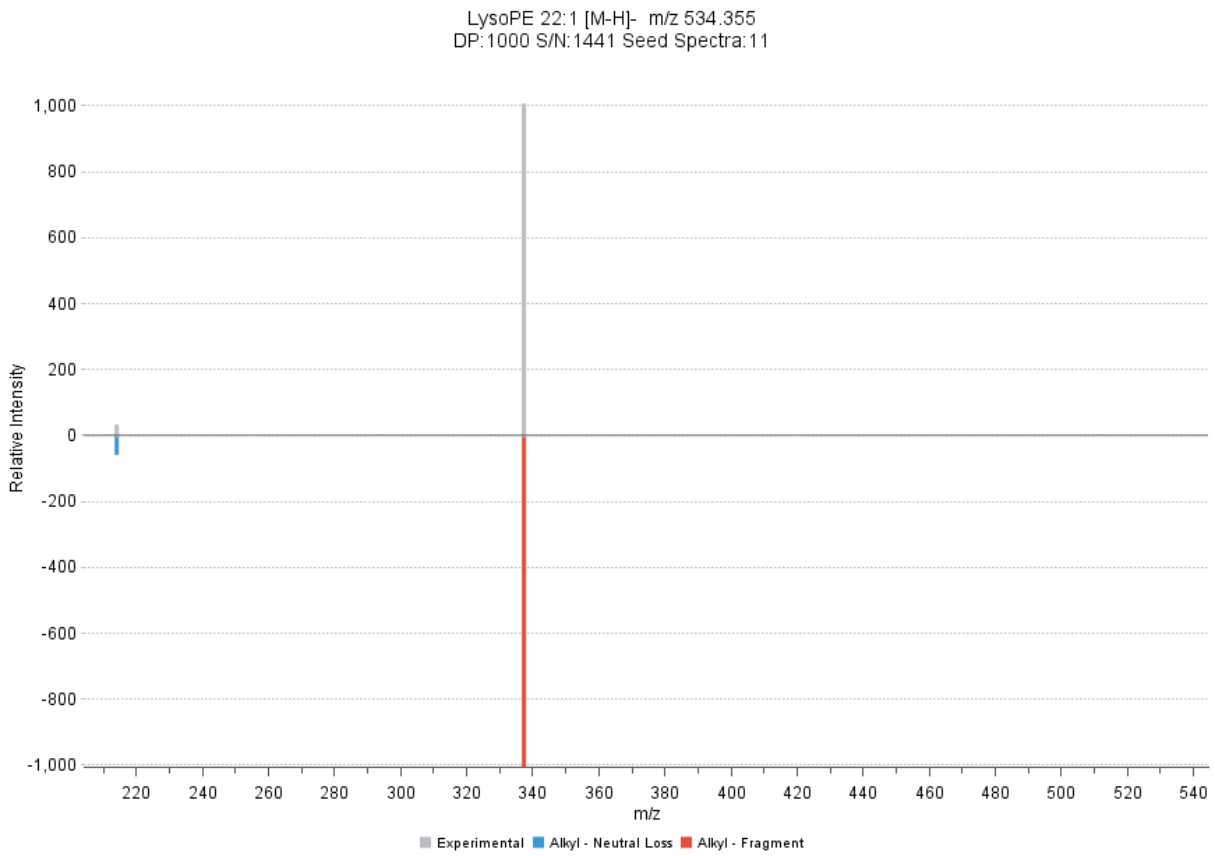
Supplementary Figure 12: Representative experimental (grey) and *in silico* (colored) LysoPC $[M+Ac-H]^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



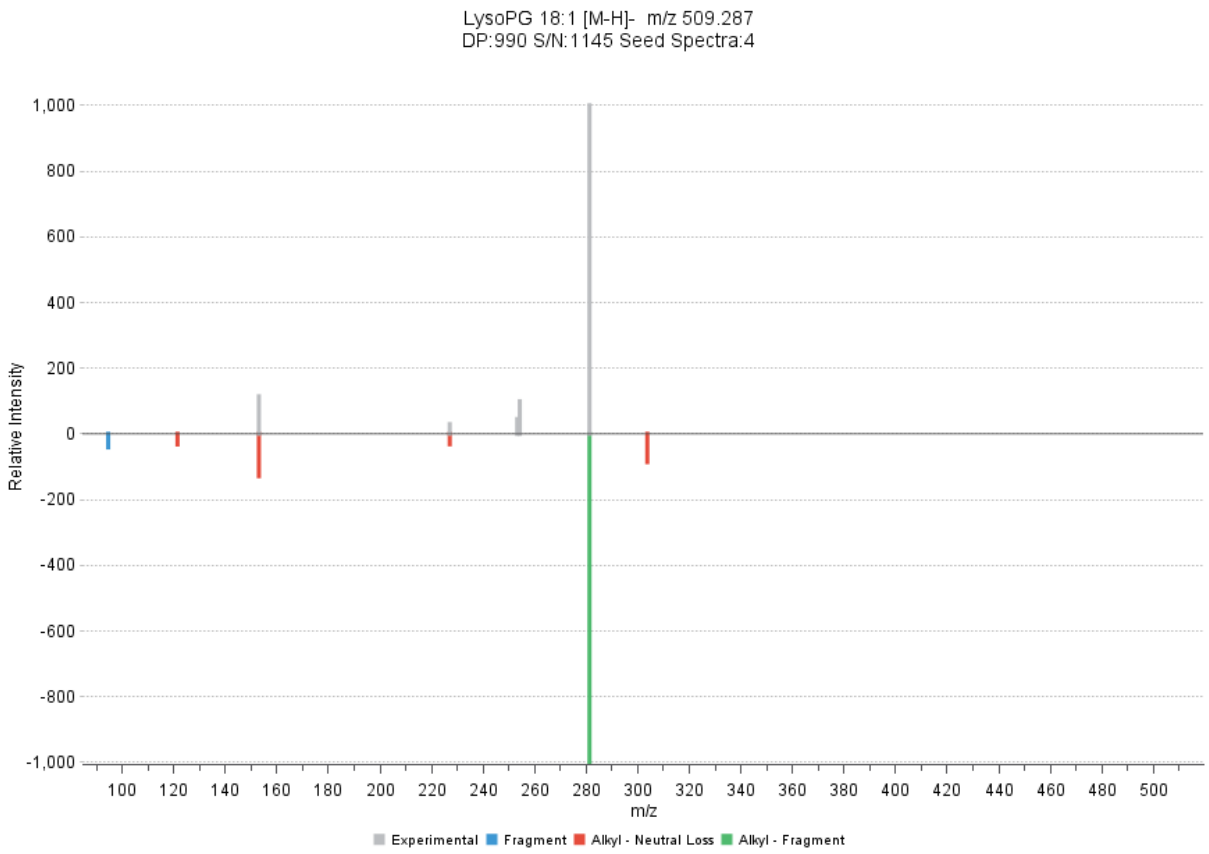
Supplementary Figure 13: Representative experimental (grey) and *in silico* (colored) LysoPC $[M+H]^+$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



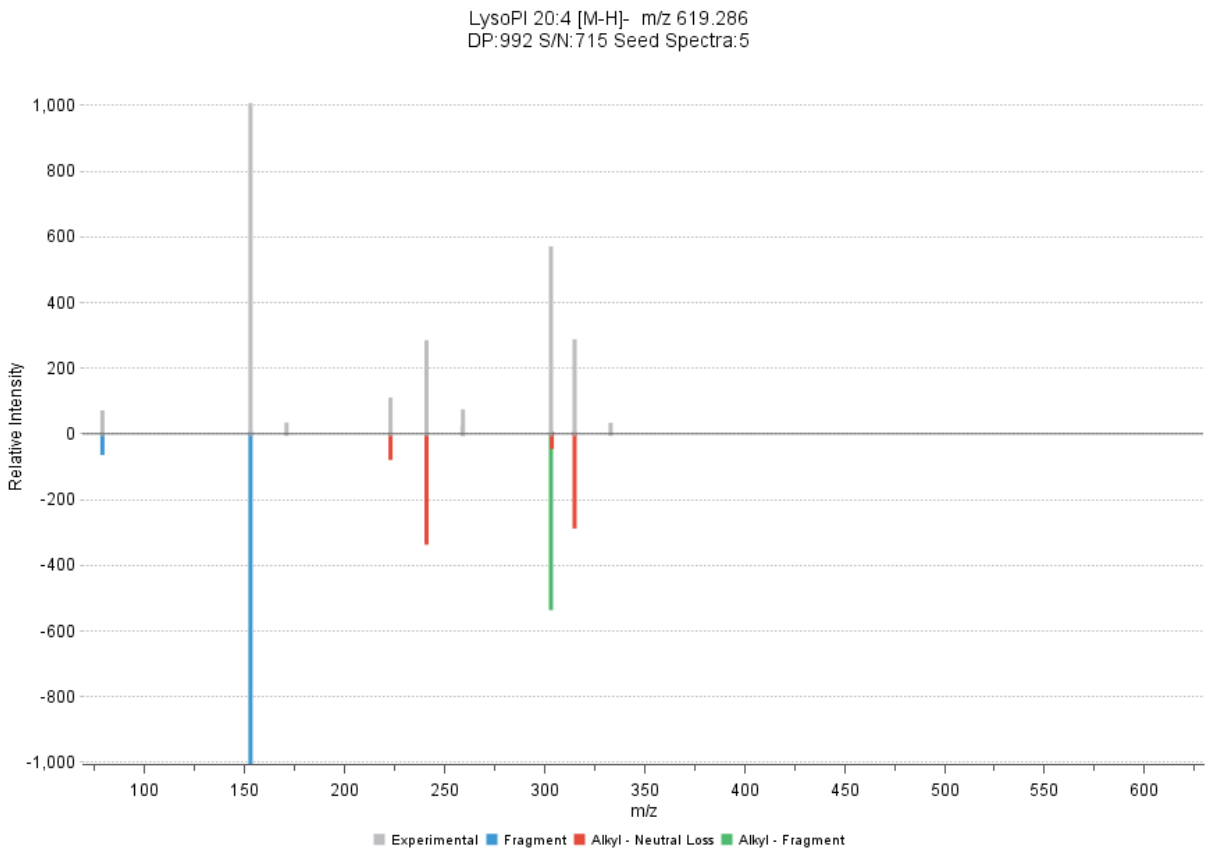
Supplementary Figure 14: Representative experimental (grey) and *in silico* (colored) LysoPE $[M+H]^+$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



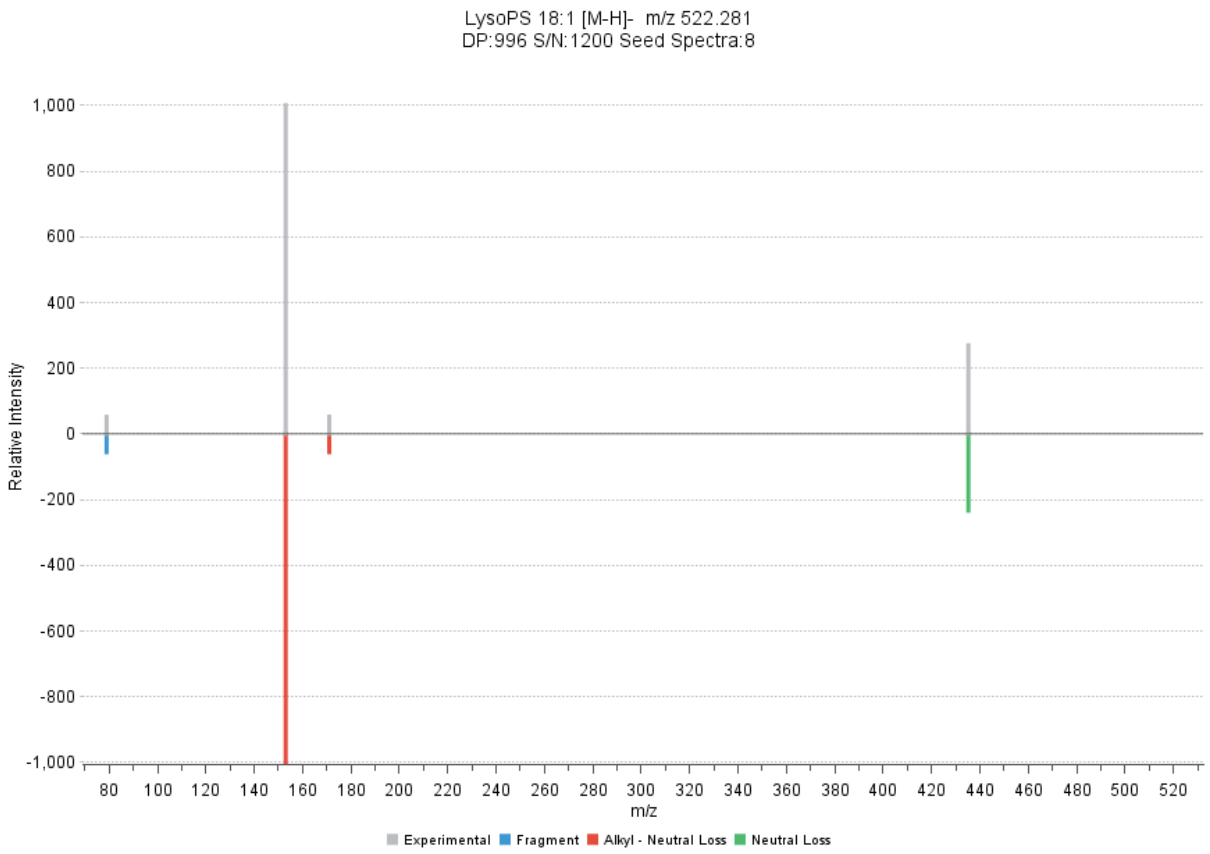
Supplementary Figure 15: Representative experimental (grey) and *in silico* (colored) LysoPE $[M-H]^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



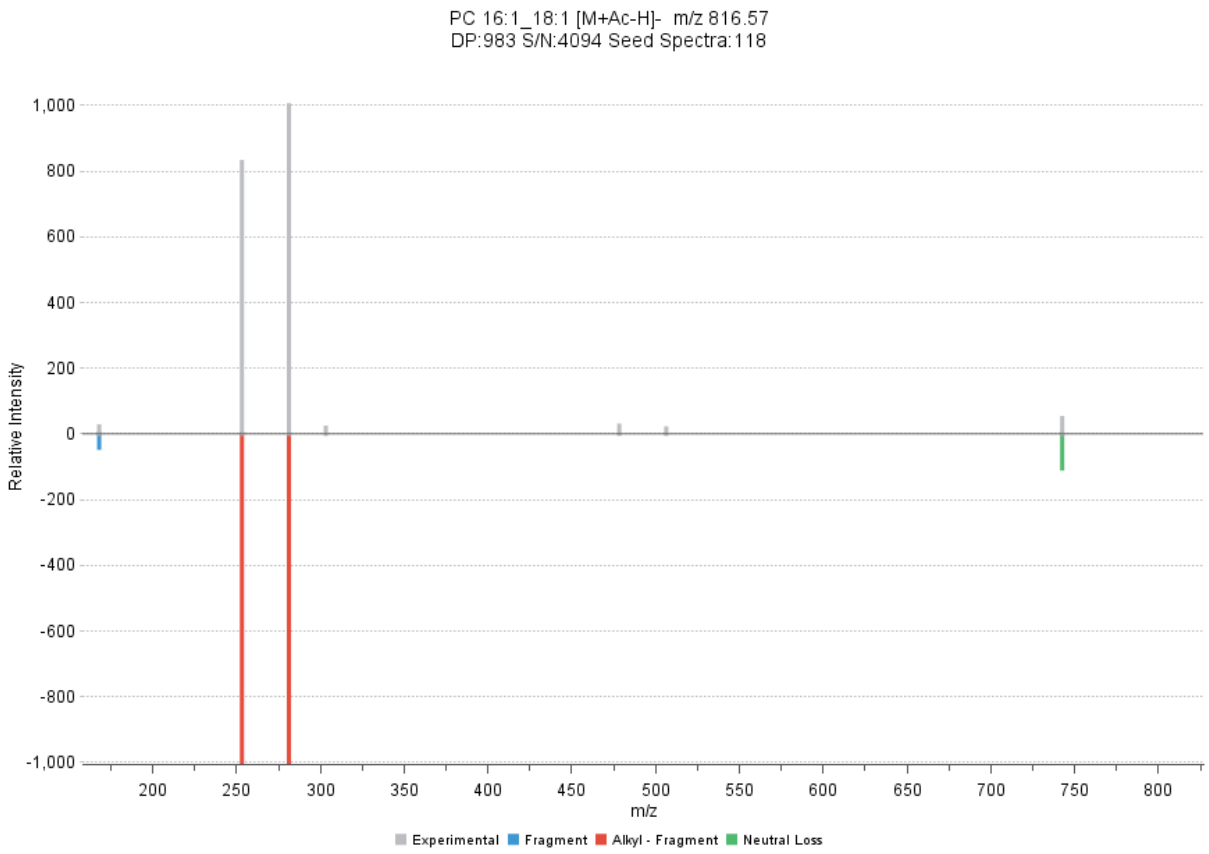
Supplementary Figure 16: Representative experimental (grey) and *in silico* (colored) LysoPG $[M-H]^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



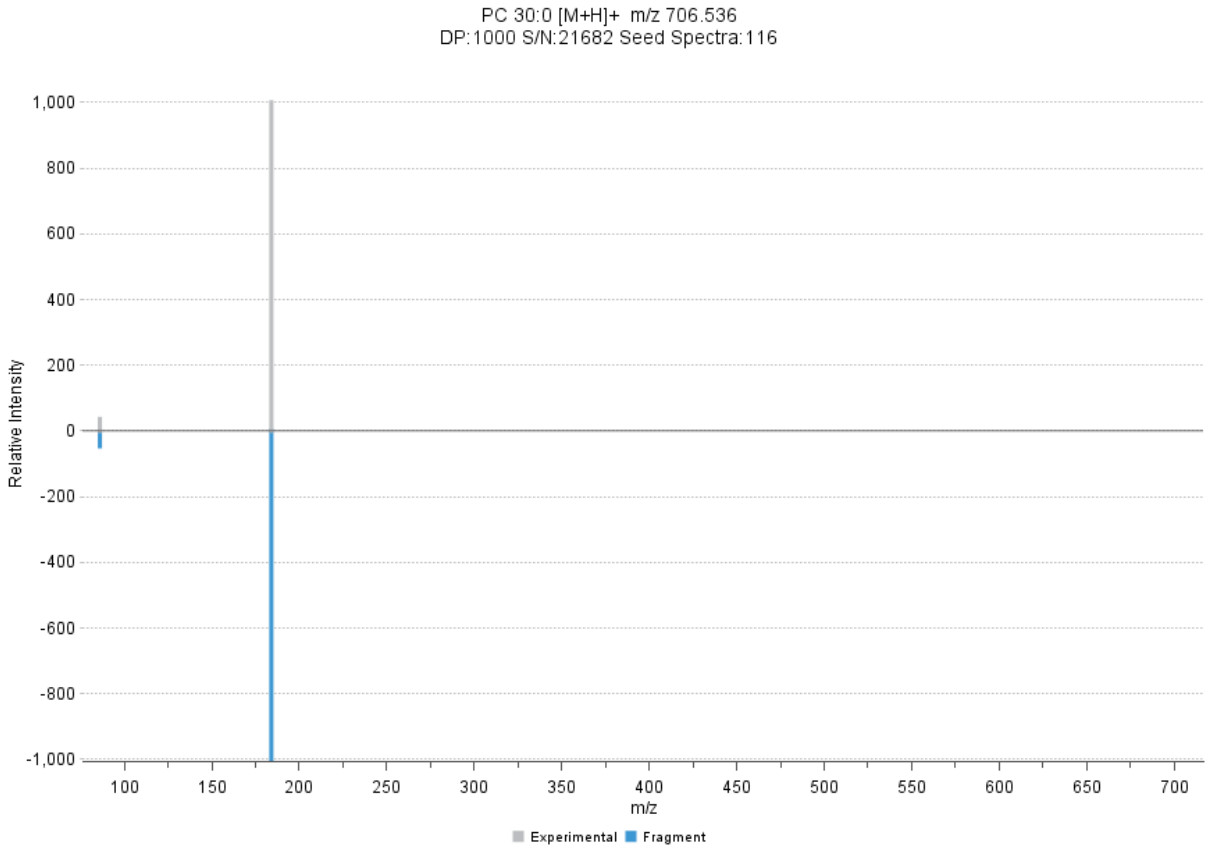
Supplementary Figure 17: Representative experimental (grey) and *in silico* (colored) LysoPI [M-H]$^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



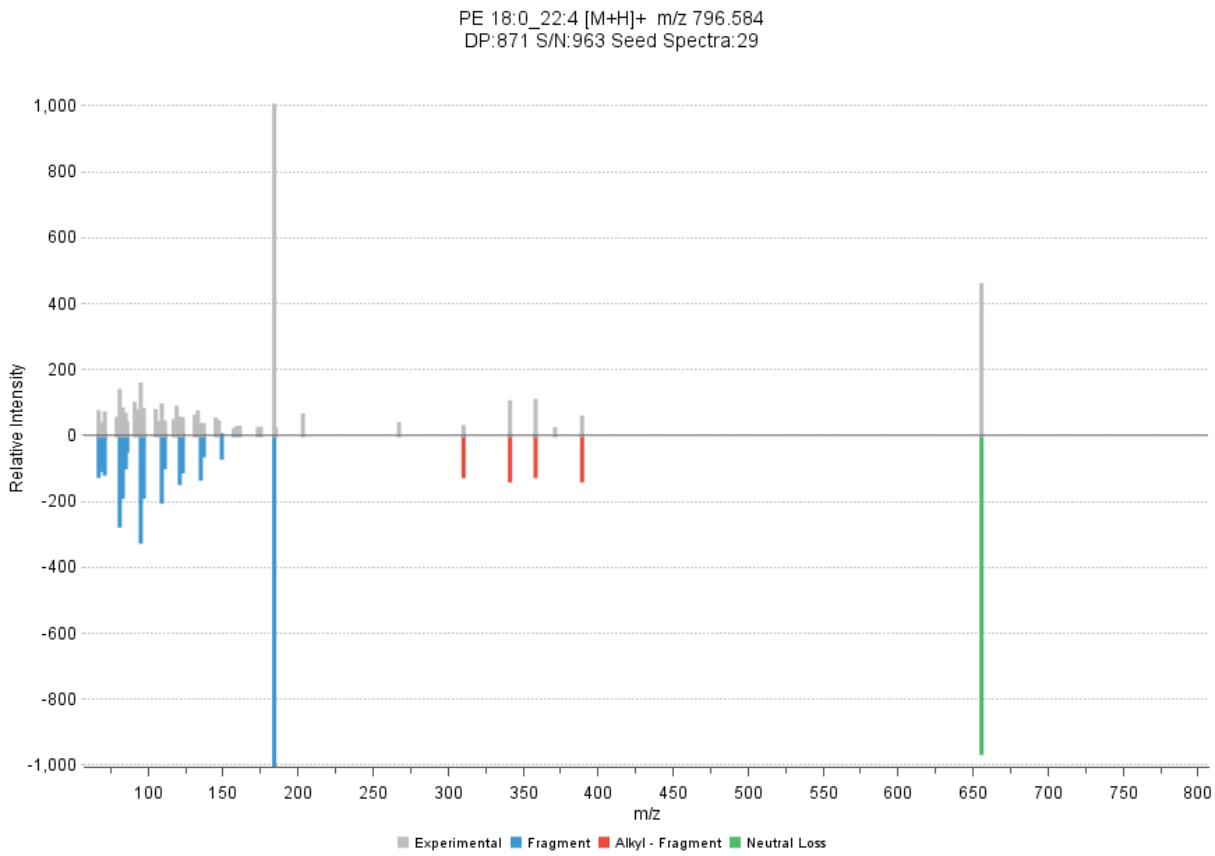
Supplementary Figure 18: Representative experimental (grey) and *in silico* (colored) LysoPS $[M-H]^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



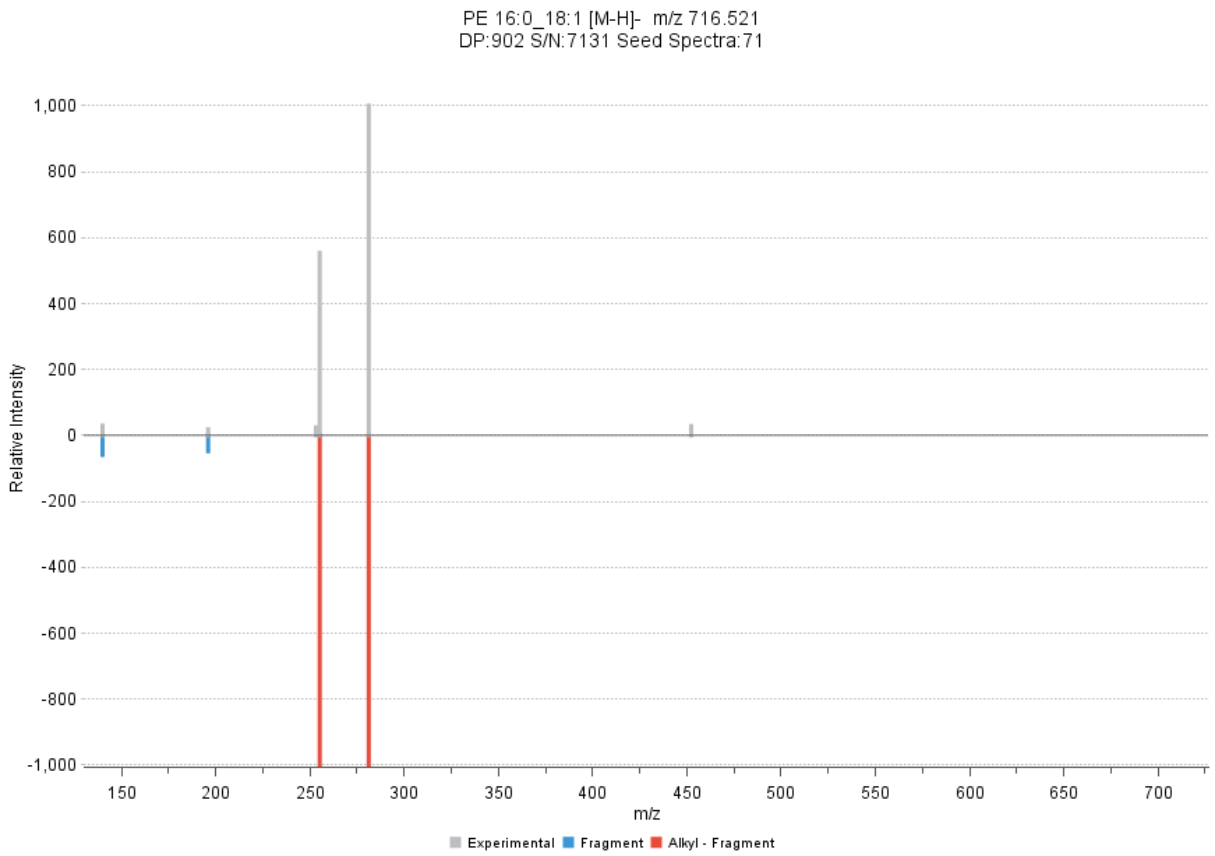
Supplementary Figure 19: Representative experimental (grey) and *in silico* (colored) PC [M+Ac-H]$^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



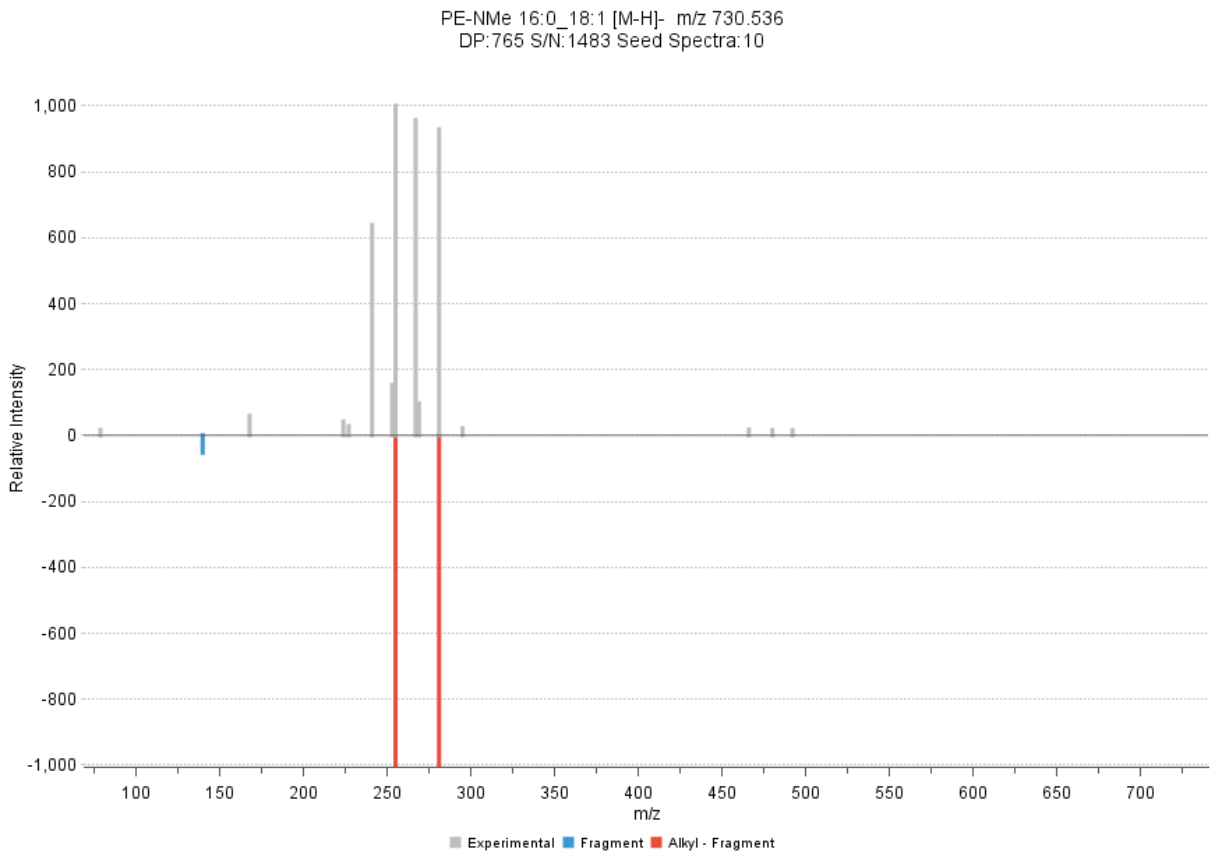
Supplementary Figure 20: Representative experimental (grey) and *in silico* (colored) PC $[M+H]^+$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



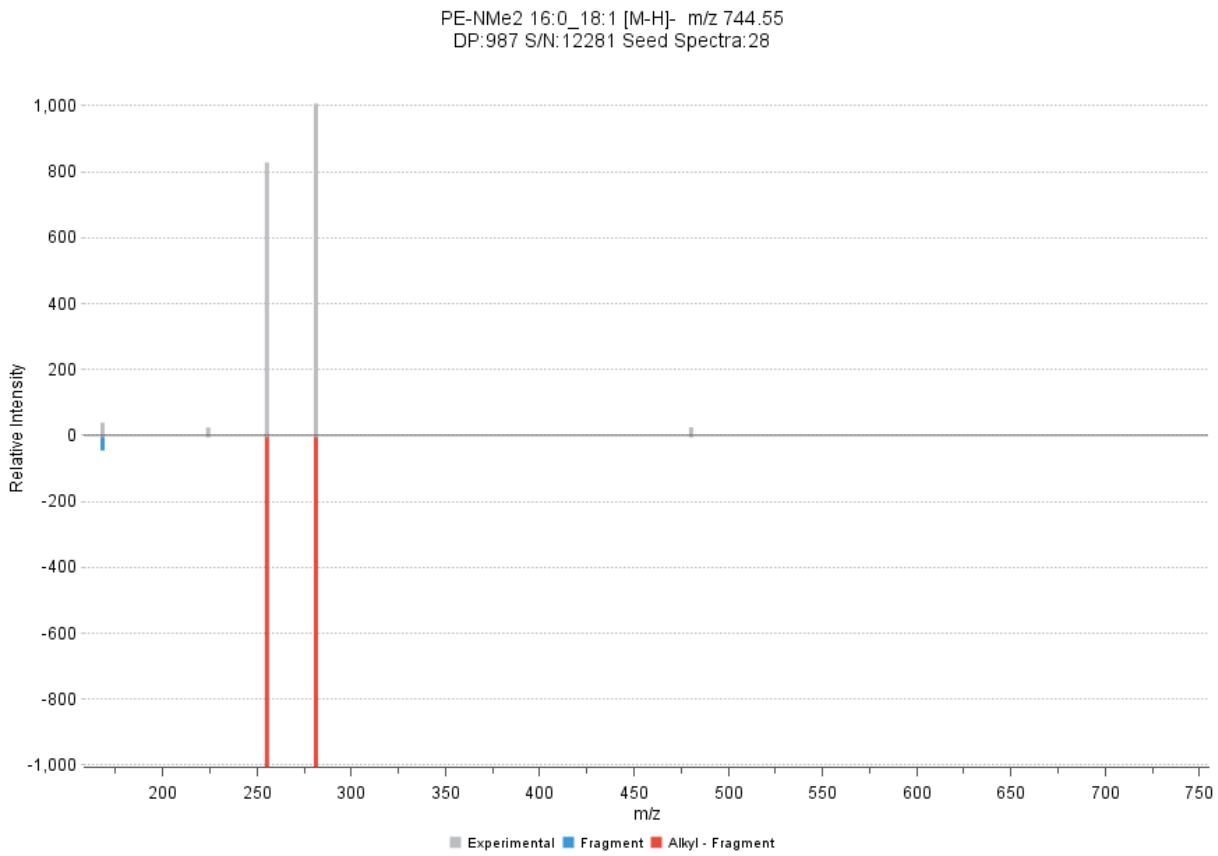
Supplementary Figure 21: Representative experimental (grey) and *in silico* (colored) PE $[M+H]^+$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library. Note: the presence of an annotated peak at *m/z* 184 is due to the prevalent co-fragmentation of PE and PC species.



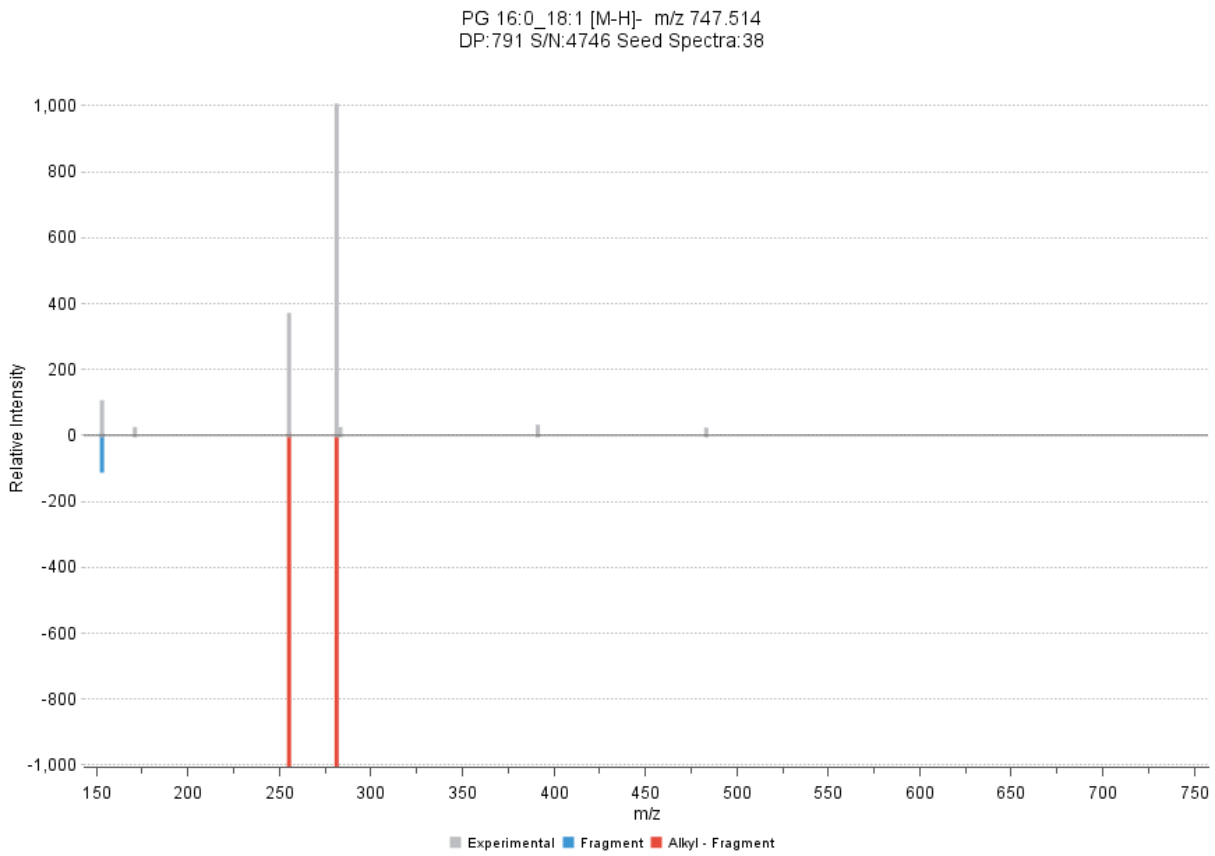
Supplementary Figure 22: Representative experimental (grey) and *in silico* (colored) PE $[M-H]^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



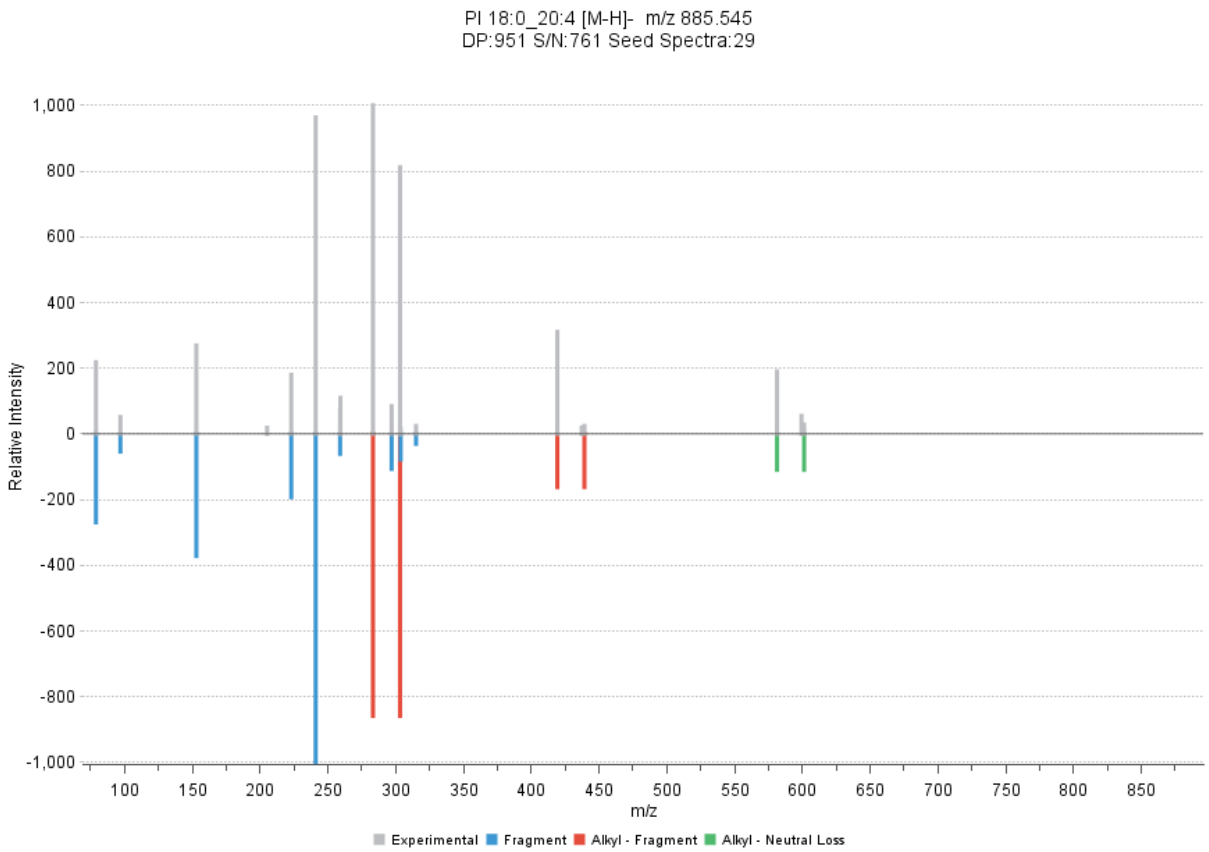
Supplementary Figure 23: Representative experimental (grey) and *in silico* (colored) PE-NMe $[M-H]^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



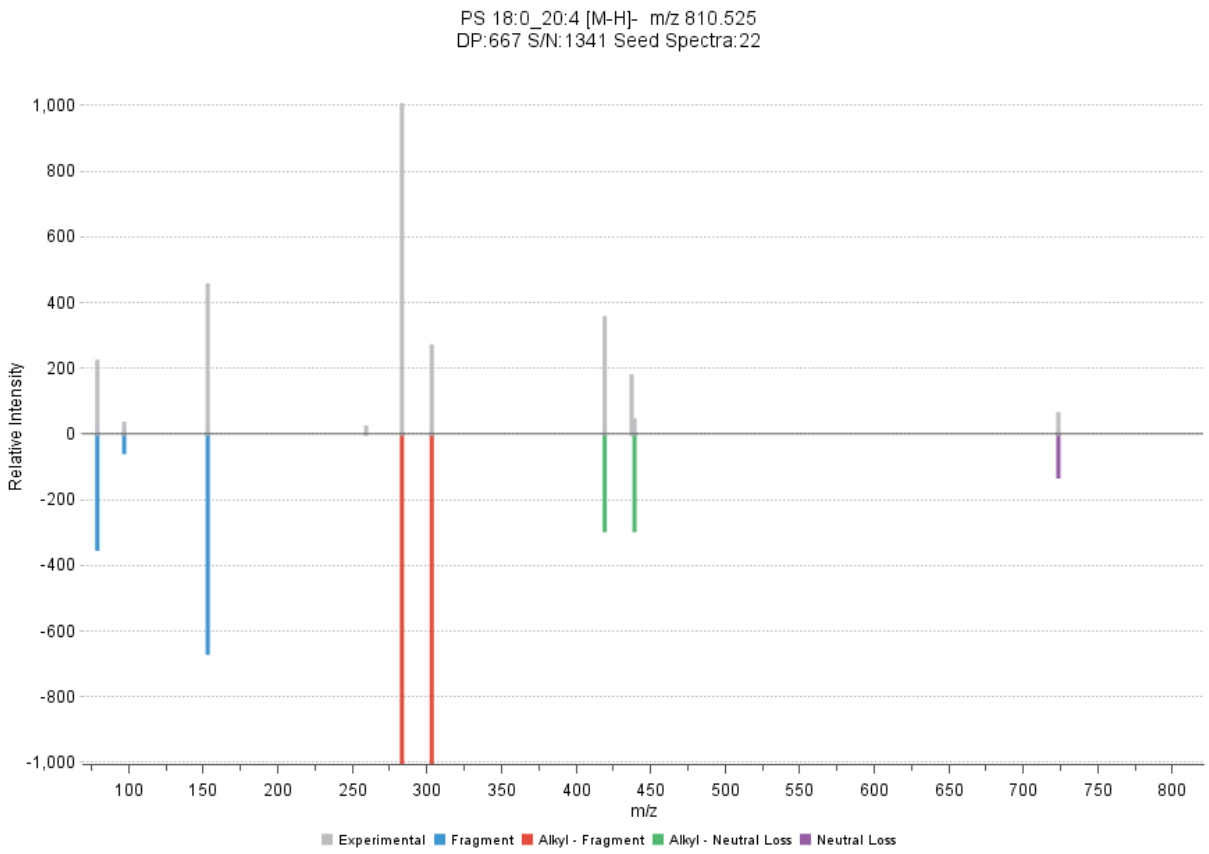
Supplementary Figure 24: Representative experimental (grey) and *in silico* (colored) PE-NMe2 $[M-H]^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



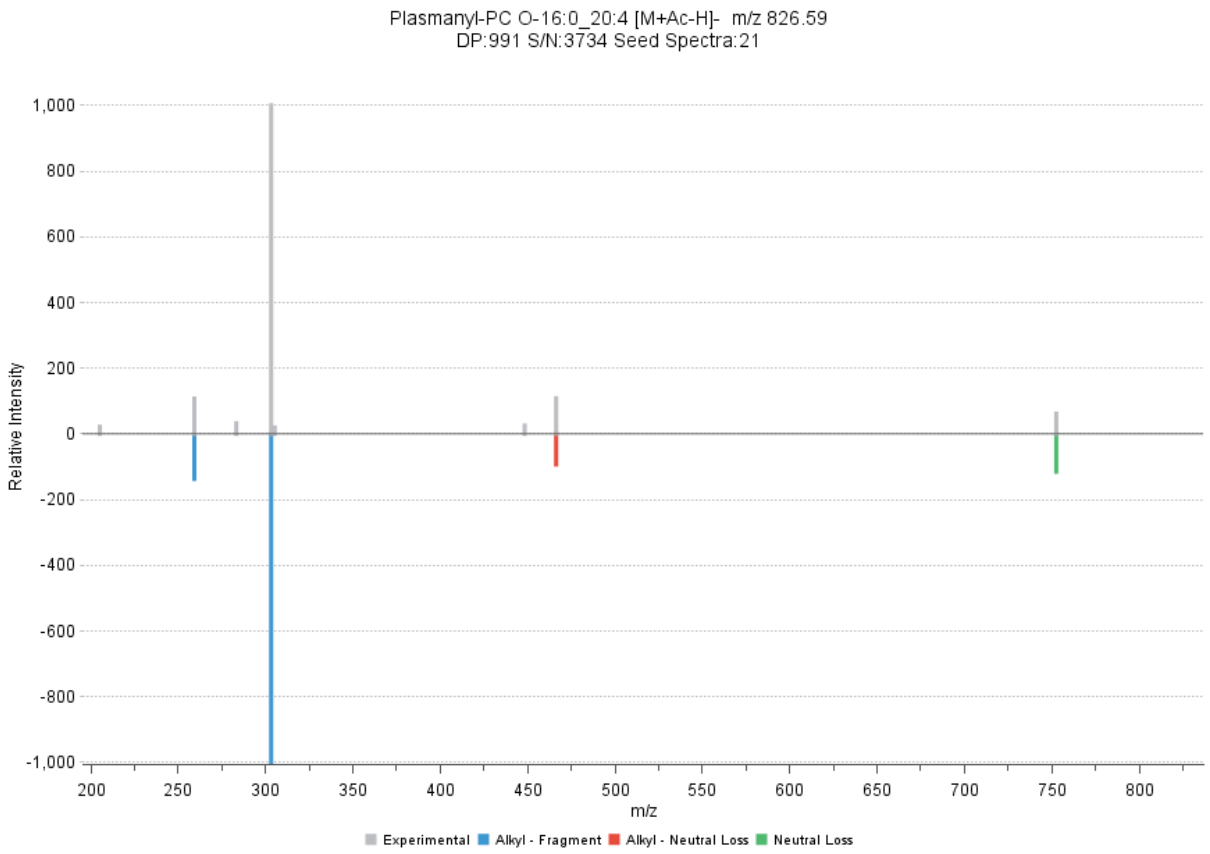
Supplementary Figure 25: Representative experimental (grey) and *in silico* (colored) PG $[M-H]^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



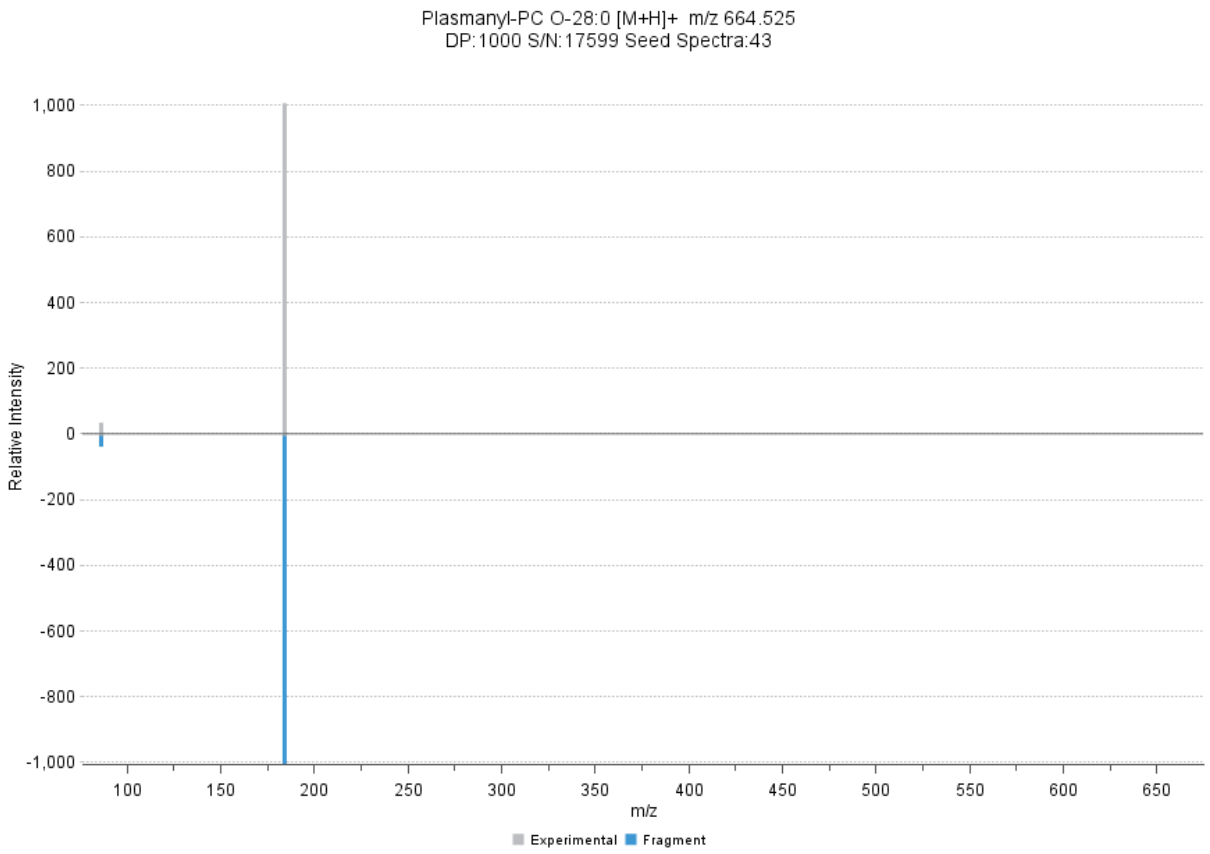
Supplementary Figure 26: Representative experimental (grey) and *in silico* (colored) PI $[M-H]^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



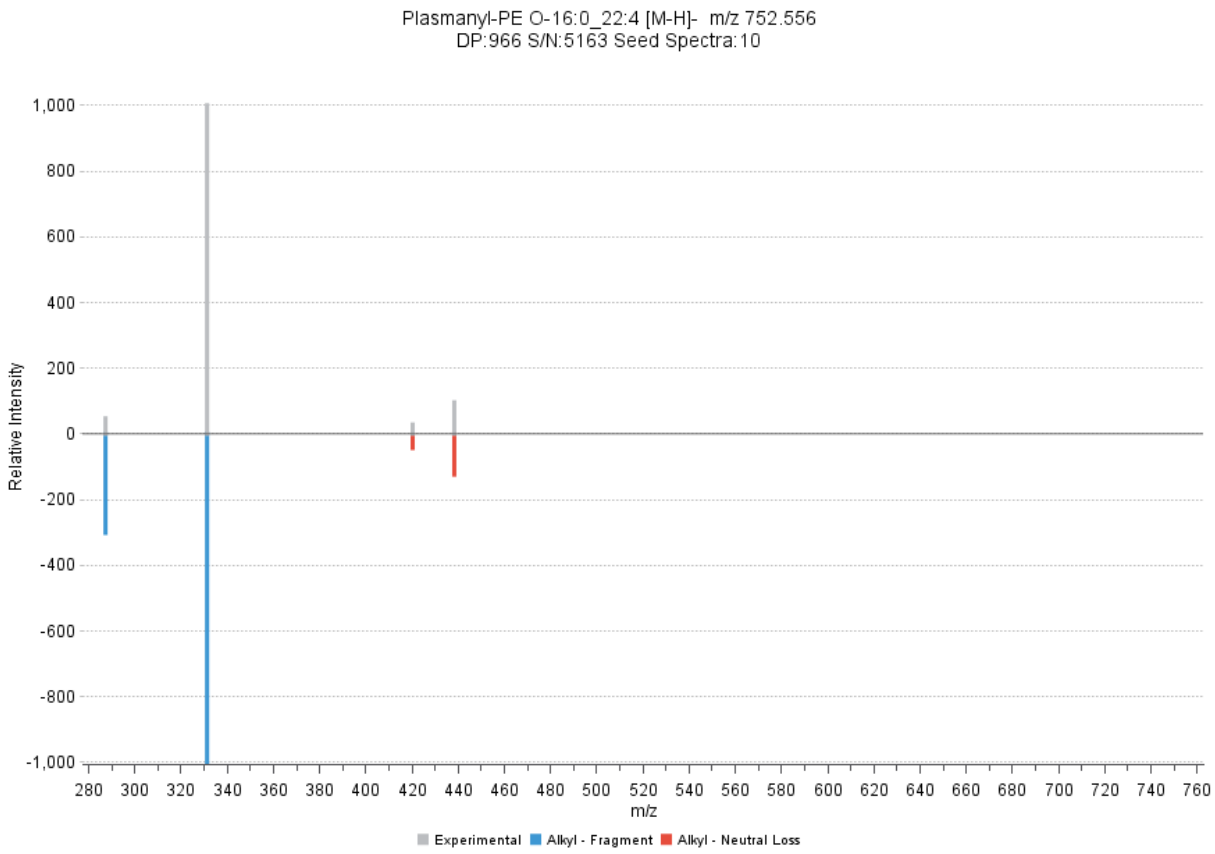
Supplementary Figure 27: Representative experimental (grey) and *in silico* (colored) PS $[M-H]^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



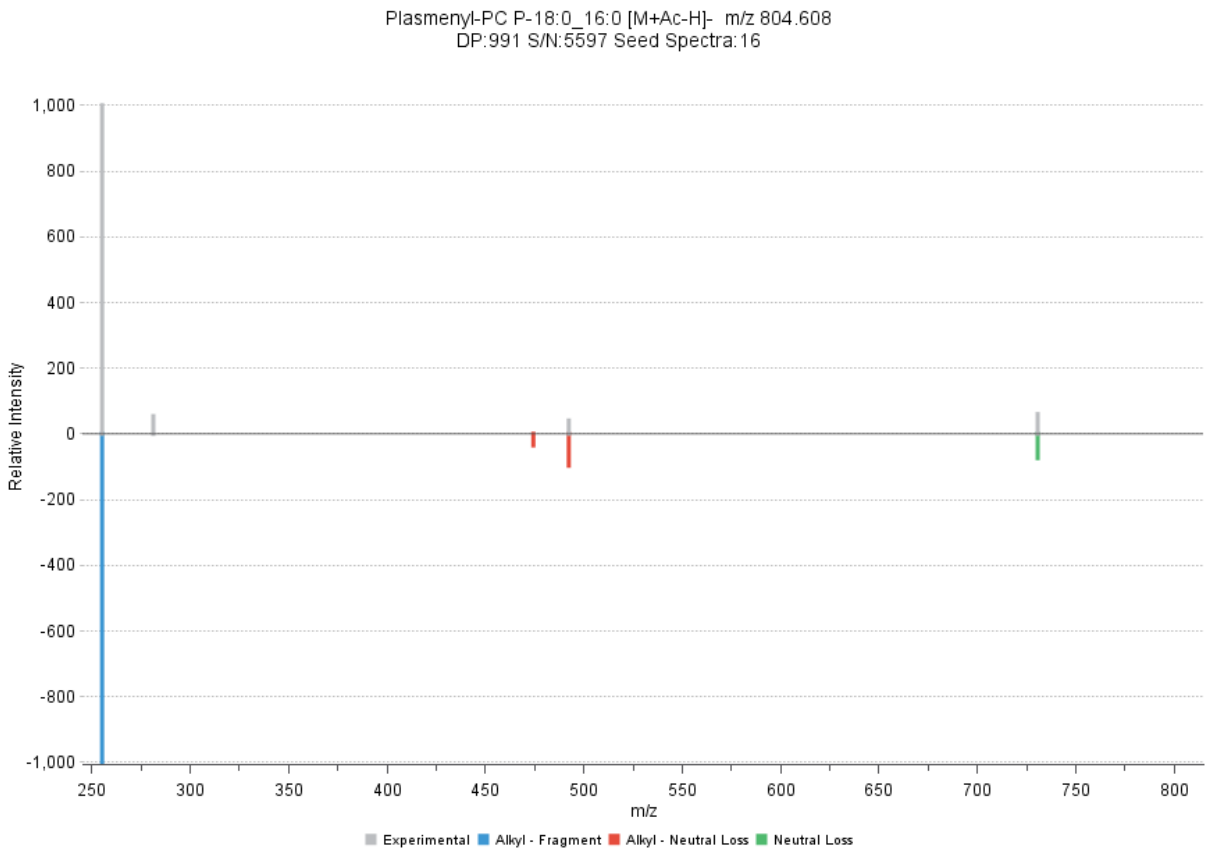
Supplementary Figure 28: Representative experimental (grey) and *in silico* (colored) Plasmanyl-PC $[M+Ac-H]^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



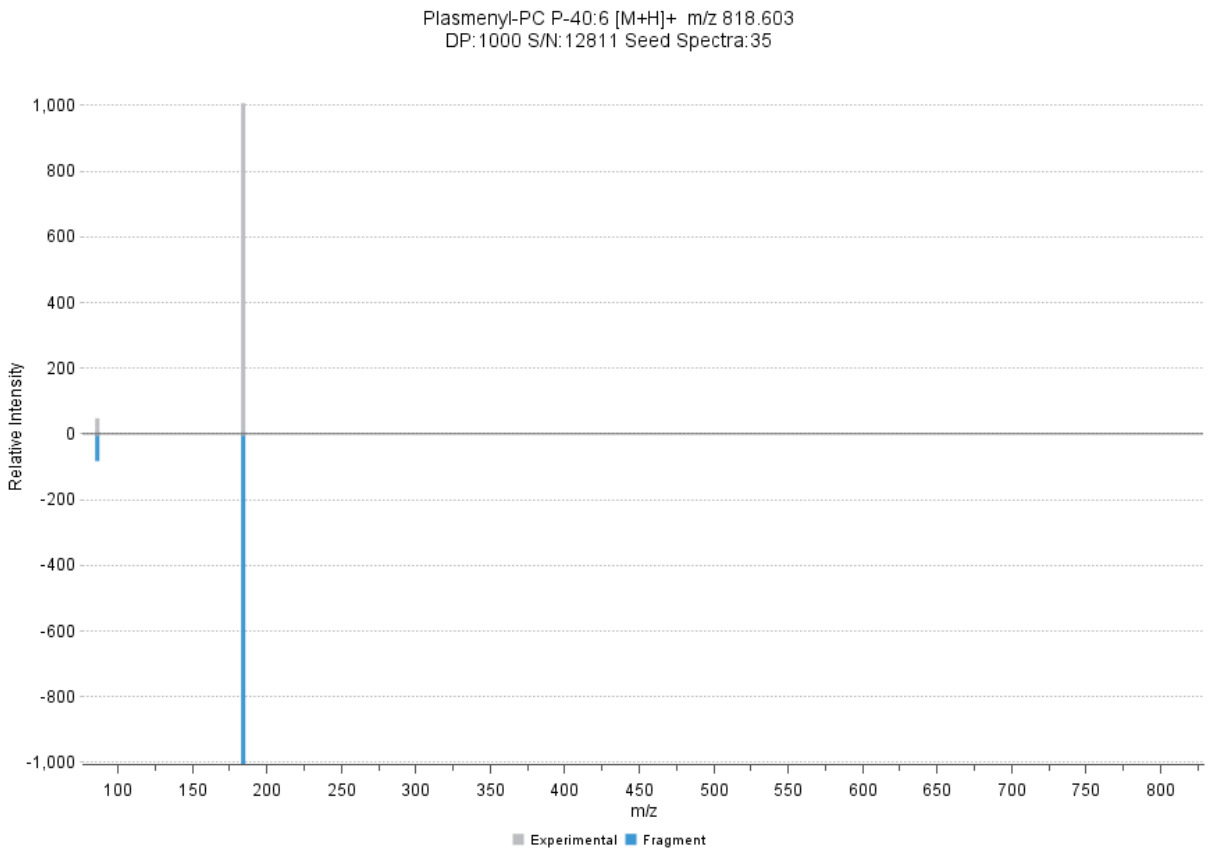
Supplementary Figure 29: Representative experimental (grey) and *in silico* (colored) Plasmanyl-PC $[M+H]^+$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



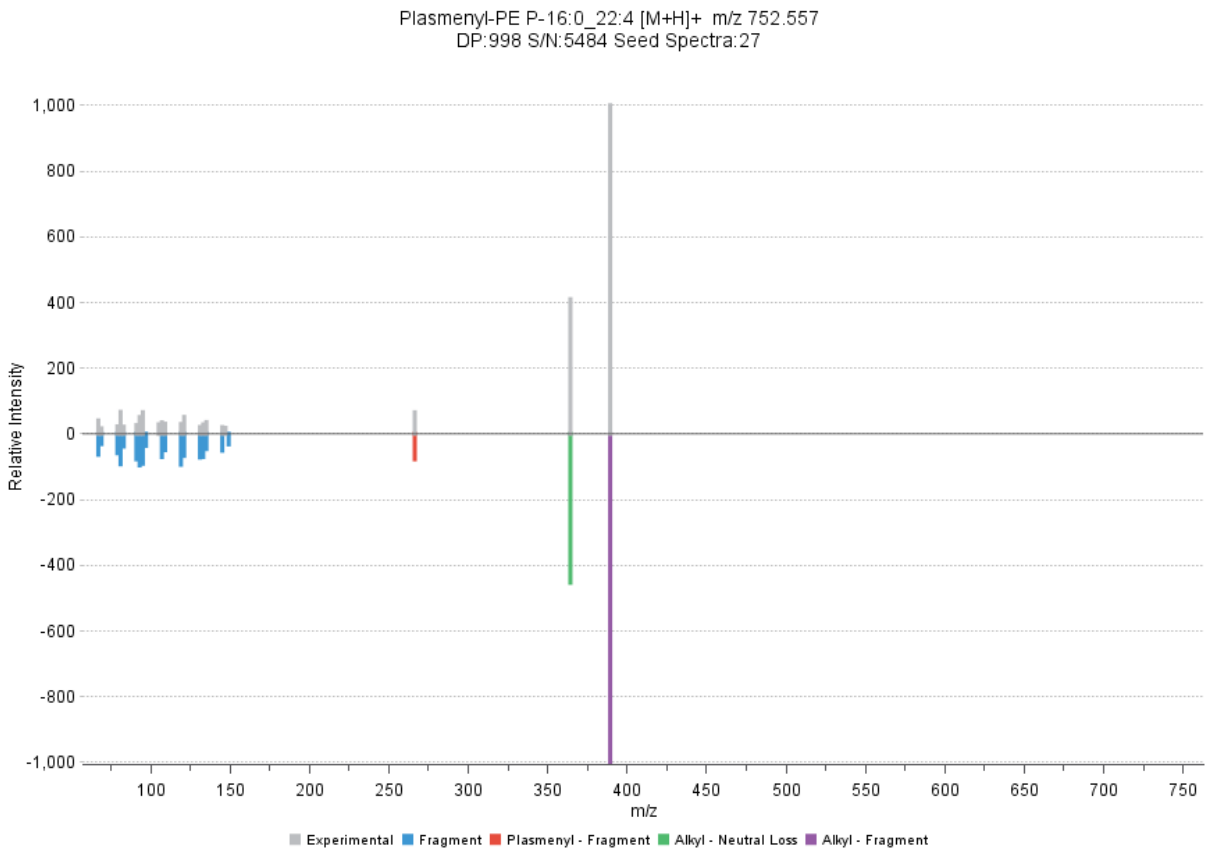
Supplementary Figure 30: Representative experimental (grey) and *in silico* (colored) Plasmanyl-PE $[M-H]^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



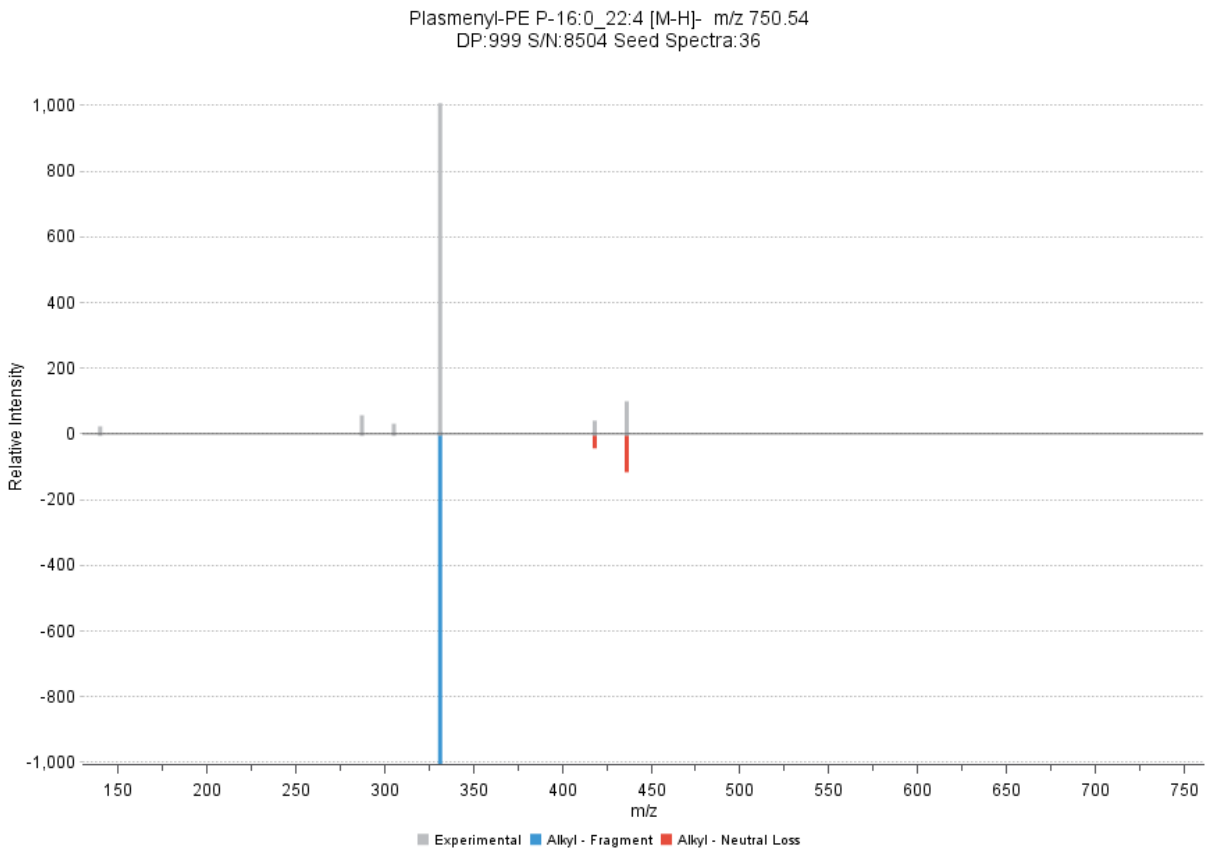
Supplementary Figure 31: Representative experimental (grey) and *in silico* (colored) Plasmenyl-PC $[M+Ac-H]^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



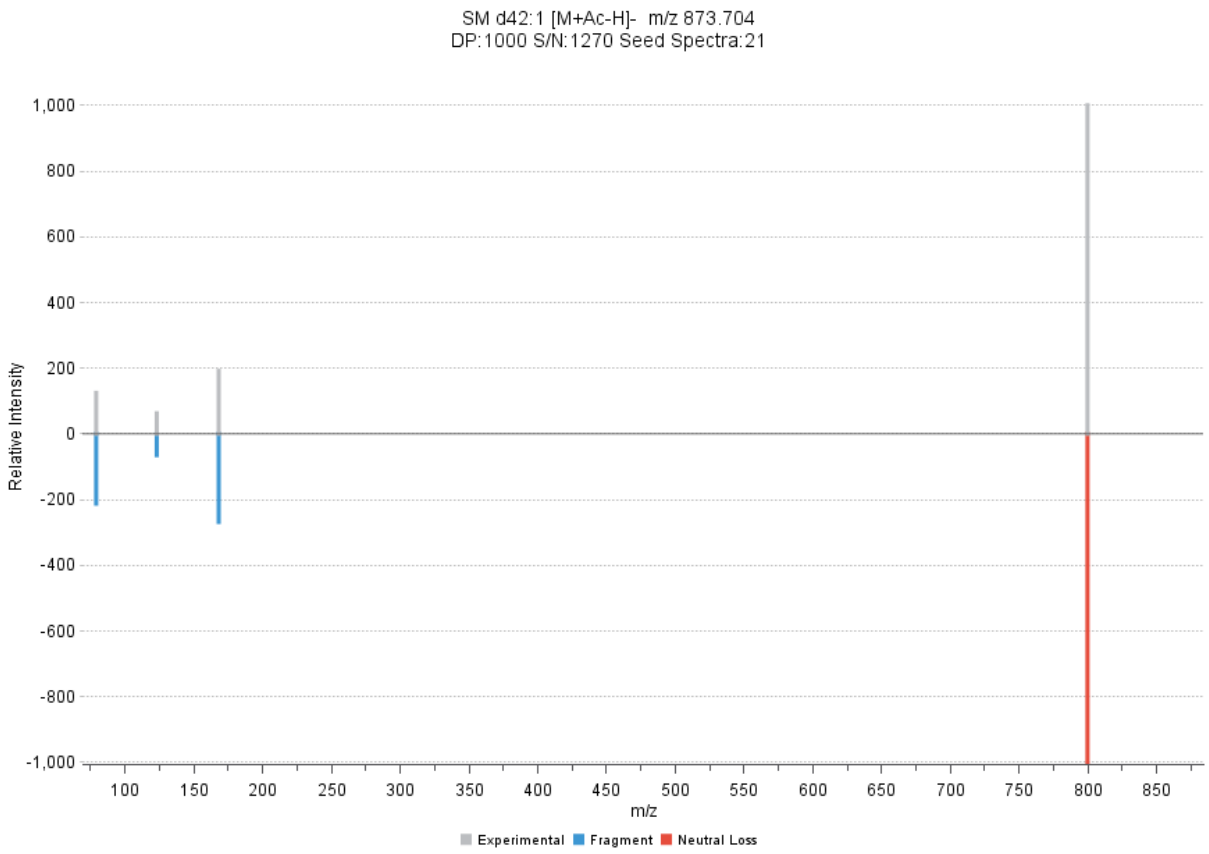
Supplementary Figure 32: Representative experimental (grey) and *in silico* (colored) Plasmenyl-PC $[M+H]^+$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



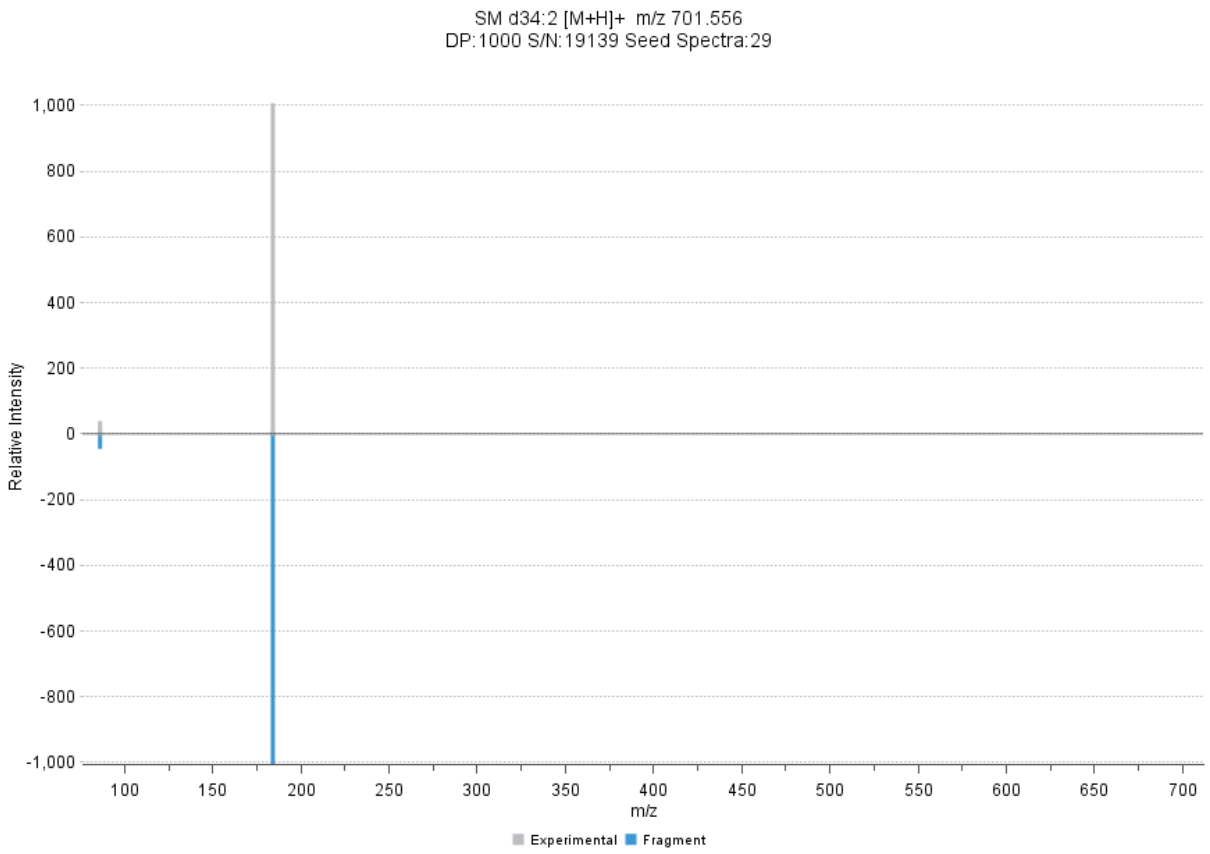
Supplementary Figure 33: Representative experimental (grey) and *in silico* (colored) Plasmenyl-PE $[M+H]^+$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



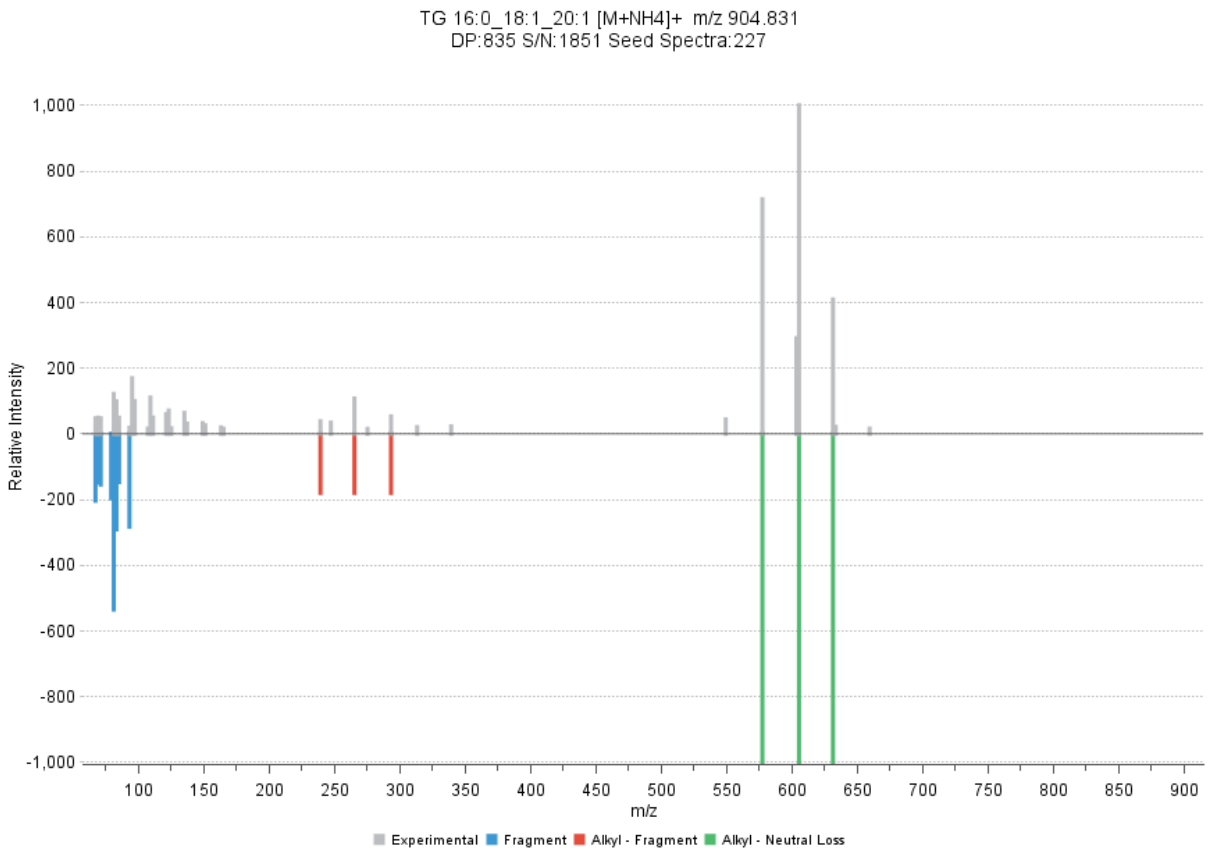
Supplementary Figure 34: Representative experimental (grey) and *in silico* (colored) Plasmenyl-PE $[M-H]^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



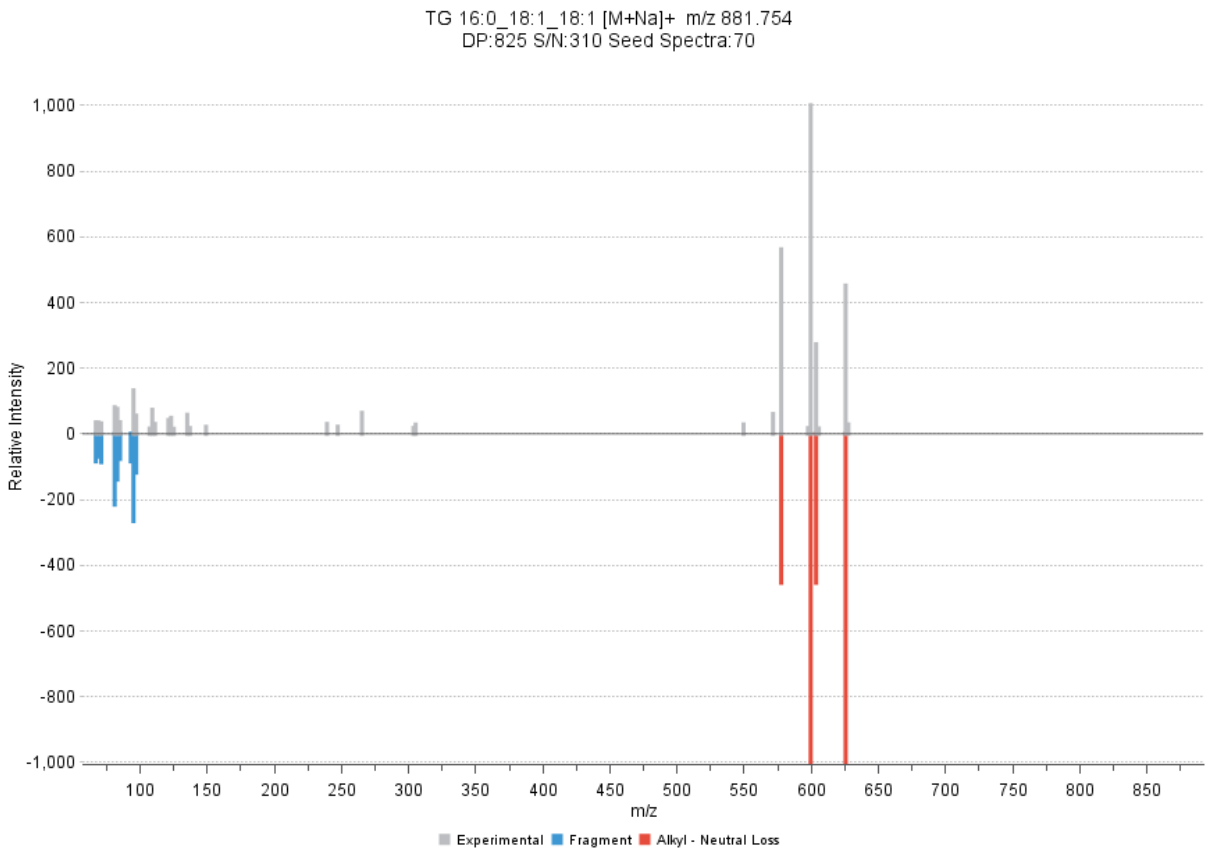
Supplementary Figure 35: Representative experimental (grey) and *in silico* (colored) SM [M+Ac-H]$^-$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



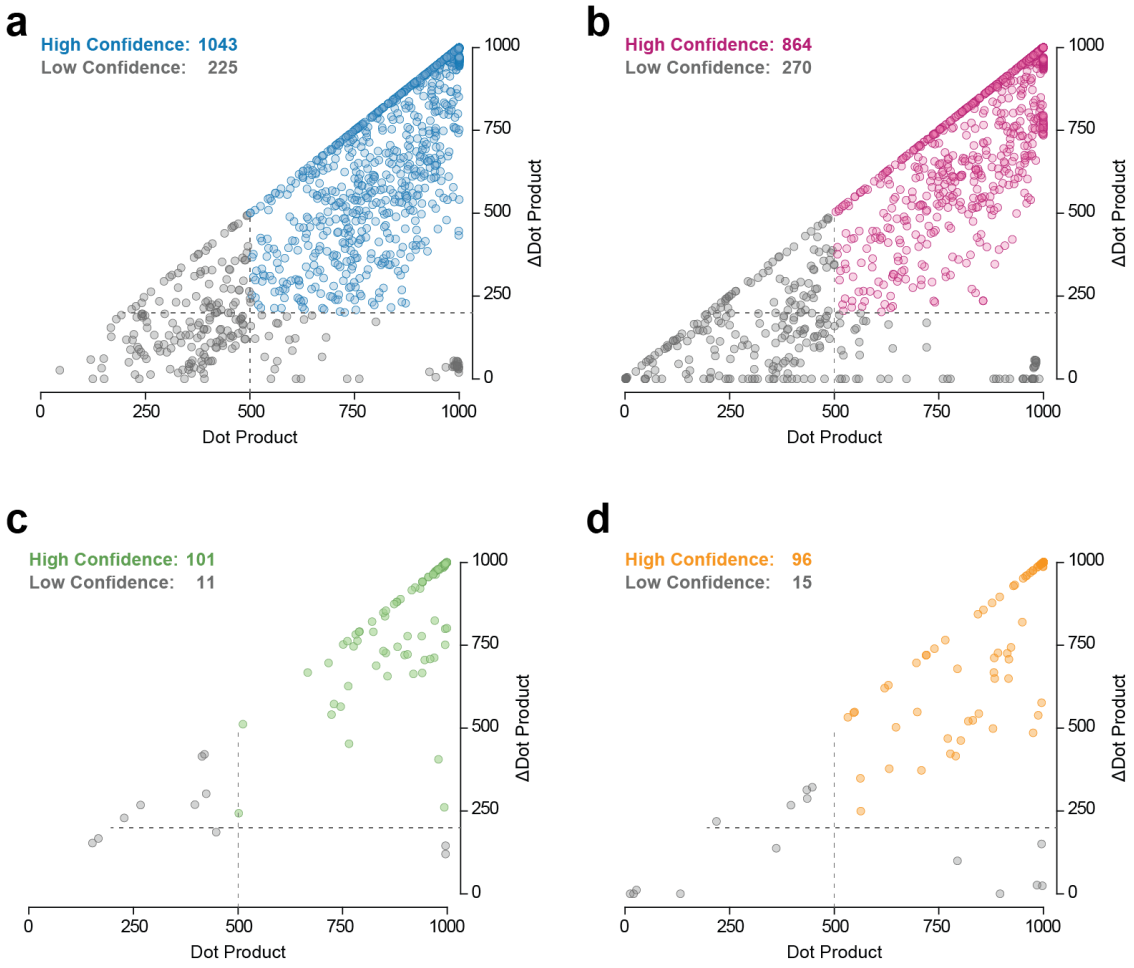
Supplementary Figure 36: Representative experimental (grey) and *in silico* (colored) SM $[M+H]^+$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



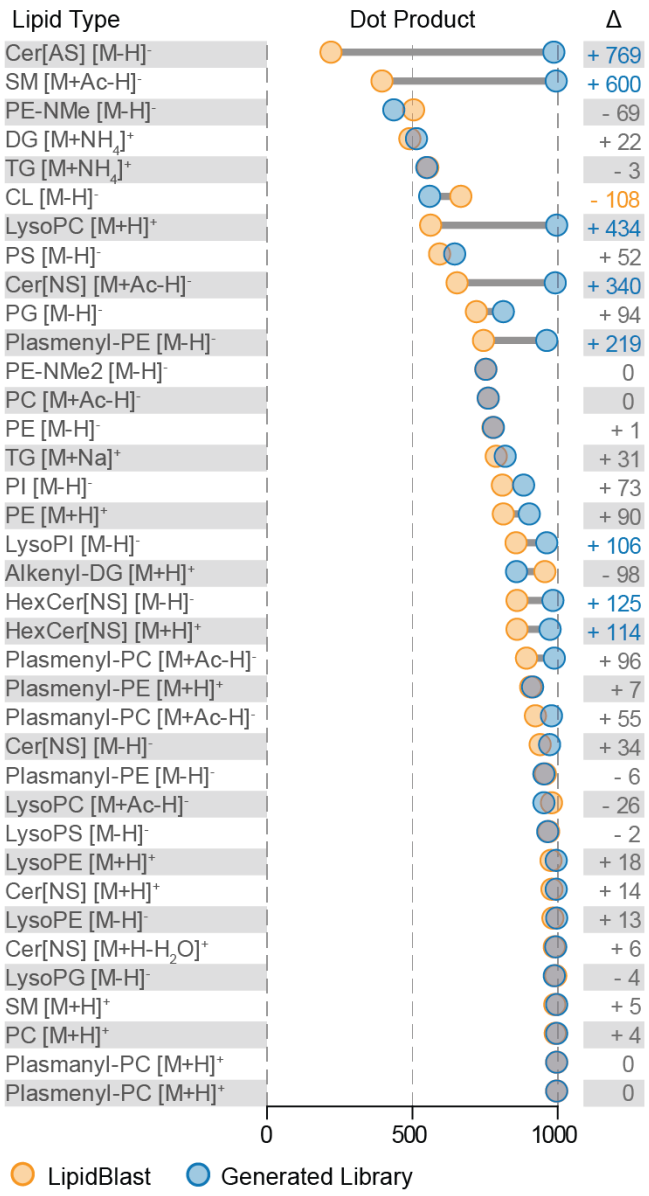
Supplementary Figure 37: Representative experimental (grey) and *in silico* (colored) TG $[M+NH_4]^+$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



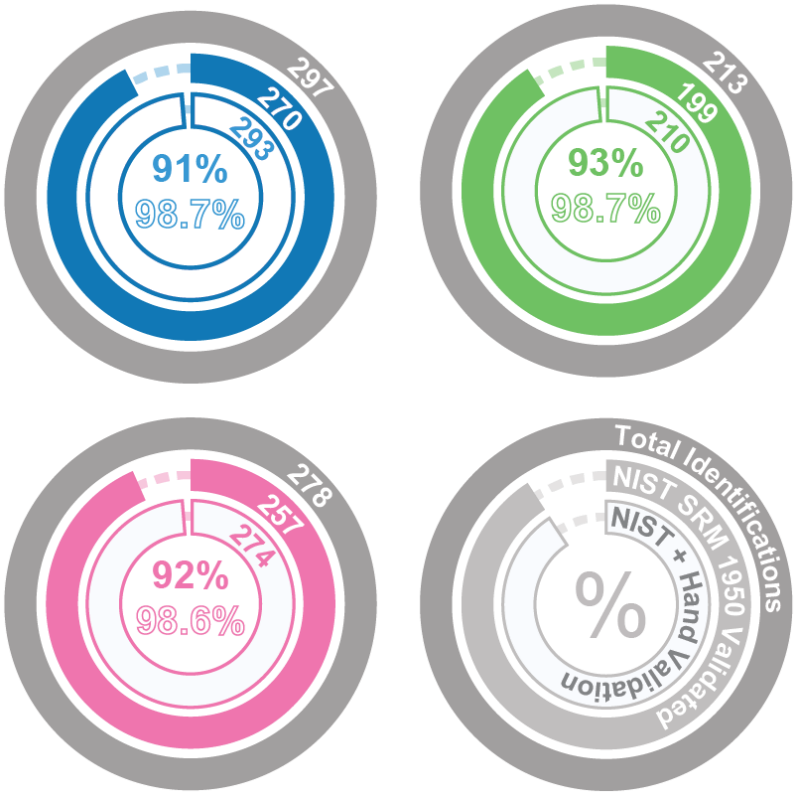
Supplementary Figure 38: Representative experimental (grey) and *in silico* (colored) TG $[M+NA]^+$ spectra generated from the complex HAP1 extract using the LipiDex HCD Acetate library as the seed library.



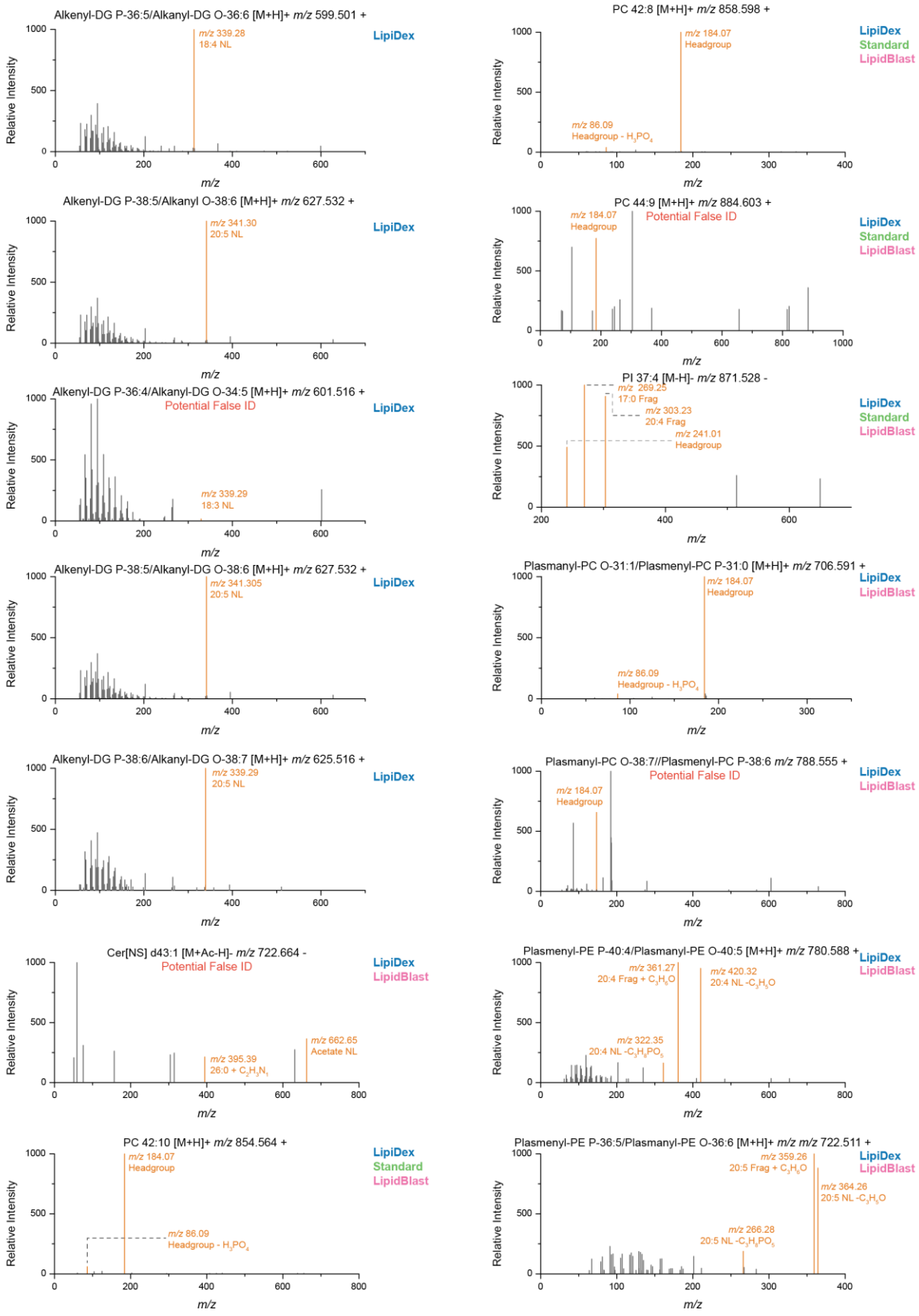
Supplementary Figure 39: Identification Confidence using Generated Libraries. Dot product returned and dot product difference between the correct match and the top scoring incorrect match for spectra generated from (A) complex Hap1 cell extract using LipiDex spectral library, (B) complex Hap1 cell extract using LipidBlast spectral library, (C) reference standard mix using HCD, and (D) reference standard mix using CAD. Note: for complex spectra which contained co-fragmented lipid species, these identifications were not counted as incorrect. The score cutoffs used to determine high confidence are meant to give a semi-quantitative value to the number of likely identifications and not a rigorous statistical determination of FDR.

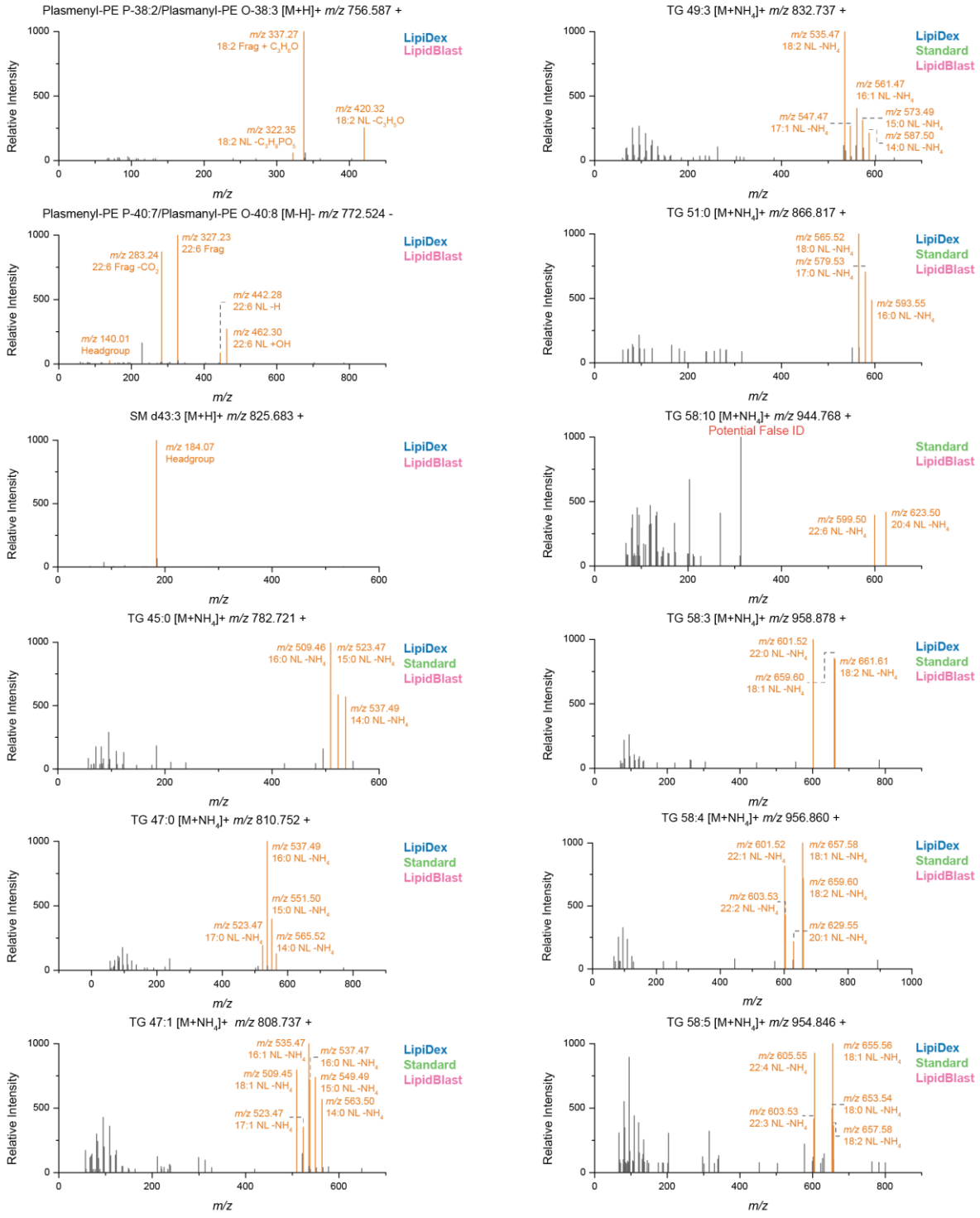


Supplementary Figure 40: Dot Product Enhancement after Fragmentation Mapping. Median dot product returned for each lipid class/adduct combination mapped when searching seed MS/MS spectra against the original LipiDex HCD Acetate spectral library (orange) and generated spectral library (blue). The difference between each group is displayed on the right with significantly improved changed dot product scores (>100) highlighted in blue or orange.

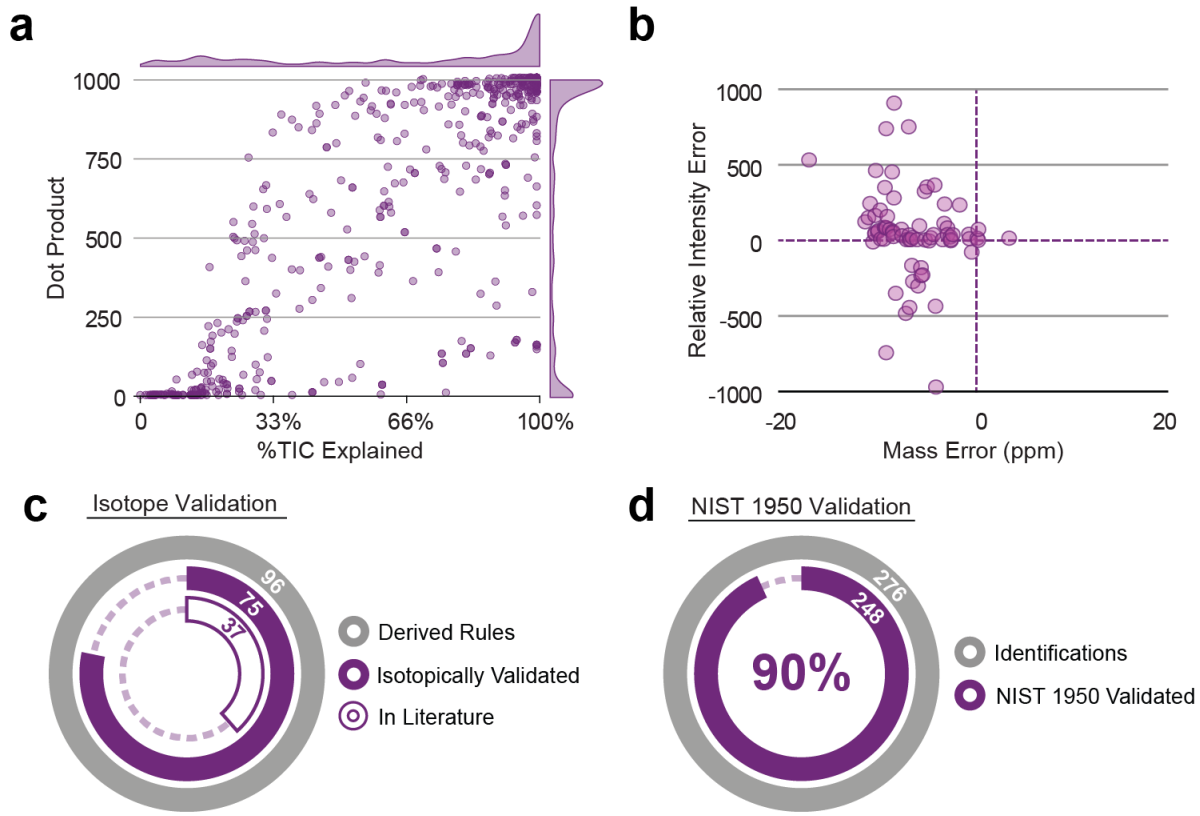


Supplementary Figure 41. NIST 1950 SRM Validation. Radial bar charts depicting the number of lipid identifications from LC-MS/MS analysis of the NIST 1950 SRM identified by at least three laboratories (solid color) and additionally hand-validated in **Supplementary Figure 42** (outlined) for the libraries generated *in silico* from complex HAP1 cell extract using the *LipiDex* spectral library (blue), complex HAP1 cell extract using the LipidBlast spectral library (magenta), and reference standard mix using HCD (green).


Alkenyl-DG P-36:5/Alkanyl-DG O-36:6 [M+H]+ m/z 599.501 +
m/z 339.28
18:4 NL
LipiDex
PC 42:8 [M+H]+ m/z 858.598 +
m/z 184.07
Headgroup
m/z 86.09
Headgroup - H3PO4
LipiDex
Standard
LipidBlast
Alkenyl-DG P-38:5/Alkanyl O-38:6 [M+H]+ m/z 627.532 +
m/z 341.30
20:5 NL
LipiDex
PC 44:9 [M+H]+ m/z 884.603 +
m/z 184.07
Headgroup
Potential False ID
LipiDex
Standard
LipidBlast
Alkenyl-DG P-36:4/Alkanyl-DG O-34:5 [M+H]+ m/z 601.516 +
Potential False ID
m/z 339.29
18:3 NL
LipiDex
PI 37:4 [M-H]- m/z 871.528 -
m/z 269.25
17:0 Frag
m/z 303.23
20:4 Frag
m/z 241.01
Headgroup
LipiDex
Standard
LipidBlast
Alkenyl-DG P-38:5/Alkanyl-DG O-38:6 [M+H]+ m/z 627.532 +
m/z 341.305
20:5 NL
LipiDex
Plasmanyl-PC O-31:1/Plasmenyl-PC P-31:0 [M+H]+ m/z 706.591 +
m/z 184.07
Headgroup
m/z 86.09
Headgroup - H3PO4
LipiDex
LipidBlast
Alkenyl-DG P-38:6/Alkanyl-DG O-38:7 [M+H]+ m/z 625.516 +
m/z 339.29
20:5 NL
LipiDex
Plasmanyl-PC O-38:7//Plasmenyl-PC P-38:6 m/z 788.555 +
Potential False ID
m/z 184.07
Headgroup
LipiDex
LipidBlast
Cer[NS] d43:1 [M+Ac-H]- m/z 722.664 -
Potential False ID
m/z 662.65
Acetate NL
m/z 395.39
26:0 + C2H3N1
LipiDex
LipidBlast
Plasmenyl-PE P-40:4/Plasmanyl-PE O-40:5 [M+H]+ m/z 780.588 +
m/z 361.27
20:4 Frag + C3H6O
m/z 420.32
20:4 NL -C3H6O
m/z 322.35
20:4 NL -C3H9PO5
LipiDex
LipidBlast
PC 42:10 [M+H]+ m/z 854.564 +
m/z 184.07
Headgroup
m/z 86.09
Headgroup - H3PO4
LipiDex
Standard
LipidBlast
Plasmenyl-PE P-36:5/Plasmanyl-PE O-36:6 [M+H]+ m/z m/z 722.511 +
m/z 359.26
20:5 Frag + C3H6O
m/z 364.26
20:5 NL -C3H6O
m/z 266.28
20:5 NL -C3H9PO5
LipiDex
LipidBlast
Relative Intensity
m/z



Supplementary Figure 42. NIST 1950 SRM Identification Manual Validation. MS/MS spectra for all identifications from LC-MS/MS analysis of the NIST 1950 SRM which were not identified by at least 3 labs from the interlaboratory study. Annotated spectral peaks (orange) which confirm identification are labeled with fragment type. The lipid spectral libraries which generated the identification are displayed adjacent to each spectrum.



Supplementary Figure 43. LDA 2 Identification Dataset. (A) Dot product returned and %TIC explained for seed MS/MS spectra after *in-silico* library generation from complex HAP1 cell extract using LDA 2 lipid identifications. For clarity, the dot product and %TIC explained are additionally displayed as density plots (bin width = 10) on each corresponding axis. (B) Scatter plot depicting the relative intensity error out of 1000 and mass error (ppm) for each derived fragmentation rule. (C) Radial bar chart depicting the number of fragmentation rules validated via isotopically-labeled standards (solid color) and present in external spectral libraries (outlined). (D) Radial bar chart depicting the number of lipid identifications from LC-MS/MS analysis of the NIST 1950 SRM identified by at least three laboratories (solid color).

References

1. Hutchins, P.D., Russell, J.D., Coon, J.J.: LipiDex: An Integrated Software Package for High-Confidence Lipid Identification. Cell Syst. 6, 621–625.e5 (2018). doi:10.1016/j.cels.2018.03.011