MS-DIAL 5 multimodal mass spectrometry data mining unveils lipidome complexities

Authors

Hiroaki Takeda[1,$], Yuki Matsuzawa[1,$], Manami Takeuchi[1], Mikiko Takahashi[2], Kozo Nishida[1], Takeshi Harayama[3,4,*], Yoshimasa Todoroki[1], Kuniyoshi Shimizu[1], Nami Sakamoto[1], Takaki Oka[1], Masashi Maekawa[5], Mi Hwa Chung[1], Yuto Kurizaki[1], Saki Kiuchi[1], Kanako Tokiyoshi[1], Bujinlkham Buyantogtokh[1], Misaki Kurata[1], Aleš Kvasnička[6,7], Ushio Takeda[8], Haruki Uchino[5,9], Mayu Hasegawa[10], Junki Miyamoto[10], Kana Tanabe[11], Shigenori Takeda[11], Tetsuya Mori[2], Ryota Kumakubo[1], Tsuyoshi Tanaka[1], Tomoko Yoshino[1], Makoto Arita[5,9,12,*], Hiroshi Tsugawa[1,2,9,12,*]

Affiliations

1. Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, 2-24-16 Nakamachi, Koganei-shi, Tokyo, 184-8588, Japan
2. RIKEN Center for Sustainable Resource Science, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
3. Institut de Pharmacologie Moléculaire et Cellulaire, Université Côte d'Azur - CNRS UMR7275 - Inserm U1323, 660 Route des Lucioles, 06560 Valbonne, France
4. Institute of Global Innovation Research, Tokyo University of Agriculture and Technology, 2-24-16 Nakamachi, Koganei-shi, Tokyo, 184-8588, Japan
5. Graduate School of Pharmaceutical Sciences, Keio University, Minato-ku, Tokyo, 105-8512, Japan
6. Laboratory for Inherited Metabolic Disorders, Department of Clinical Biochemistry, University Hospital Olomouc, Zdravotníků 248/7, 779 00 Olomouc, Czech Republic
7. Faculty of Medicine and Dentistry, Palacký University Olomouc, Hněvotínská 3, 779 00 Olomouc, Czech Republic
8. K.K. ABSciex Japan, Shinagawa Tokyo, 140-0001, Japan
9. RIKEN Center for Integrative Medical Sciences, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
10. Department of Applied Biological Science, Tokyo University of Agriculture and Technology, 3-5-8 Saiwai-cho, Fuchu, Tokyo 183-8509, Japan
11. Innovative Technology Laboratories, AGC Inc., 1-1 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan
12. Graduate School of Medical Life Science, Yokohama City University, Yokohama, Japan

$ Hiroaki Takeda and Yuki Matsuzawa contributed equally to this work.

Corresponding Authors

* E-mail: T.H. (harayama@ipmc.cnrs.fr), M.A. (marita@keio.jp) and Hiroshi T. (htsugawa@go.tuat.ac.jp)

ABSTRACT

Lipidomics and metabolomics communities comprise various informatics tools; however, software programs that can handle multimodal mass spectrometry (MS) data with structural annotations guided by the Lipidomics Standards Initiative are limited. Here, we provide MS-DIAL 5 to facilitate the in-depth structural elucidation of lipids through electron-activated dissociation (EAD)-based tandem MS, as well as determine their molecular localization through MS imaging (MSI) data using a species/tissue-specific lipidome database containing the predicted collision-cross section (CCS) values. With the optimized EAD settings using 14 eV kinetic energy conditions, the program correctly delineated the lipid structures based on EAD-MS/MS data from 96.4% of authentic standards. Our workflow was showcased by annotating the *sn*- and double-bond positions of eye-specific phosphatidylcholine molecules containing very-long-chain polyunsaturated fatty acids (VLC-PUFAs), characterized as PC n-3-VLC-PUFA/FA. Using MSI data from the eye and HeLa cells supplemented with n-3-VLC-PUFA, we identified glycerol 3-phosphate (G3P) acyltransferase (GPAT) as an enzyme candidate responsible for incorporating n-3 VLC-PUFAs into the *sn*-1 position of phospholipids in mammalian cells, which was confirmed using recombinant proteins in a cell-free system. Therefore, the MS-DIAL 5 environment, combined with optimized MS data acquisition methods, facilitates a better understanding of lipid structures and their localization, offering novel insights into lipid biology.

MAIN TEXT

Untargeted lipidomics has emerged as a crucial biotechnology approach, enabling comprehensive lipidomic analysis of various biospecimens[1]. Tandem mass spectrometry (MS/MS) of lipids ionized by electron spray ionization (ESI), followed by collision-induced dissociation (CID)-based fragmentation, provides detailed substructure information. This allows the characterization of lipid structures at the molecular species level, characterizing lipid subclasses in addition to carbon and double bond numbers in the individual acyl chains[2]. In addition, advanced techniques such as electron-based method[3], Paterno-Buchi reaction[4], ultraviolet photodissociation[5], and ozone- or hydroxyl-radical reactions[6,7] offer deeper insights into lipid structures by annotating the *sn*-position and double bond (C=C) locations. Moreover, spatial lipidomics, such as matrix-assisted laser desorption ionization (MALDI) coupled with MS, facilitate the determination of lipid molecule localizations[8]. Consequently, the screening, in-depth structural annotation, and spatial mapping of lipids are now feasible using state-of-the-art analytical chemistry tools. Given these advancements, the development of an informatics environment that fully leverages the potential of advanced MS techniques has become a pressing need, propelling lipid-centric biological research forward.

Despite the development of various informatics tools within the lipidomics and metabolomics communities[9], only a limited number of software programs can handle multimodal MS data with structural annotations guided by the Lipidomics Standards Initiative (LSI). In this study, we introduce MS-DIAL 5, an advanced environment that builds upon its predecessor, MS-DIAL 4[10], which supports diverse MS methodologies and has an improved user interface utility, as summarized in **Table 1**. This environment excels in multimodal MS data analysis, enabling the in-depth elucidation of lipid structure with electron-activated dissociation (EAD)[3] and facilitating spatial lipidomics through a tissue/species-specific lipid CCS database constructed using a machine learning method on datasets acquired from CID-based untargeted lipidomics studies.

Table 1. Informatics software and tools in lipidomics

Software name	MS-DIAL 5	MS-DIAL 4	Lipostar 2.0	LDA	LipidFinder	XCMS online	MZMine3	LipidHunter2
DI-MS (Full MS)	✓						✓	
DI-MS (DDA)	✓						✓	✓
DI-MS (DIA)	✓							
IM-MS (Full MS)	✓						✓	
IM-MS (DDA)	✓						✓	
IM-MS (DIA)	✓							
LC-MS (Full MS)	✓	✓	✓	✓	✓	✓	✓	
LC-MS (DDA)	✓	✓	✓	✓			✓	✓
LC-MS (DIA)	✓	✓	✓				✓	
LC-IM-MS (Full MS)	✓	✓	✓				✓	
LC-IM-MS (PASEF)	✓	✓	✓				✓	
LC-IM-MS (diaPASEF)	✓						✓	
Supporting EAD	✓							
MSI	✓						✓	

The summary was created based on the publication of Ni, Z. et al. Nat Methods 20, 193–204 (2023) with the actual website's information on December 11, 2023.
DI: direct or flow injection (for shotgun lipidomics), IM: ion mobility (IM), IM-MS: direct or flow injection with IM separation
LC: liquid chromatography, DDA: data dependent acquisition, DIA: data independent acquisition, MSI: mass spectrometry imaging

We initially assessed the informational content of EAD-MS/MS by examining the spectra of 716 unique small molecules, which revealed that a kinetic energy (KE) of 15 eV yielded the most MS/MS information, as determined by the spectrum entropy values and molecular spectrum networks[11,12] (**Figure 1a and Supplementary Figure 1**). Subsequently, we analyzed the MS/MS spectra of 65 lipids at KEs from 8 to 20 eV and selected 14 eV as the optimal KE for lipid structure analysis for three main reasons:

first, the sensitivity of the product ions was higher in the 14 eV KE than in those of 8–10 eV KE, which was utilized in the previous report as the optimal parameter[3] (**Figure 1b**). Second, while the V-shaped pattern, whose valley corresponds to the C=C position, is an important criterion for interpreting the C=C position, the MS/MS spectrum of the 14 eV KE maintains the pattern in various phospholipid subclasses (**Figure 1b-d and Supplementary Figure 2**). The potential mechanism for the increase in product ions abundance adjacent to the double bond is the stabilization of the fragment ion by the McLafferty rearrangement or allyl radical formation[13]. As the increased peaks can be utilized as marker ions for structure elucidation, they are termed "C=C high peak" in this study (see **Online Method**). Finally, the 14 eV KE condition provided unique diagnostic ions for characterizing polyunsaturated fatty acids (PUFA) in addition to V-shaped patterns. We discovered a significant increase in the hydrogen gain (H-gain) fragment ions in phospholipids with methylene-interrupted PUFAs containing more than three double bonds, including arachidonic acid (ARA) and docosahexaenoic acid (DHA) (**Figure 1d**). An increase in the abundance of H-gain fragment ions was observed under high KE conditions (14–18 KEs), which is likely due to the McLafferty rearrangement facilitating the removal of acidic protons from the methylene moiety between double bonds. The principle of C=C position determination in EAD relies on charge-remote fragmentation (CRF), producing three ion types—H-loss, radicals, and H-gain—at each carbon-carbon cleavage, complicating PUFA structural elucidation. Thus, the distinctive pattern of PUFA-specific H-gain fragments at the 14–18 eV KEs provides a key criterion for structure elucidation. Moreover, we observed benefits in the annotation of sphingolipids, with facilitated characterizations of hydroxy (OH) positions and *N*-acyl chain compositions, and in glycerolipids, enabling the discrimination of polar head isomers such as 1,2-diacylglyceryl-3-*O*-2'-(hydroxymethyl)-(*N*,*N*,*N*-trimethyl)-*β*-alanine (DGTA) and 1,2-diacylglyceryl-3-*O*-4'-(*N*,*N*,*N*-trimethyl)-homoserine (DGTS), as evidenced by their 14 eV KE spectra (see **Online Methods**, **Supplementary Figure 3, and Supplementary Figure 4**).

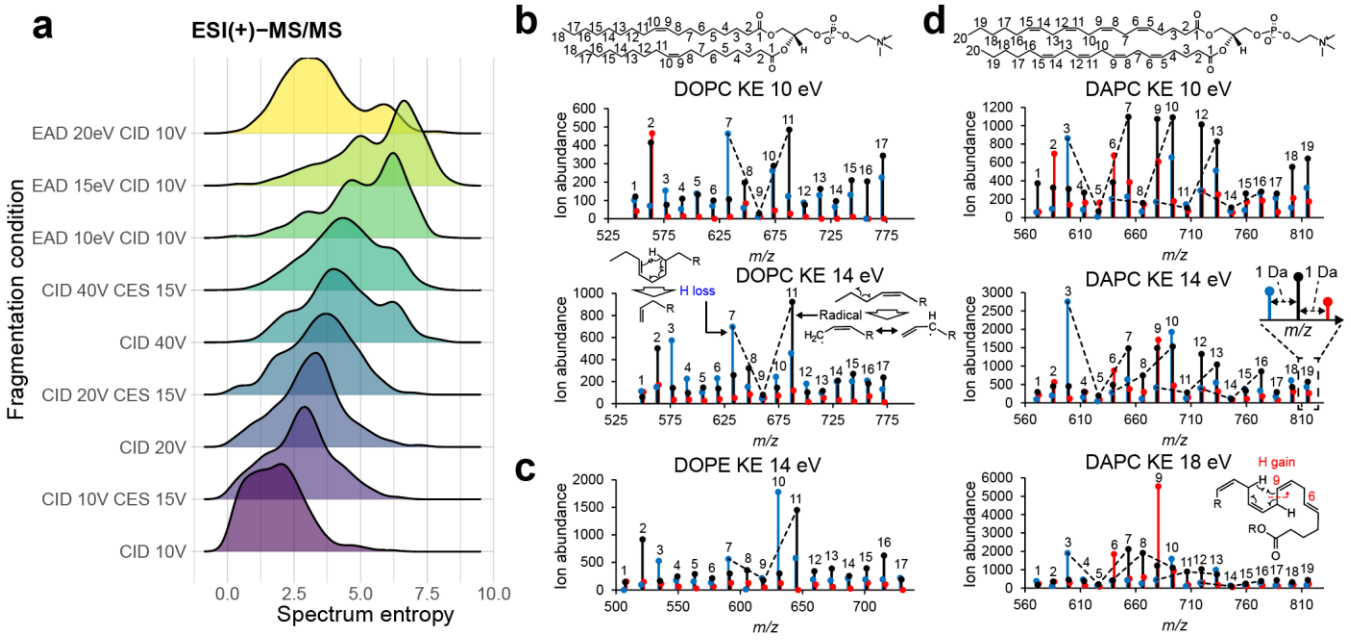


Figure 1. Electron-activated dissociation (EAD)-based tandem mass spectrum facilitates efficient lipid structure elucidation. (a) Spectrum entropy value distributions for 716 small molecules, with the x- and y-axes representing spectrum entropy and fragmentation conditions, respectively. (b) EAD-MS/MS

spectra of 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) at kinetic energies (KE) of 10 and 14 eV, highlighting only the hydrogen (H) loss (blue), radical (black), and H-gain (red) fragment ions related to acyl chain properties. The proposed mechanism explaining the increased abundance of H-loss and radical fragments is also depicted. (c) EAD-MS/MS spectrum of 1,2-dioleoyl-*sn*-glycero-3-phosphatidylethanolamine (DOPE) at a 14 eV KE. (d) EAD-MS/MS spectra of 1,2-diarachidonoyl-*sn*-glycero-3-phosphocholine (DAPC) at KEs of 10, 14, and 18 eV. The mechanism behind the observed increase in H-gain fragment abundance at the delta-6 and 9 carbon positions is also illustrated. Numbers atop each fragment ion denote the carbon count remaining in a single acyl chain.

Through detailed investigations of the MS/MS spectra, we developed a decision tree algorithm for the 14 eV KE spectra to elucidate lipid structures that was implemented in MS-DIAL 5. Our algorithm first annotated the molecular species level of lipids, such as PC 16:0_20:4 and ceramide (Cer) 18:1;O2/16:0. If only species-level annotations, such as PC 36:4 and Cer 34:1;O2, were assigned, no further details were examined. The *sn*-, OH-, and C=C positions were independently evaluated. For *the sn*- and OH-position assessments, the candidates were ranked based on the abundance of the diagnostic product ions, with no positional assignments made if they did not match the theoretical *m/z* values. The C=C positions within complex lipids were considered when product ions labeled as 'C=C high peak' were detected. Finally, candidates were ranked based on the local correlation coefficient between the experimental and computationally generated *in silico* spectra related to the acyl chains, which included heuristic H-loss, radical, and H-gain ions that reflect the V-shape and increase in PUFA-H-gain ions. To develop the MS-DIAL annotation algorithm, the peak heights of the diagnostic product ions were assessed using authentic standards of 1,2-dilinoleoyl-*sn*-glycero-3-phosphocholine (DLPC) and 1-palmitoyl-2-arachidonoyl-*sn*-glycero-3-phosphocholine (PAPC) (**Figure 2a**). The results suggested that a large amount of lipids need to be injected into the LC-MS system for determining *sn*- and C=C positions of complex lipids (500–1000 femtomoles for PC). Therefore, we designed an EAD spectral annotation program to rank candidates from the well-characterized lipid chemical space (see **Online Method** for the targeted chemical space), rather than to discover new structures. This enables one to perform comprehensive annotation of 27 subclasses, including 15 subclasses of glycerophospholipids (GP), five sphingolipids (SP), and seven glycerolipids (GL).

We evaluated the MS-DIAL program using LC-MS/MS data from a dilution series of a mixture containing 13 lipid standards and samples to which 110 lipid standards were added at various concentrations to the lipid extract of a stable isotope-labeled plant which was used for the background matrix (**Figure 2b and 2c**). EAD spectra were acquired in data-dependent acquisition (DDA) mode. The results demonstrated that both the *sn*- and C=C-positions in glycerophospholipids (GP) and both the OH- and C=C-positions in sphingolipids (SP) were achieved at high concentrations of PC, Cer, and sphingomyelin (SM). However, determining *sn*-positions for other GP lipid subclasses proved challenging, likely because of the reduced CRF reaction associated with less charge bias in the proton and ammonium ion forms[14]. The incorporation of metal ions, such as sodium, facilitates the acquisition of simpler spectra for structural elucidation owing to a stronger charge bias in the polar head moiety. As specialized sample preparation and analytical conditions are required to efficiently acquire the metal ion forms of lipids, a comprehensive evaluation of the entire analytical system for metal-ion-based lipid structure elucidation will be conducted in future work, while we offer a platform for EAD spectral annotation of sodium adduct forms in this study. As a result of program evaluation of lipid spectra of proton or ammonium adduct

forms, wherein MS-DIAL annotates lipids at species or molecular species levels with annotations of *the sn*-position, OH-position, and C=C-position based on the quality of MS/MS spectra, the program achieved a 96.4% accuracy rate by such decision-tree-based diagnostics in the annotation process. Misannotations of the C=C positions were often observed in triglycerides (TG), phosphatidylinositols (PI), and diglycerides (DG) with different acyl chains (**Supplementary Figure 5**). Except for the case in which all acyl chains were identical, the results indicated that the ammonium adduct form of non-cationic lipids provides inefficient CRF ions, suggesting the use of metal ions, such as $[M+Na]^+$, for structure elucidation. Although a misannotation of the *sn*-position was observed in PC-d5 17:0/16:1(9), the product ion spectrum could be interpreted as a mixture of *sn1*-17:0 and *sn1*-16:1 structures, even after manual inspection. This could be due to the synthesis of byproducts from the standard compounds. Furthermore, we assessed the effect of metabolite co-elution on lipid structure annotation using a mixture of DLPC and PAPC (**Supplementary Figure 6**). The structural description of PAPC was more affected by the presence of the DLPC spectrum, with proper annotation of PAPC occurring when its concentration was 5-10 times higher than that of DLPC. This effect was attributed to the identical *sn1/sn2* acyl chains of DLPC, which doubled the intensity of each product ion related to the acyl chains. Thus, our findings emphasize the importance of separation and lipid enrichment for in-depth structural elucidation using EAD.

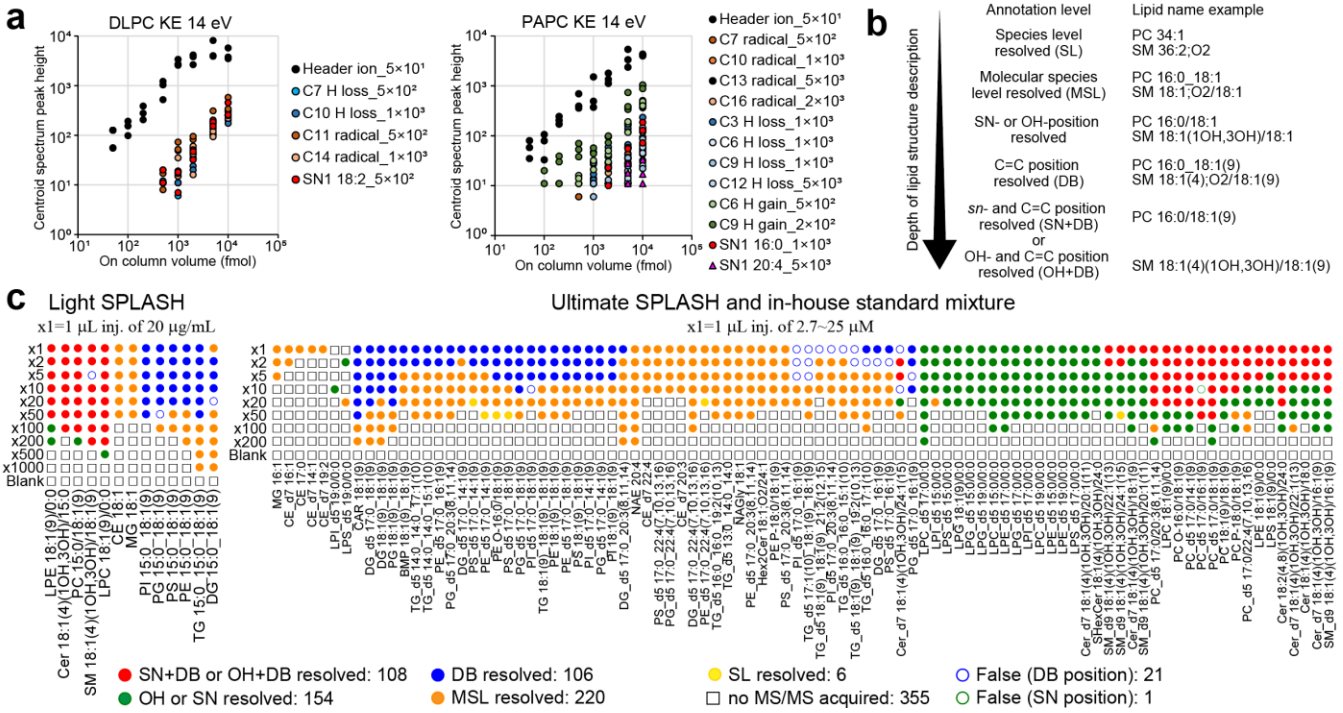


Figure 2. Evaluation of dynamic ranges for in-depth lipid annotation with MS-DIAL 5. (a) Dynamic range and limit of detection (LOD) required to confirm the presence of diagnostic ions for lipid structure elucidation. The x-axis represents the on-column volume of 1,2-dilinoleoyl-*sn*-glycero-3-phosphocholine (DLPC) and 1-palmitoyl-2-arachidonoyl-*sn*-phosphatidylcholine (PAPC), while the y-axis shows the peak heights of diagnostic ions from the centroided product ion spectrum obtained from the product ion scanning mode. The response values of important fragment ions were investigated; for instance, "C10 H loss_1x10^3" and "SN1 18:2_5x10^2" denote the LOD values of the H-loss fragment ion at the C10 position and the neutral loss (NL) of *sn1*-18:2 to characterize the *sn*-position at 1,000 and 500 femtomoles (fmol)

on-column volumes, respectively. Even in the authentic standard of PAPC, ions related to *sn*1-20:4 are detected due to chemical impurities or conformational changes during sample preparation. (b) Relationship between annotation level terminology and lipid description. The cases of PC and SM were described. (c) Validation of the MS-DIAL 5 environment for lipid structure description based on EAD-MS/MS spectra quality. Dilution series were analyzed three times at each concentration. The representative annotation was determined as follows: if the same lipid name was annotated in at least two of the three replicates, that name was used as the representative annotation. If the annotation results differed across all three replicates, the lipid with the highest score was adopted as the representative. For example, "x10" indicates a dilution 10 times less concentrated than the original, denoted as "x1." For sphingolipids, green and red circles represent annotations where OH-positions and both OH- and C=C positions are resolved, respectively. For glycerophospholipids, green and red circles indicate annotations of resolved *sn*-positions and both *sn*- and C=C-positions, respectively. Blue, orange, and yellow circles represent annotations at the C=C position resolved, molecular species, and species levels, respectively. If the MS/MS spectrum was not assigned to the precursor ion by DDA, a square shape is used. Incorrect annotations are shown as white fills with a border color indicating the source of the misannotation.

We demonstrated the capabilities of our in-depth lipidomic platform by characterizing phospholipids containing very long-chain polyunsaturated fatty acids (VLC-PUFAs) in the retinal tissue of mice (**Figure 3a**)[15]. Importantly, the VLC-PUFAs are mostly contained in PC in the tissue according to our investigation. Through an optimized solid-phase extraction procedure, we achieved an in-depth structural elucidation of 250 peaks in total and characterized 3, 20, and 10 molecules of VLC-PUFA PC at the molecular species, *sn*-position resolved, and both *sn*- and C=C-position resolved levels, respectively. The most abundant peak in the VLC-PUFA PC fraction was characterized as PC 34:6(16,19,22,25,28,31)/22:6(4,7,10,13,16,19). Moreover, our results indicated that all the top hit candidates with *sn*- and C=C-positional information contained n-3-VLC-PUFA at the *sn1*-position. Previous indirect evidence for VLC-PUFA PC structures using phospholipase enzymes and gas chromatography-MS predicted the major forms of *sn*- and C=C-positions to be *sn1* and methylene-interrupted n-3 fatty acid, respectively[16]. In contrast, the present study is the first to directly identify structures in their native form, and our result strongly suggests that the VLC-PUFAs are enriched at the *sn1*-position of PC in the retinal tissue. To investigate the localization of VLC-PUFAs, we reanalyzed a public MALDI-MSI dataset of eye tissues from C57B6/J and acyl-coenzyme A synthetase (ACSL) 6 knockout mice, where ACSL6 is known to have substrate specificity for DHA, by utilizing an eye-specific lipid database containing predicted CCS values generated by a machine learning model and publicly available CID-based untargeted lipidomic data (**Figure 3b and 3c**)[10,17]. The analysis revealed that the VLC-PUFA PC containing stearic acid, annotated as PC 34:6(16,19,22,25,28,31)/18:0, was not significantly reduced in ACSL6 KO mice, which was consistent across both MSI and untargeted lipidomics data (**Figure 3d**). These results suggested that n-3-VLC-PUFA does not undergo the same enzymatic substrate recognition as DHA. Furthermore, n-3-VLC-PUFA incorporation into lysophosphatidic acid (LPA), PC, diacylglycerol (DG), triacylglycerol (TG), and cholesteryl ester (CE) was confirmed in HeLa cells when n-3-VLC-PUFA was supplemented to the culture medium (**Figure 3e**). These findings indicate that n-3-VLC-PUFA is not incorporated into phospholipids by retina-specific enzymes. Instead, the structure of n-3-VLC-PUFA, with more than 32 carbons, resembles that of saturated fatty acids, such as palmitic acid, from the carboxylic acid terminus to the first C=C-position. We

hypothesized that VLC-PUFAs are recognized by glycerol 3-phosphate acyltransferase (GPAT), which prefers saturated fatty acids as substrates and incorporates an acyl chain at the *sn1*-position[18]. A cell-free system assay[19] showed that VLC-PUFAs are converted to VLC-PUFA-LPAs by recombinant GPAT1 ($GPAT1^{WT}$) which is highly expressed in the mouse retina (**Supplementary Figure 7**), where acyl-coenzyme A (CoA) is synthesized by an acyl CoA synthetase in a cell-free system using wheat extract (**Figure 3f and 3g**). In contrast, the LPA molecule was not synthesized by the inactive mutant ($GPAT1^{H230A}$), the DHA-LPA molecule was not detected, and the DHA-CoA metabolite was synthesized at high levels in the cell-free system compared to those of the VLC-PUFA-CoA molecule. While the enzymatic pathways for VLC-PUFA incorporation into lipids were previously unknown, our results suggest enzymes with preferences for saturated substrates, such as GPAT1, mobilize VLC-PUFAs. The lack of double bonds near the carboxyl group makes VLC-PUFAs appear more 'saturated-like,' potentially contributing to their dominant localization at *sn1*-positions by GPATs. We believe that the discovery that VLC-PUFAs are incorporated into PC via GPAT is challenging to validate because the knockout or knockdown of all GPATs in cells or organisms is lethal. Nevertheless, once the enzymes responsible for VLC-PUFA uptake in the lipid remodeling pathway are identified, the contribution and role of de novo and remodeling pathways in VLC-PUFA PC biosynthesis in the eye can be elucidated.

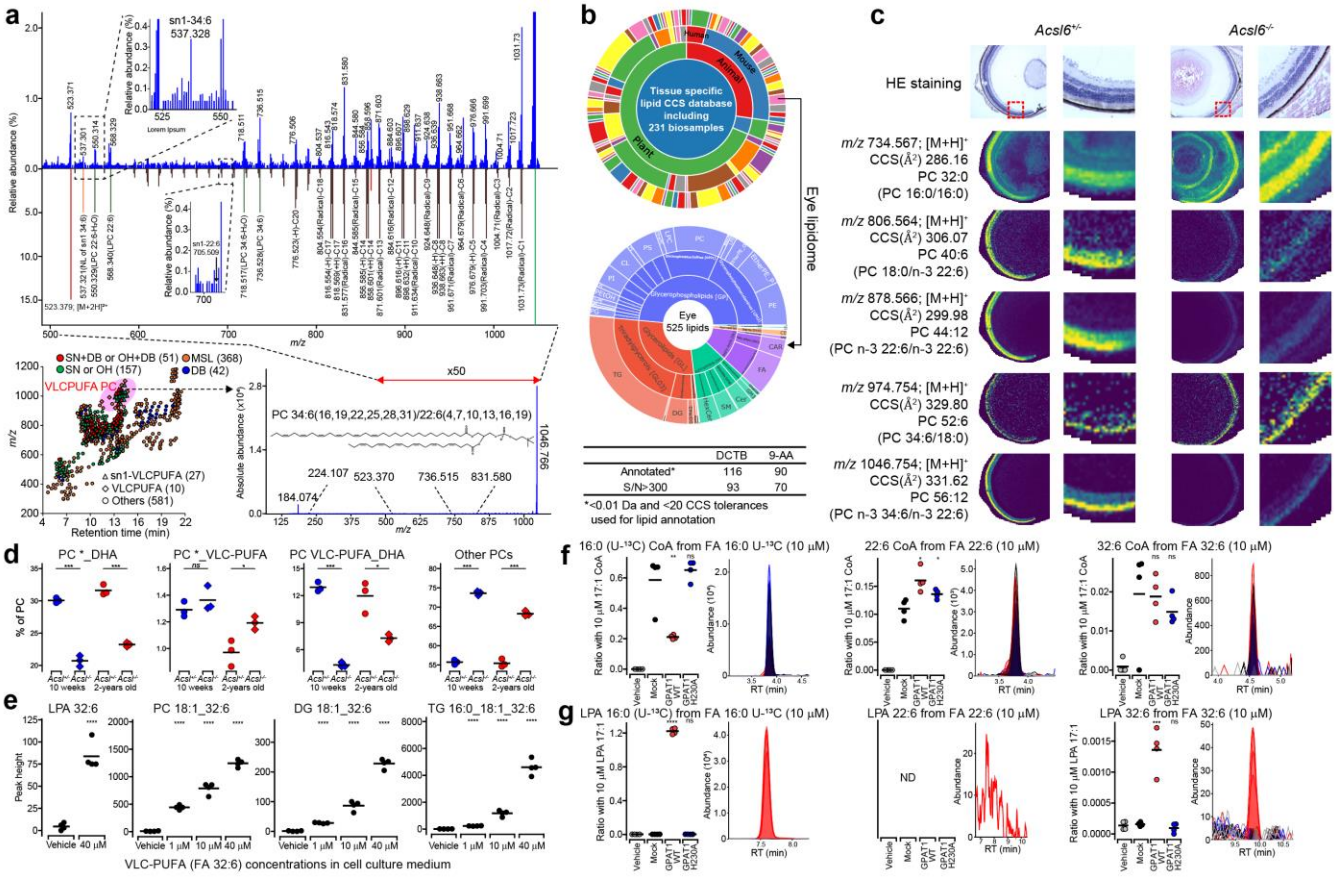


Figure 3. Structural and pathway elucidation of very long-chain polyunsaturated fatty acid (VLC-PUFA)-containing phosphatidylcholine (PC). (a) EAD-MS/MS structure elucidation with MS-DIAL annotation. The result of in-depth lipidome profiling is shown by the scatter plot of retention time- and *m/z* axis. The annotation results of molecular species (MSL), double-bond (DB) resolved, sn- or OH-

positions (SN or OH) resolved, and both sn- and DB- or both OH- and DB-position (SN+DB or OH+DB) resolved levels are described by the same color charts used in Figure 2b. The sn1-position determined or uncharacterized for VLC-PUFA are described by triangle and diamond symbols, respectively. The bottom-right panel describes the experimental spectrum of the lipid ion annotated as PC 34:6(16,19,22,25,28,31)/22:6(4,7,10,13,16,19), where the *E/Z* isomer definition in acyl chains is unsupported, yet a representative form is shown. The top panel displays a 50-fold zoomed experimental spectrum and a 10-fold zoomed *in silico* spectrum of the assigned lipid in the upper and lower panels, respectively. Brown, green, orange, and red spectral peaks represent ions related to homolytic cleavages in acyl chains, lyso PC substructures, neutral loss of *SN1*-34:6 moiety, and precursor- or polar head-specific fragments, respectively. The precursor *m/z* value in the survey MS1 scan was 1046.755, with a theoretical value of 1046.757. (b) A sunburst plot summarizing species/tissue-specific lipid database statistics containing collision-cross section (CCS) values. The database comprises lipidomes from 231 biosamples, including humans, mice, plants, and microorganisms. An eye-lipidome table with *m/z* and CCS values for 525 unique lipids was used to annotate lipids in MSI data analysis. A summary table of peak annotations in the analyzed MSI data is also provided. (c) Hematoxylin and eosin (HE) staining and MSI data in eye tissues from *Acsl6*$^{+/-}$ and *Acsl6*$^{-/-}$ mice. Ion distributions for five lipid molecules are shown, with annotations performed within 0.01 Da and 20 Å^2 tolerances from the *m/z* and CCS references, respectively. The reference *m/z* and CCS values for each lipid molecule are listed, with EAD-MS/MS-based annotations for each precursor *m/z* value in parentheses. (d) Reanalysis of publicly available LC-CID-MS/MS-based untargeted lipidomics data examining eye tissues from *Acsl6*$^{+/-}$ and *Acsl6*$^{-/-}$ mice at 10 weeks and 2 years of age. Here, “22:6” denotes DHA, while “28:6,” “30:5,” “32:4,” “32:5,” “32:6,” “34:4,” “34:5,” “34:6,” “36:6,” and “38:6” are defined as VLC-PUFAs. An asterisk indicates acyl chains other than DHA and VLC-PUFA, with the sum of lipid molecules labeled '*_DHA' or '*_VLC-PUFA'. 'Other PCs' refers to the total abundance of PC molecules not containing DHA or VLC-PUFA. (e) HeLa cell lipid profiling with VLC-PUFA (FA n-3-32:6) supplementation. The normalized lipid abundances of PC 18:1_32:6, DG 18:1_32:6, LPA 32:6, and TG 18:1_18:1_32:6 at final concentrations of 1, 10, or 40 μM of FA n-3-32:6 supplementation are depicted. While LPA was analyzed by a derivatization method using trimethylsilyl-diazomethane, which converts LPA to bis-methyl LPA (BisMeLPA), other molecules were analyzed using conventional untargeted lipidomics methods. (f, g) Acyl CoA (f) and LPA (g) profiling for the glycerol 3-phosphate acyltransferase 1 (GPAT1) recombinant enzyme assay. The acyl CoAs and LPAs were analyzed with vehicle, mock (native plasmid vector), active GPAT1^{WT}, and the inactive GPAT1 mutant (GPAT1^{H230A}), supplied with glycerol 3-phosphate and coenzyme in addition to ^{13}C-uniformly labeled palmitic acid (FA 16:0 U-^{13}C), docosahexaenoic acid (DHA, FA 22:6), or FA n-3-32:6, in the cell-free system enzymatic reaction. The fatty acid was supplied at final concentration of 10 μM, and the same amount of 17:1 CoA (Figure 3f) and LPA 17:1 (Figure 3g) was supplied as the internal standards. The putative ratio between the converted product and the internal standard was used for the y-axis value of dot plots. Significances were adjusted by false discovery rate in the student's t-test, with the following definitions: ns, P>0.05; *, P<0.05; **, P<0.01; ***, P<0.001; ****, P<0.0001.

Lipidomics has become an essential tool in systems biology, and is widely used in basic research and clinical studies. Despite structural complexity of lipids, the multimodal mass spectrometry techniques allow one to illuminate the diversity of lipids by using various methods, including untargeted analysis, in-depth structural elucidation through fragmentation methods, and spatial lipidomics. Concurrently, with

the evolution of measurement techniques, the development of informatics technology has become indispensable. Since 2015, we have been developing MS-DIAL[20], enhancing not only the algorithm based on feedback from the metabolomics and lipidomics community, but also creating a user-friendly interface for beginners and mature scientists. In this study, we demonstrated that MS-DIAL enables straightforward knowledge extraction from EAD spectra, spatial lipidomics, and publicly available untargeted lipidomic data, leading to new insights into lipid biology. Our group aims to continue contributing to data standardization in various omics sciences and developing data-driven knowledge generation platforms, facilitating machine learning and natural language processing research utilizing omics data.

Data availability

The spectral data of the authentic standards are available in MSP format files on the RIKEN DROPMet website (http://prime.psc.riken.jp/menta.cgi/prime/drop_index) under the index number DM0054, and the details are described in the readme file. All raw LC-MS data were available with the same index number as the RIKEN DROPMet. The Source Data for the figures is also available in the Supplementary Information. The MS-DIAL source code is available at https://github.com/systemsomicslab/MsdialWorkbench. The MS-DIAL tutorial and the demonstration data are available at https://systemsomicslab.github.io/msdial5tutorial/ and https://zenodo.org/communities/msdial, respectively.

Authors contributions

M.A. and Hiroshi T. designed this study. Hiroaki T., Manami T., Y. T., and Hiroshi T. performed the LC-MS/MS analyses. Y.M., M.T., K.N., and Hiroshi T. developed MS-DIAL, and Manami T., K.N., Y.T., T.O., Y.K., S.K., and A.K. created the tutorial, and contributed to MS-DIAL 5 improvements. U.T. provided technical support for the EAD-MS/MS experiments. M.H. and J.M. performed the mouse experiments. N.S. performed the imaging MS data analysis. Manami T., T.H., K.S., M.M., and M.H.C. performed cloning and enzyme assays of GPAT proteins. H.U., K.T., S.T., T.M., K.S., T.T., and T.Y. prepared authentic standards and biologically created definable lipid molecules. H.T. wrote the manuscript. All authors have thoroughly discussed this project and helped improve the manuscript.

Acknowledgments

This study represents a portion of the dissertation submitted by Yuki Matsuzawa to the Tokyo University of Agriculture and Technology in partial fulfillment of the requirement for his Ph.D. This study was supported by the JSPS KAKENHI (21K18216, H.T.), National Cancer Center Research and Development Fund (2020-A-9, H.T.), AMED Japan Program for Infectious Diseases Research and Infrastructure (21wm0325036h0001, H.T.), AMED Brain/MINDS (JP15dm0207001, H.T. and H.T.), JST National Bioscience Database Center (NBDC, H.T.), JST ERATO “Arita Lipidome Atlas Project” (JPMJER2101, M.A. and H.T.) and Technologically Advanced research through Marriage of Agriculture and engineering as Groundbreaking Organization (TAMAGO to J.M. and H.T.). TH was supported by the ATIP-Avenir program (CNRS/Inserm) and the Global Innovation Research funds of Tokyo University of Agriculture and Technology.

Supplementary Figures

Figure S1. Results of molecular spectrum networking of 716 metabolites based on the spectra from collision-induced dissociation (CID) and electron-activated dissociation (EAD). (a) Summary of nodes connectivity. If two nodes have the same ontology term, the count is incremented. The ontology terms were generated by the ClassyFire program. The term "parent" means the direct parent term of metabolite defined by the ClassyFire program. (b) Molecular networks based on the spectra of CID 40±15 V (left panel) and kinetic energy (KE) 15 eV with CID 10 V.

Figure S2. Relationships between the kinetic energy 14 eV spectrum and the lipid structure in the glycerophospholipids category. The top and bottom panels show the experimental- and computer-generated spectrum in MS-DIAL, respectively. NL means neutral loss. The spectra of cardiolipin (CL) and hemi-BMP are also described to clarify the inadequate information in the product ion spectra for the determination of *sn*- and C=C-positions.

Figure S3. Relationships between the kinetic energy 14 eV spectrum and the lipid structure in the glycerolipids and fatty acyls categories. The layout and the terms used are the same as those in Figure S2. For free fatty acid (FA) and FAHFA, the spectra of the derived forms using 2-dimethylaminoethylamine (DMED) are described.

Figure S4. Relationships between the kinetic energy 14 eV spectrum and the lipid structure in the sphingolipids category. The layout and the terms used are the same as those in Figure S2. For the ceramide-AS type containing alpha-hydroxy fatty acid as the *N*-acyl chain, the OH position in the *N*-acyl chain was not characterized, while the OH positions in the sphingobase moiety were characterized.

Figure S5. Details of misannotations in the molecules of phosphatidylcholine (PC), phosphatidylinositol (PI), and triacylglycerol (TG). (a) The MS/MS spectrum of the protonated form of PC-d5 17:0/16:1(9). The spectra of the entire (bottom panel) and zoomed regions (top panel) are shown, where the diagnostic ions of *m/z* 466.3246 and *m/z* 482.3559 that determine the *sn*-position for *sn1*-17:0 and *sn1*-16:1, respectively, are described. The annotation was incorrect due to a lower abundance of the ion related to sn1-17:0 than that of sn1-16:1. (b) The MS/MS spectrum of the protonated form of PC-d5 17:0/ 22:4(7,10,13,16) was correctly annotated, although the contamination of sn1-22:4 related ion existed. The spectra of the entire (bottom panel) and zoomed regions (top panel) are shown, where the diagnostic ions of *m/z* 544.3715 and *m/z* 482.3559 that determine the *sn*-position for *sn1*-17:0 and *sn1*-22:4, respectively, are described. (c) MS/MS spectrum of the ammonium adduct form of PI 18:1(9)/18:1(9). The upper and lower panels show the experimental- and *in silico* MS/MS spectra. The V-shape pattern of product ion spectrum that determine the C=C-position as C9 was described. (d) MS/MS spectrum of the ammonium adduct form of PI-d5 17:0/16:1(9). The correct annotation is PI-d5 17:0/16:1(9), while the MS-DIAL program annotated the spectrum as PI-d5 17:0/16:1(7) due to the absence of a C=C high peak for 16:1(9). (e) MS/MS spectrum of the ammonium adduct form of TG 18:1(9)_18:1(9)_18:1(9), where the V-shape pattern for 18:1(9) is also described. (f) MS/MS spectrum of the ammonium adduct form of TG-d5 16:0_16:0_17:1(10). The spectrum was misannotated because the local correlation value for 17:1(5) was higher than that of 17:1(9). The V-shape patterns for 17:1(5) and 17:1(9) are described while the in silico MS/MS spectrum of TG-d5 16:0_16:0_17:1(10) is described in the lower panel.

Figure S6. Annotation results in the co-elution situation of two lipids having the same *m/z* value. The left panel shows the result of "DLPC fixed" where PAPC concentrations varied at 0.1, 0.2, 0.5, 1.0, 2.0, 5.0, and 10 μM. The right panel shows the result of "PAPC" fixed where DLPC concentrations adjusted to 0.1, 0.2, 0.5, 1.0, 2.0, 5.0, and 10 μM. The term "Full description" means that both *sn*- and C=C-positions were characterized by the MS-DIAL program. The black and gray colors indicate the mis-annotation not providing DLPC no PAPC.

Figure S7. Expression levels of GPAT genes. The GPAT expression levels were downloaded from http://biogps.org/?full#goto=welcome on January 21, 2024. When multiple data sets were available for the expression tables, the data set with the highest expression level for each GPAT enzyme was selected.

Figure S8. Coomassie brilliant blue (CBB) staining for recombinant proteins.

Supplementary Tables

Table S1. Detail of 953 authentic standards.
Table S2. Detail of lipid authentic standards or biologically created metabolites.
Table S3. Details of LightSPLASH and the lipid description that can be characterized in this study.
Table S4. MS-DIAL parameters used for lightsplash- and ultimatesplash/in-house standards spectral data.
Table S5. Details of ultimate splash, in-house standards, and lipid description characterized in this study.
Table S6. Database of species/tissue-specific m/z and collision-cross section values of lipids.
Table S7. MS-DIAL parameters used in the untargeted lipidomics for public data.
Table S8. Base sequence and primer details in GPAT1 enzyme assay.

Supplementary Notes

Note 1. Lipidomics minimal reporting checklist for MS-DIAL EAD spectral annotation.
Note 2. Lipidomics minimal reporting checklist for characterization of very long chain PUFA (VLC-PUFA) containing PC in the eye tissue of mice using EAD.
Note 3. Lipidomics minimal reporting checklist for HeLa lipid profiling for HeLa cells with the supplementation of VLC-PUFA (FA 32:6).
Note 4. Lipidomics minimal reporting checklist for lysophospholipid profiling by using trimethylsilyl-diazomethane for HeLa cells.
Note 5. Lipidomics minimal reporting checklist for lysophospholipid profiling by using trimethylsilyl-diazomethane for GPAT enzyme assay.
Note 6. Lipidomics minimal reporting checklist for acyl-CoA profiling for GPAT enzyme assay.

ONLINE METHODS

MS-DIAL development environment

The MS-DIAL development environment was redesigned to enhance scalability, sustainability, and community contribution to the program package. The software was programmed in C#. The underlying algorithm was constructed in . NET Standard 2.0 framework, while the user interface is developed using the Windows Presentation Foundation (WPF) and follows the Model-View-ViewModel (MVVM) architecture. As a result, MS-DIAL functions as an operating system (OS)-independent command line tool, although its graphical user interface is available only for Windows OS. The source code is publicly accessible from the GitHub repository (https://github.com/systemsomicslab/MsdialWorkbench). The mzML[21] and netCDF parsers are deposited as a NuGet-package; however, owing to licensing restrictions, raw data providers for proprietary formats, such as WIFF, RAW, and .D remain private. Nonetheless, MS-DIAL supports the direct import of vendor formats, such as SCIEX, Bruker, ThermoFisher, Shimadzu, Waters, Agilent, and Kanomax. MS-DIAL employs the vendor's API to convert the original profile spectra into centroid spectra. The raw data parser is designed to retrieve centroid spectra. Consequently, users importing vendor-specific raw data into MS-DIAL should select “centroid” as the data acquisition type for both MS1 and MS2. Conversely, when working with mzML- and Analysis Base File (ABF) format data containing profile mode spectra, users should select “profile.”

MS-DIAL 5 major functional updates when compared to MS-DIAL 4

The latest update of MS-DIAL includes significant enhancements to both the algorithmic (backend) and user interface (front-end) components. The software supports data processing for direct infusion mass spectrometry (DI-MS), direct infusion coupled with ion mobility separation (IM-MS), and mass spectrometry imaging (MSI) data. Additionally, it accommodates the data-independent acquisition of IM-MS/MS and LC-IM-MS/MS, such as the diaPASEF (parallel accumulation-serial fragmentation) used in Bruker instruments[22]. For metabolite annotation, the program includes an annotation pipeline for lipidomics employing oxygen attachment dissociation (OAD)[7] and electron-activated dissociation (EAD)[3], with the algorithm for EAD-based lipid structure elucidation being validated in this study (refer to subsequent sections).

MS-DIAL 5 is designed to cater to a wide range of users, ranging from novices to experienced analysts. Here, we introduce the front-end features tailored for seasoned analysts. Unlike MS-DIAL 4, which allows only one library file for metabolite annotation, version 5 has no such limitation. Users can load multiple MSP files containing spectra from standards generated in their laboratories, public and commercial libraries such as MassBank[23] or NIST, and in silico spectra produced using tools such as CFM-ID[24]. These files can be searched for various parameters and assigned priority levels, thereby enhancing the accuracy of the metabolite annotations and reducing false annotations in untargeted analyses.

Additionally, a companion application, "rawdataviewer.exe," is bundled with the MS-DIAL software packages. This utility provides a platform similar to SeeMS from the ProteoWizard community[21], allowing users to view raw data and adjust peak-picking parameters. For instance, the default threshold for the 'minimum peak amplitude,' which is a critical peak-picking parameter, was set to 1000. However, this value may not be optimal for instruments such as Orbitrap MS and FT-ICR-MS, where suitable thresholds often range from 10,000 to 100,000. The optimal thresholds also vary depending on the number of biospecimens and the sample matrix background. Thus, assessing the relationship between the number

of detected peaks and the "minimum peak amplitude" threshold is crucial in untargeted analyses. Furthermore, exploring the relationship between the peak height/area and the signal-to-noise ratio is vital. The "rawdataviewer.exe" application facilitates access to this information. For detailed descriptions of other utilities, please visit https://systemsomicslab.github.io/msdial5tutorial/.

Investigating the kinetic energies to acquire information-rich EAD spectra of small molecules

MS/MS spectral records for 953 standard compounds, including the IROA large-scale metabolite library (https://www.iroatech.com/large-scale-metabolite-library-of-standards/) and an in-house natural product compound library which have been reported previously[25] were investigated. Chemical ontologies were defined using ClassyFire program[26]. These records were acquired under various fragmentation conditions (**Supplementary Table 1**). For EAD data acquisition, a liquid chromatography (LC) system consisting of a SCIEX Exion LC system and mass spectrometry (MS) detection of molecules were performed using quadrupole/time-of-flight MS (ZenoTOF 7600; SCIEX, Framingham, MA, USA). An InertSustainSwift C18 column (30 mm × 2.1 mm; 3 μm) from GL Sciences, Japan, was used, maintained at 40 °C with a flow rate of 0.2 mL/min. The mobile phases were composed of (A) water with 0.1% formic acid and (B) acetonitrile (ACN) with 0.1% formic acid, utilizing an isocratic mode of 80% B. Sample temperature was kept at 4 °C. A targeted MS/MS scanning mode, referred to as "MRM HR" by SCIEX, was used. Target precursor *m/z* values were calculated for the protonated form of each molecule. The TOF mass range was set from a starting mass of 50 to an end mass equal to the precursor *m/z*. Nine fragmentation conditions were explored: three in collision-induced dissociation (CID) mode with collision energy (CE) set at 10, 20, and 40 volts (V) with no CE spread, three in CID mode with the same CEs but with a CE spread of 15 V, and three in EAD mode with CE set at 10 V and electron kinetic energy (KE) at 10, 15, and 20 electron volts (eV), all with a CE spread of 0. The accumulation times were 50 and 100 ms for the CID and EAD modes, respectively. In the EAD mode, the electron beam current and electron transfer coefficient (ETC)% were set to 7000 nA and 100%, respectively. Other parameters included: ion source gas 1 at 50; ion source gas 2 at 50; curtain gas at 35; CAD gas at 7; temperature at 450 °C; spray voltage at 5500 V; and declustering potential at 80 V. Mass calibration was conducted automatically using the SCIEX Calibration Delivery System (CDS). Finally, the spectra of 716 chemicals were successfully captured and accessible at the RIKEN DROPMet.

The acquired spectrum was processed as follows: MS/MS spectral peaks accumulated across the elution fraction from the left to the right edge of the molecule's peak. A bin size of 0.05 Da was set for accumulation, and the average intensity for each bin was used as the representative value. Subsequently, the spectrum from a retention time of approximately 30 s, where no peaks were observed, was accumulated using the same method and subtracted from the compound spectrum. If the subtracted intensity was below zero, it was replaced with zero. The EAD-MS/MS spectra data was assessed using the spectrum-entropy calculation method[11]. Density plots for each fragmentation condition were visualized using ggplot2 and related packages in the R programming environment. In addition, the results of molecular spectral networking from CID at 40 V with a collision energy spread (CES) of 15 and EAD at 15 eV are detailed. The MS/MS spectral similarity among the compounds was calculated using a modified dot product score as previously reported[12]. The parameters for this calculation were as follows: relative abundance cutoff of 0.1%, absolute abundance cutoff of 50, product ion mass tolerance of 0.05, mass binning value of 1.0, intensity scale factor of 0.5, and a maximum scale value of 100. The source code for this process can be found in the MsScanMatching.cs file in the MsdialWorkbench repository.

Investigating the kinetic energies to obtain information-rich EAD spectra for lipid structure elucidation

MS/MS spectral records were obtained for 34 glycerophospholipid (GP), 18 molecules of sphingolipid (SL), 9 molecules of glycerolipid (GL), 1 molecule of sterol lipid (SL), and 3 molecules of fatty acyls (FA) molecules (**Supplementary Table 2**). The mobile phases included (A) a mixture of ACN, methanol (MeOH), and water in a 1:1:3 (v/v/v) ratio with 5 mM ammonium acetate and 10 nM ethylenediaminetetraacetic acid (EDTA) and (B) a 1:9 (v/v) mixture of ACN and isopropanol (IPA) with 5 mM ammonium acetate and 10 nM EDTA. The other liquid chromatography conditions remained unchanged. Nine fragmentation conditions were explored during targeted MS/MS scanning mode, including one CID setting (45 eV) and various EAD settings with different kinetic energies (8, 10, 12, 14, 16, 18, and 20 eV KE). A constant CE spread of 0, collision energy of 12 for EAD, and a time-of-flight (TOF) start mass of 170 were applied. The accumulation time was 100 ms for both CID and EAD modes. The ion source temperature was maintained at 275 °C, while other parameters were consistent with those described above.

The charge remote fragment (CRF) ion pattern of the lipid molecules was elucidated using the EAD-MS/MS spectra of the three lipid metabolites, as shown in **Figure 1** of the main text. CRF ions arise from the homolytic cleavage of chemical bonds, producing three types of fragment ions: hydrogen loss (H-loss), radical, and hydrogen gain (H-gain) derived from each of carbon-carbon bond cleavage. As demonstrated in a previous report[3], a KE of 10 eV yields a distinct V-shaped pattern in the product ion intensities around the double-bond position. For instance, the ion abundance at the "C9" position from the C9-C10 bond cleavage in dioleoyl PC (DOPC) is reduced, while the ion abundances at the "C11" and "C7" positions from the C11-C12 and C7-C8 bond cleavages, respectively, are increased (**Figure 1b**). The term "C=C low peak" is used in this paper to denote the low-intensity peak. The amplified abundance of radical ions at the "C11" position is interpreted as stabilization of the fragment ion by resonance structure formation. Additionally, the significant increase in the C7 fragment ion intensity suggests that the hydrogen of C6 is more readily transferred by the electron pair of the double bond between C9 and C10 because of the McLafferty rearrangement. Such markedly increased fragment ions are referred to as "C=C high peak" in this study. Fragment ion abundances in EAD-MS/MS became more pronounced at KE 14 eV than at KE 10 eV, while preserving the V-shaped pattern. Comparable patterns were observed in dioleoyl PE (DOPE) and diarachidonoyl PC (DAPC) (**Figure 1c and 1d**). However, when complex lipids include polyunsaturated fatty acids (PUFAs) such as DAPC, pinpointing the valley in the V-shaped product ion pattern is challenging owing to increased spectral complexity.

In the EAD-MS/MS spectrum derived from the KE 18 eV condition, unique fragmentation patterns emerged that were not present at KE 10 eV. Specifically, in the KE 18 eV MS/MS spectrum of DAPC, there was a marked increase in the abundance of the hydrogen gain (H-gain) fragment ion at the C9 position (**Figure 1d**). This phenomenon of increased H-gain fragment ion is also notable in structures like linolenic acid, arachidonic acid, and docosahexaenoic acid, which possess multiple methylene-interrupted C=C bonds, confirmed by the *Z* isomer. The hydrogen atom on the methylene group situated between the double bonds is more acidic than the typical C-H bond, which is known to be an acidic proton. The observed intensification of the H-gain fragment ion is attributed to the effective transfer of this acidic proton to the electron of the double bond, facilitated by McLafferty rearrangement, as demonstrated in the acyl chain of arachidonic acid. Thus, the pronounced increase in the H-gain fragment ion abundance under KE 18 eV conditions serves as a diagnostic marker to differentiate C=C-positional isomers, such as n-3, n-6, and n-9 fatty acid chains. This increase in H-gain fragment ion is described as "C=C PUFA high" in this study. Notably, an increase in the abundance of H gain fragment ions was detected in the EAD-MS/MS spectrum at KE 14 eV. Therefore, the KE 14 eV setting provides (1) enhanced sensitivity of the

product ions, (2) a characteristic V-shaped pattern in lipids with C=C bonds, and (3) a distinct H-gain ion behavior from PUFAs with more than three C=C bonds.

Overview of MS-DIAL lipid annotation for EAD-MS/MS

This program employs a decision-tree-based method to annotate each lipid subclass. Given that a collision energy (CE) of approximately 10 V is commonly applied to ion transfer within mass spectrometers, including in EAD mode, the product ion spectrum from EAD represents a composite of CID-based fragmentation and charge-remote fragmentation (CRF) effects on lipids. Consequently, the program initiates the annotation process by searching for a diagnostic ion or neutral loss characteristic of the lipid subclass, which aids in determining the acyl chain attributes. The feasibility of identifying *sn*-positions and hydroxy (OH) groups varies with the lipid subclass and the type of adduct formed. For instance, the fragment ion from the homolytic cleavage at the C1-C2 bond of the glycerol backbone, indicative of the *sn*-position in glycerolipids (GLs) and glycerophospholipids (GPs), is readily detectable in the protonated form of PCs. However, the neutral loss fragment of the *sn1*-acyl chain from the protonated or ammonium adducts in other GLs and GPs may be less distinguishable from the acyl chain CRF- or noise ions. Because the cationic moiety in lipids tends to stabilize in the sodium adduct form, the EAD-MS/MS spectra for sodium adducts tend to be less convoluted than those for the $[M+H]^+$ and $[M+NH_4]^+$ forms. In this study, the annotation pipeline in MS-DIAL was assessed for $[M+H]^+$ and $[M+NH_4]^+$ adducts, which are predominant in conventional untargeted lipidomics. On the contrary, the annotation for $[M+Na]^+$ adduct is available in the current MS-DIAL 5 platform, with its validation for sodium adducts in combination with the optimization of sample preparation and analytical chemistry to be reported elsewhere. Furthermore, the elucidation of the OH position, which is vital for understanding sphingolipid metabolism, was enhanced by EAD in tandem with the MS-DIAL computational framework, given the distinct visibility of OH-position-related fragment ions.

The procedure for determining the position of the carbon-carbon double bonds (C=C) is as follows. Initially, the search is conducted for the presence of "C=C high" peaks in the product ion spectrum. Typically, for each C=C location, two such peaks are expected, with an additional "C=C PUFA high" peak observed for polyunsaturated fatty acid (PUFA) acyl chains containing more than three C=C bonds. If any of the expected "C=C high" peaks are absent from the product ion spectrum, the candidate molecule is eliminated from consideration. In cases where more than five "C=C high" peaks are predicted, the algorithm allows one such peak to be omitted. Subsequently, the correlation between the experimental and *in-silico* spectra of the lipid molecules was calculated, focusing on the CRF ions of the acyl chains. The *in-silico* spectrum generation involved computing *m/z* values for hydrogen loss (H-loss), radicals, and hydrogen gain (H-gain) fragment ions for each homolytic cleavage along the acyl chain. The intensity ratios for the H-loss, radical, and H-gain fragment ions were set to 0.5, 1, and 0.05, respectively, for saturated fatty acyl chains. These ratios are adjusted to 0.25, 0.5, and 0.05 for a "C=C low peak," and to 2.0, 4.0, and 0.05 to reflect an H-loss peak increase, and 4.0, 2.0, and 0.05 to indicate a radical peak increase in a "C=C high peak." A reverse dot product similarity score, utilizing the *in-silico* spectrum as the library template, was employed as a measure of correlation. The candidates were then ranked according to their reverse dot product scores, with the highest-scoring candidate designated as the representative C=C isomer in the EAD-MS/MS spectrum.

The structural diversity of the C=C positional isomers generated in MS-DIAL is inherently limited, and the configuration is seemingly optimized for mammalian cells. For monounsaturated fatty acids (MUFAs) with *O*-acyl and *N*-acyl chains, the potential C=C positions were derived from those listed in the LIPID MAPS Structure Drawing Tool for glycerophospholipid structures. Positions defined as multiples of three from the omega terminus were included as potential sites. The C=C positions of

sphingoid bases reference the candidate list from the LIPID MAPS tool for sphingolipids, typically including delta 4, 8, and 14, but deliberately excluding delta 6. In this study, only delta 4 position is considered when the sphingoid base contains one double bond. For polyunsaturated fatty acids (PUFAs), candidate structures featuring a methylene-interrupted C=C sequence starting from multiples of three from the omega terminus were generated. The current version of MS-DIAL is not geared towards the discovery of new C=C positions; rather, it is designed to identify the most plausible candidate structures from a range of known double bonds and hydroxyl positions recognized in lipid biology. Naturally, the program can be tailored for other species such as plants and microorganisms by modifying the range of double-bond positions defined in the XML format in the source code. Although it is recognized that the current capabilities of EAD-MS/MS techniques may not be sufficient for untargeted approaches, the development of structure annotation programs for interpreting EAD-MS/MS spectra remains crucial for advancing the standard-free structure elucidation of lipids. By integrating computational mass spectrometry techniques with targeted analyses, wherein lipid enrichment is followed by highly sensitive measurements, EAD-MS/MS can be leveraged to uncover new lipid structures.

Elucidation of EAD-MS/MS spectra for glycerophospholipids in MS-DIAL

MS-DIAL provides an in-depth annotation pipeline for a wide range of glycerophospholipids. These include phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylglycerol (PG), phosphatidylinositol (PI), phosphatidylserine (PS), bis(monoacylglycero)phosphate (BMP), lyso-type forms (LPC, LPE, LPG, LPI, and LPS), and plasmenyl/plasmanyl species (PC P-, PE P-, PC O-, and PE O-). While hemi-BMP (HBMP) and cardiolipin (CL) molecules were also examined, only molecular species-level annotations, such as CL 16:0_18:1_18:1_18:2, were feasible in EAD-MS/MS. This limitation was due to the poor sensitivity of diagnostic ions related to the C=C- and *sn*-positions, even with injections exceeding 1 pmol of the on-column volume on a conventional C18 analytical column. The annotation of the adduct forms $[M+H]^+$, $[M+NH_4]^+$, and $[M+Na]^+$ was supported by MS-DIAL, although we evaluated the annotation accuracy of the $[M+H]^+$ and $[M+NH_4]^+$ product ion spectra.

In the EAD-MS/MS spectra, the product ions of the polar head (PH; $X+H_2PO_4$), $PH+C_3H_4$, and $PH+C_2H_2O$, along with the neutral loss (NL) of the polar head (PH) group, are commonly observed in many glycerophospholipid subclasses. Here, “X” denotes the specific formula for each lipid subclass, for instance, $C_5H_{12}N^+$ for PC. Notably, the ion abundance of NL in the PC polar head group is typically lower than that in other phospholipids. Double-charged ions of the molecules were distinctly detected in PC and PE. The fragment ion from homolytic cleavage of the C1-C2 bond in the glycerol backbone, identified as "NL of sn1+CH2," is detectable across most lipid subclasses. However, practical diagnosis using this ion is only viable for protonated PCs or sodium adduct phospholipids. The determination of *sn*-positional isomers is feasible for PC-O by confirming the fragment ion from the homolytic cleavage of the C1-C2 bond in the glycerol backbone. The distinction between PE-O and PE-P was as clear as in CID-MS/MS. Differentiation between PC-O and PC-P relies on scoring the C=C isomer candidates. For lysophospholipids (LPLs), *sn1/sn2* isomer characterization is based on the neutral loss of CH_2OH, which is specific to *sn2*-LPL. Further details on the relationship between phospholipid structure and EAD-MS/MS spectra are available in **Supplementary Figure 2** and in the description of the lipidomics minimal reporting checklist[27] (**Supplementary Note 1**).

Elucidation of EAD-MS/MS spectra for sphingolipids in MS-DIAL

This program provided an in-depth annotation pipeline for five sphingolipids: sphingomyelin (SM), ceramide (Cer), hexosylceramide (HexCer), dihexosyl ceramide (Hex2Cer), and sulfatide (SHexCer). Other sphingolipids such as gangliosides and globosides have also been annotated at the molecular species

level. EAD-MS/MS offers two distinct advantages over CID-MS/MS for the annotation of sphingolipids. First, the product ion of the sphingobase (SPB) and the neutral loss (NL) of the *N*-acyl chain was clearly observed, serving as crucial diagnostic markers to define lipids at the molecular species level, such as SM 18:1;O2/16:0. Second, the fragment ion resulting from the cleavage of each carbon-carbon bond containing a hydroxy moiety was distinctly detected, aiding in the annotation of OH positions in the sphingobase backbone. Consequently, EAD-MS/MS enhanced the OH-resolved sphingolipid profiling. Annotation of the C=C-position followed the same methodology as that used for glycerophospholipids. Further details on the relationship between sphingolipid structure and EAD-MS/MS spectra are provided in **Supplementary Figure 3** and the lipidomics minimal reporting checklist (**Supplementary Note 1**).

Elucidation of EAD-MS/MS spectra for glycerolipids, sterols, and fatty acyls in MS-DIAL

The program offers the in-depth annotation pipeline for diacylglycerol (DG), triacylglycerol (TG), acylcarnitine (CAR), 1,2-diacylglyceryl-3-*O*-2'-(hydroxymethyl)-(*N*,*N*,*N*-trimethyl)-β-alanine (DGTA), 1,2-diacylglyceryl-3-*O*-4'-(*N*,*N*,*N*-trimethyl)-homoserine (DGTS), and their lyso forms (LDGTA and LDGTS). Monoacylglycerol (MG) was characterized at the molecular species level using EAD-MS/MS. The program also accommodates the detailed structural elucidation of fatty acyl esters of hydroxy fatty acids (FAHFA) and free fatty acids derivatized with 2-dimethylaminoethylamine (DMED), whose detailed methodology includes sample preparation and lipid enrichment will be reported elsewhere. For DG and TG, the *sn1/sn2* positional isomers were determined from the EAD-MS/MS spectra of the sodium adducts. In contrast, the characterization of C=C positional isomers involves scoring candidates based on the reverse dot product similarity value between the experimental and *in silico* spectra related to acyl chains, as described above. Spectral information on DGTS and DGTA was obtained from cultivating two algal species: *Chlamydomonas reinhardtii*, which predominantly produces DGTS, and *Fistulifera solaris*, which generates DGTA. Cells were cultured according to a previously established protocol[28,29]. The distinction between DGTS and DGTA isomers, which was not feasible in CID, was enabled in EAD by comparing the ion abundance ratios of *m/z* 204.123 ($C_9H_{18}NO_4$) and *m/z* 236.149 ($C_{10}H_{22}NO_5$). Notably, the ion abundance at *m/z* 236.149 surpassed that at *m/z* 204.123 in DGTS, whereas this ratio was inverted in DGTA. Details of the EAD-MS/MS spectra of glycerolipids are available in **Supplementary Figure 4**, as well as the description of the lipidomics minimal reporting checklist (**Supplementary Note 1**).

Evaluation of calibration curve using DLPC and PAPC

The authentic standards 1,2-dilinoleoyl-*sn*-glycero-3-phosphocholine (DLPC) and 1-palmitoyl-2-arachidonoyl-*sn*-glycero-3-phosphocholine (PAPC) were purchased from Avanti Polar Lipids. Each compound was dissolved using 1:1 MeOH:$CHCl_3$ (v/v). A series of dilutions were prepared at concentrations of 10, 5, 2, 1, 0.5, 0.2, 0.1, 0.05, 0.02, and 0.01 µM for each compound. Given that 1 µL of each sample was injected, the on-column volume for the LC-MS method utilized in this study is estimated at 10,000 fmol, 5,000, 2,000, 1,000, 500, 200, 100, 50, 20, and 10. The LC-MS conditions employed were mostly identical to those detailed in "*Investigating the kinetic energies to acquire information-rich EAD spectra of small molecules*". Nine fragmentation conditions were explored during targeted MS/MS scanning mode, including one CID setting (45 eV) and various EAD settings with different kinetic energies (8, 10, 12, 14, 16, 18, and 20 eV KE). A constant CE spread of 0, collision energy of 10 for EAD, and a time-of-flight (TOF) start mass of 170 were applied. For CID, a constant CE spread of 15 was applied. The ion source temperature was maintained at 250 °C, while other parameters were consistent with those described above. Each sample was analyzed three times (technical replicates = 3). The peak heights from the product ion chromatogram peak tops were used for quantification.

The peak heights of the diagnostic ions used to determine the *sn*-position and C=C-position of the lipids were investigated for DLPC and PAPC. For DLPC, the peak heights for product ions at *m/z* 184.073 (mandatory to define the PC lipid subclass), *m/z* 489.321 (NL of sn1+CH_2; diagnostic ion to define the *sn*-position), *m/z* 630.413 (H-loss at 18:2 C7 "C=C high peak"), *m/z* 670.444 (H-loss at 18:2 C10 "C=C high peak"), *m/z* 685.468 (radical at 18:2 C11 "C=C high peak"), and *m/z* 725.499 (radical at 18:2 C14 "C=C high peak") were investigated. For PAPC, the peak heights for product ions at *m/z* 184.073 (mandatory to define the PC lipid subclass), *m/z* 465.321 (NL of sn1-20:4+CH_2; diagnostic ion to define the *sn1* 20:4), *m/z* 513.321 (NL of sn1-16:0+CH_2; diagnostic ion to define the *sn1* 16:0), *m/z* 550.350 (H-loss at 20:4 C3 "C=C high peak"), *m/z* 590.382 (H-loss at 20:4 C6 "C=C high peak"), *m/z* 592.397 (H-gain at 20:4 C6 "C=C PUFA high peak"), *m/z* 605.405 (radical at 20:4 C7 "C=C high peak"), *m/z* 630.413 (H-loss at 20:4 C9 "C=C high peak"), *m/z* 632.429 (H-gain at 20:4 C9 "C=C PUFA high peak"), *m/z* 645.436 (radical at 20:4 C10 "C=C high peak"), *m/z* 670.444 (H-loss at 20:4 C12 "C=C high peak" and H-gain at 16:0 C8), *m/z* 685.468 (radical at 20:4 C13 "C=C high peak"), and *m/z* 725.499 (radical at 20:4 C16 "C=C high peak" and radical at 16:0 C12), were examined. In this study, the product ion's peak was recognized as "not detected" if the peak height was zero in two of three samples.

Evaluation of annotation results of co-eluted lipid molecules using a mixture of DLPC and PAPC

Two sets of the mixtures were prepared. The first set, termed "DLPC fixed," comprised seven mixtures containing DLPC and PAPC, where the DLPC concentration was consistently maintained at 1 µM, while PAPC concentrations varied at 0.1, 0.2, 0.5, 1.0, 2.0, 5.0, and 10 µM. The second set, termed "PAPC fixed," similarly consisted of seven mixtures. Here, the PAPC concentration was fixed at 1 µM, with DLPC concentrations adjusted to 0.1, 0.2, 0.5, 1.0, 2.0, 5.0, and 10. Each sample was analyzed three times (technical replicates = 3). The same mass spectrometer conditions described in the previous section were used. Flow injection, involving no column installation, was used to ensure the co-elution of the two metabolites. The other liquid chromatography conditions were consistent with those previously described (see the section of "*Evaluation of calibration curve using DLPC and PAPC*"). The spectra of the co-eluted lipids were elucidated using MS-DIAL.

Evaluation of MS-DIAL program by using LightSPLASH mixture

The LightSPLASH mixture (https://avantilipids.com/product/330732) containing 13 authentic lipid standards at 100 µg/mL each was purchased from Avanti Polar Lipids (**Supplementary Table 3**). This mixture was initially diluted fivefold with a 1:1 $CHCl_3$:MeOH (v/v) solvent. This diluted solution served as the starting point for a subsequent series of dilutions. The initial mixture was further diluted by factors of 2, 5, 10, 20, 50, 100, 200, 500, and 1000, using the same 1:1 $CHCl_3$:MeOH solvent.. Each sample was analyzed thrice using LC-EAD (KE 14)-MS/MS. The lipid separation was carried out with the column of Unison UK-C18 MF (50 × 2.0 mm, 3 µm, Imtakt Corp., Kyoto, Japan) and the mobile phases of (A) acetonitrile (ACN):MeOH:H_2O (1:1:3, v/v/v) and (B) ACN:IPA (1:9, v/v). Both the solvents contained 10 nM ethylenediaminetetraacetic acid and 5 mM ammonium acetate. The injection volume, flow rate, sample rack temperature, and column oven temperature were set to 1 µL, 300 µL/min, 4 °C, and 45 °C, respectively. The gradient condition is as follows: 0.1% (B) (1 min), 0.1–40% (B) (4 min), 40–64% (B) (2.5 min), 64–71% (B) (4.5 min), 71–82.5% (B) (0.5 min), 82.5–85% (B) (6.5 min), 85–99.9% (B) (0.1 min), 99.9% (B) (1.4 min), 99.9–0.1% (B) (0.1 min), 0.1% (B) (4.4 min). A data-dependent MS/MS acquisition mode, called information-dependent acquisition in SCIEX, was used. The conditions for the EAD are as follows: MS1 scan range, *m/z* 70-1250; MS/MS scan range, *m/z* 150-1250; MS1 accumulation time, 200 ms; MS2 accumulation time, 100 ms; electron beam current, 7000 nA; ETC%, 100%; TOF start

mass, 150; KE, 14 eV; CE, 10 V; CES, 0 V; ion source gas 1, 40; ion source gas 2, 80; curtain gas, 30; CAD gas, 7; temperature, 250; spray voltage, 5500; declustering potential, 80. Mass calibration was automatically performed using a SCIEX calibration delivery system. The mass spectra were analyzed using MS-DIAL version 5. The WIFF format files were directly imported into MS-DIAL. The following parameters were selected from the measurement setting page view: ionization mode, soft ionization, fragmentation method, EIEIO, target omics, lipidomics, MS1, centroid, MS2, and centroid. Details of the other parameter settings are listed in **Supplementary Table 4**. In this evaluation, the representative annotation was determined as follows. If the same lipid name was annotated in at least two of the three replicates, that name was used as the representative annotation. If the annotation results differed across all three replicates, the lipid with the highest score was adopted as representative.

Evaluation of MS-DIAL program by using a mixture of UltimateSPLASH and in-house lipid standards

A standard mixture was prepared to evaluate the performance of the MS-DIAL algorithm for the EAD spectral annotation (**Supplementary Table 5**). The Ultimate SPLASH, containing 69 lipid molecules, was purchased from Avanti Polar Lipids. The concentrations of the compounds in the original solutions varied from 26.87 µM 192.5 µM. This solution was diluted by factors of 2, 5, 10, 20, 50, 100, and 200 in 1:1 MeOH:$CHCl_3$ (v/v). Furthermore, an in-house mixture containing 41 lipid standards, previously employed in a different study[30], was also utilized, where each lipid was adjusted to a concentration of 50 µM. This solution was subjected to the same dilution process as that used for Ultimate SPLASH.

A leaf lipid extract from uniformly ^{13}C-labeled (>97 atom % ^{13}C) *Nicotiana tabacum* was used as the background matrix. The plant materials were purchased from IsoLife (Wageningen, Netherlands). The lipid extraction protocol was performed according to a previous study[25]. Briefly, the plant material in a 2.0 mL microcentrifuge tube was milled by shaking at 900 rpm for 3 min on a Shake Master Neo (BMS, Tokyo, Japan) using zirconia beads. From the frozen powdered plant material, 5 mg was measured and transferred into a new 2.0 mL tube. To the tube, 1 mL of a solvent mixture consisting of 5:2:2 MeOH:H_2O:$CHCl_3$ (v/v/v) was added. After stirring on a vortex mixer vigorously, the homogenate was incubated for 30 min at 1200 rpm at 25 °C, followed by the addition of 400 µL of H_2O for liquid-liquid separation. Twenty microliters of the bottom solvent layer was transferred into a new 2.0-mL tube, where a total of eight tubes were prepared. Fifty microliters of each dilution ratio from the dilution series of Ultimate SPLASH and in-house standard mixture solutions were added to each tube. A total of 100 µL of the solvent used to create the dilution series, namely a 1:1 MeOH:$CHCl_3$ (v/v) solvent, was added to the remaining one tube. The samples were dried with a vacuum dryer and resuspended in 50 µL of MeOH, including 1 µL of EquiSPLASH mixture. The LC-MS/MS settings, MS-DIAL settings, and evaluation methods were the same as those described in the previous section.

Characterization of very long chain PUFA (VLC-PUFA) containing PC in the eye tissue of mice

The animal experiments were performed in accordance with the ethical protocol approved by the Tokyo University of Agriculture and Technology (R5-50). Nine-week-old C57BL/6J male mice were purchased from SLC (Shizuoka, Japan). The mice were fed the chow of CE-2 (CLEA Japan, Tokyo, Japan) for 2 weeks. The eye organ was harvested and immediately frozen after dissection and stored at −80 °C until lipid extraction. Samples were lyophilized prior to lipid extraction. For the extraction, an incision was made in the mouse eye with scissors, and a single 5-mm diameter zirconia bead was inserted. All procedures were performed on ice. The samples were homogenized using a mixer mill (MM 301; Retsch, Germany) at 20 Hz for 2 min. The mixer mill rack was pre-chilled with liquid nitrogen prior to homogenization. A 2–5-mg sample of the mouse eye was subjected to the Bligh and Dyer method. Samples were mixed with 1,000 µL of an ice-cold MeOH/$CHCl_3$/H_2O (10:4:4, v/v/v) solvent. Lipids were

extracted using a vortex mixer for 1 min and then ultrasonicated for 5 min. The solution was centrifuged at 16,000 ×*g* for 5 min at 4 °C, and 700 µL of the supernatant was transferred to a clean tube. The supernatant was mixed with 235 µL of $CHCl_3$ and 155 µL of H_2O using a vortex mixer for 1 min. After a subsequent centrifugation at 16,000 ×*g* for 5 min at 4 °C, 330 µL of the organic (bottom) layer was collected. Finally, lipid extracts were dried using a centrifugal evaporator. Lipid extracts were dried using a centrifuge evaporator.

The samples were analyzed with and without enrichment for VLC-PUFA PC molecules. The enrichment process was performed as follows: A dried lipid extract was dissolved by applying 60 µL of MeOH with 1% formic acid. A MonoSpin Phospholipid (GL Sciences Inc., Tokyo, Japan) solid-phase extraction (SPE) column was activated with 200 µL of MeOH with 1% formic acid, followed by the application of 50 µL of the sample. Centrifugation during SPE was performed at 3,000 ×*g* for 1 min. After washing the SPE column with 200 µL of 100% MeOH, the phospholipids were fractionated twice using 200 µL of 9:10:1 IPA:H_2O:NH_3 (v/v/v) solvent. After solvent evaporation using a vacuum dryer, the residue was resuspended in 50 µL of 95% MeOH and further fractionated using a MonoSpin C18 SPE column (GL Sciences Inc., Tokyo, Japan). The column was conditioned with 200 µL of 100% MeOH followed by 200 µL of H_2O. Subsequently, 50 µL of the sample was applied. The column was then washed with 200 µL of H_2O, 200 µL of 95% MeOH, and twice with 200 µL of 96.5% MeOH. Finally, the VLC-PUFA PC fraction was eluted using 200 µL of 100% MeOH and 200 µL of 1:1 MeOH:$CHCl_3$, with the solvent subsequently evaporated using a vacuum dryer. These processes were omitted from analyses without lipid enrichment. After the sample was dissolved in 50 µL of 100% MeOH containing 1 µL of EquiSPLASH and 1 µM FA 16:0-d3 and FA 18:0-d3, it was transferred to an LC-MS vial. Four biological replicates were analyzed. The same LC-MS/MS conditions used for the evaluation of the MS-DIAL program were used, and 1 µL from each vial was injected. Samples without lipid enrichment were analyzed using the DDA method of EAD 14 eV KE. The same MS settings as those described for the LightSPLASH and UltimateSPLASH analyses were used for the EAD, while the LC gradient condition was slightly different. The gradient condition is as follows: 0.5% (B) (1 min), 0.5–40% (B) (4 min), 40–64% (B) (2.5 min), 64–71% (B) (4.5 min), 71–82.5% (B) (0.5 min), 82.5–85% (B) (6.5 min), 85–99% (B) (1.0 min), 99.9% (B) (2.0 min), 99.9–0.1% (B) (0.1 min), 0.1% (B) (4.9 min). This experiment is also described in the lipidomics minimal reporting checklist (**Supplementary Note 2**).

Database creation of species/tissue-specific m/z and collision-cross section values of lipids

The database for the lipid annotation of matrix-assisted laser desorption/ionization (MALDI) coupled with tapped ion mobility mass spectrometry (TIMS) data was prepared as follows: Conventional reverse-phase LC-MS/MS-based lipidomics data from MetaboBank ID MTBKS215, MTBKS216, and MTBKS217 were downloaded from the website (https://www.ddbj.nig.ac.jp/metabobank/index.html). The dataset contained 136 unique biological origins, including 27 unique tissues or cell types from mice (C57B6J or C57B6N), two cell types and human plasma, 99 unique pairs of plant species/tissues, and 7 algae species. The *m/z* and collision cross-section (CCS) database of lipids that were characterized in a specific biospecimen was created for each of the 136 unique biological origins. The full list of the characterized lipid molecules in each biological study is available in **Supplementary Table 6**, where an average of 236 lipid molecules per sample were recorded. The CCS values of $[M+H]^+$, $[M+NH_4]^+$, $[M+Na]^+$, $[M-H]^-$, $[M+HCOO]^-$, $[M+CH_3COO]^-$, $[M+H-H_2O]^+$, $[M+K]^+$, and $[M+Li]^+$ were predicted by the machine learning model created in the previous study with a small modification. Briefly, the experimental CCS values of 3601 ion forms of 2799 molecules from 95 lipid subclasses were used for the model development, where the training data set is available in the supplementary data of a previous report[10]. The descriptors and fingerprints of the molecular structure were calculated by NCDK v1.5.6

(https://kazuyaujihara.github.io/NCDK/html/e2ff06cc-99b7-4f8b-95c5-53965548639f.htm). With these variables, the XGBoost function optimized by the parameter tuning method of PicNet.XGBoost (v0.2.1; https://www.nuget.org/packages/PicNet.XGBoost/), was used to create the CCS prediction model.

Data processing of mass spectrometry imaging data

Mass spectrometry imaging data of the eye tissues from C57B6/J and acyl-coenzyme A (CoA) synthetase (ACSL) 6 knockout mice were downloaded from the RIKEN DROPMet website (http://prime.psc.riken.jp/menta.cgi/prime/drop_index), identified under index number DM0048. This dataset was obtained using the "TIMS-ON" mode, indicating that ion mobility separation was executed. In this study, only positive ion mode data generated using a matrix of DHB (2,5-dihydroxybenzoic acid) were analyzed. The detailed methodologies are available in a previous paper[17]. Timsdata.dll is necessary for reading the Bruker raw data files were downloaded from the Bruker SDK website (https://www.bruker.com/en/services/software-downloads.html). The data structure encompasses approximately 400 spectra for each MALDI spot, expandable along the drift time and *m/z* axis. The initial step in the algorithm involves accumulating all spectra from all MALDI spots, applying binning values of *m/z* 0.005 and drift time 0.01 ms, adjustable by users. The accumulated spectral data were stored in an intermediate file. The peak picking algorithm of MS-DIAL was performed for the accumulated spectra, resulting in the generation of peak features defined by *m/z*, drift time, collision cross section (CCS), and peak height. Lipid annotation was performed using the above database, which contained the *m/z* and CCS reference values of lipids detected in the eye tissues of mice. The tolerances for *m/z* and CCS were set to 0.01 Da and 20 $Å^2$, respectively. Peak features were utilized to map the ion distributions in the spatial images. For the ion abundance mapping into each of the MALDI spot pixels, the ions with tolerances of 0.01 Da and 5 $Å^2$ from the experimental values of peak features were accumulated, and the integrated data is stored as an intermediate file in the MS-DIAL application.

Re-analysis of publicly available untargeted lipidomics data analyzing the eye tissue of Acsl6 KO mouse

The untargeted lipidomics data using a reverse-phased chromatography method was downloaded from the RIKEN DROP Met website (http://prime.psc.riken.jp/menta.cgi/prime/drop_index), identified under index number DM0048, which is the same as above. The negative-ion mode data were analyzed using MS-DIAL 5, and the parameters used are listed in **Supplementary Table 7**.

HeLa cells experiment with the addition of very long-chain polyunsaturated fatty acid (VLC-PUFA)

The omega-3 VLC-PUFA compound 14*Z*,17*Z*,20*Z*,23*Z*,26*Z*,29*Z*-dotriacontahexaenoic acid (FA 32:6; ID: CAY10632) was purchased from Cayman Chemicals. HeLa cells (ATCC) were maintained at 37 °C in DMEM (Dulbecco's modified Eagle medium) supplemented with 10% EquaFETAL (Atlas Biologicals, Inc.) and 1% penicillin-streptomycin solution (Fujifilm, Wako, Japan) with 5% CO_2. For sample preparation for glycerolipid and glycerophospholipid profiling, cells (3×10^5 cells per well) were incubated in 6-well plates (Thermo Scientific, Nunc, Denmark) for 3 h. For LPA analysis, cells (5×10^5 cells/well) were incubated in 10 cm dishes (TPP, Switzerland). VLC-PUFA (FA 32:6) dissolved in 0.4% ethanol was added to the well plate at final concentrations of 1, 10, and 40 μM, with four biological replicates per condition. A solution of 0.4% ethanol was used as the vehicle control. After incubation for 24 h, the medium was removed, and the cells were washed twice with ice-cold phosphate-buffered saline (PBS) without calcium and magnesium. To profile the glycerolipids and glycerophospholipids, the cells were detached using a cell scraper with 1000 μL of ice-cold MeOH. The solvent (700 μL) was then transferred to a tube. In addition, 400 μL of ice-cold MeOH was added to each well plate, and 300 μL of the solvent was transferred to the same tube. To profile LPA, ice-cold PBS was used instead of ice-cold MeOH. The

PBS solution was discarded after centrifugation at 16,000 ×*g* for 5 min at 4 °C, and the cell pellet was stored at -80 °C until LPA analysis (see next section).

The solvent was sonicated in BIORUPTOR II (CosmoBio, Tokyo, Japan) for 10 cycles, each taking 0.5 min for sonication and 0.5 min to maintain the water temperature at 4 °C. After adding 400 μL of $CHCl_3$, lipids were extracted using a vortex mixer for 1 min and ultrasonication for 5 min. The solution was centrifuged at 16,000 ×*g* for 5 min at 4 °C, and 700 μL of the supernatant was transferred to a clean tube. After adding 300 μL of $CHCl_3$ and 400 μL of H_2O to the tube, the solution was vortexed for 1 min and ultrasonicated for 5 min. After a subsequent centrifugation at 16,000 ×*g* for 5 min at 4 °C, 400 μL of the organic (bottom) layer was collected. Lipid extracts were dried using a centrifuge evaporator. After the sample was dissolved in 60 μL of 100% MeOH containing 1 μL of EquiSPLASH and 1 μM FA 16:0-d3 and FA 18:0-d3, it was transferred to an LC-MS vial. The lipids were analyzed using ESI(+)- and ESI(-)-CID DDA modes. The mass spectrometer settings for the CID mode were as follows: MS1 and MS2 mass ranges, *m/z* 70–1250; MS1 accumulation time, 200 ms; Q1 resolution, units; MS2 accumulation time, 50 ms; maximum candidate ions, 10; CAD gas, 7; intensity threshold for DDA, 10 cps; dynamic background subtraction, ticked; and no inclusion or exclusion lists were used. The following settings were used for positive/negative ion mode, independently: ion source gas 1, 40/50 psi; ion source gas 2, 80/50 psi; curtain gas, 30/35 psi; source temperature, 250/300 °C; spray voltage, 5500/-4500 V; declustering potential, 80/-80 V; and collision energy, 40/-42 ± 15 eV. The LC condition is the same as used for eye-lipidome analysis. This experiment is also described in the lipidomics minimal reporting checklist (**Supplementary Note 3**).

LC-MS/MS analysis for lysophosphatidic acid (LPA) profiling for HeLa cells

Methyl tert-butyl ether (MTBE) and trimethylsilyl (TMS)-diazomethane were purchased from Sigma-Aldrich (Tokyo, Japan) and Tokyo Chemical Industry (Tokyo, Japan), respectively. LPA analysis was performed using a modified protocol from a previous study[31]. Next, 200 mL ice-cold MeOH with 0.1% formic acid was added to the HeLa cell pellet. The solvent was sonicated in BIORUPTOR II (CosmoBio, Tokyo, Japan) for 10 cycles, each taking 0.5 min for sonication and 0.5 min to maintain the water temperature at 4 °C. After centrifugation at 16,000 ×*g* for 5 min at 4 °C, 190 μL of the supernatant was transferred into a new tube and dried by a centrifugal evaporator. The sample was dissolved with 120 μL of MeOH containing 0.5 μM of LPA 17:1 as the internal standard. For the derivatization, 50 μL of 2M TMS-diazomethane was added and incubated for 20 min at 25 °C by 800 rpm in Ballerina NSD-12J (Tokyo Garasu Kikai Co., Ltd., Japan). After adding 3 μL of acetic acid, 400 μL of MTBE and 100 μL of H_2O were added. After vortex mixing at the maximum speed for 5 min at 25 °C in Ballerina NSD-12J, 360 μL of the supernatant was collected and dried up by centrifugal evaporator. The sample was resuspended in 30 μL MeOH containing 1 μL EquiSPLASH and 1 μM FA 16:0-d3 and FA 18:0-d3. LPA analysis was performed using the MRMHR mode, targeting the protonated form of bismethyl LPA (BisMeLPA) 32:6 under the mostly same LC-MS conditions described in the previous section. The mass spectrometer settings were as follows: MS1 mass range, *m/z* 100–1000; MS1 accumulation time, 250 ms; Q1 resolution, units; MS2 accumulation time, 100 ms. The other settings are the same as described above. This experiment is also described in the lipidomics minimal reporting checklist (**Supplementary Note 4**).

GPAT1 protein reconstitution using a wheat germ cell-free synthesis system

GPAT recombinant proteins were prepared using a cell-free system according to a previously reported protocol[19]. Complementary deoxyribonucleic acid (cDNA) encoding human GPAT1, also known as GPAM (glycerol-3-phosphate acyltransferase, mitochondrial), was cloned into the pEU vector (CellFree

Science, Japan). GPAT1 was synthesized by Integrated DNA Technologies (Coralville, IA, US). Hereafter, the native gene is shown as $GPAT1^{WT}$. The gene arrays and sequence details are listed in **Supplementary Table 8**. The cDNA encoding the mutant $GPAM^{H230A}$ was generated by site-directed mutagenesis PCR, according to the manufacturer's protocol (TaKaRa, PrimeSTAR Mutagenesis Basal Kit, Japan). The primers used for cloning in this study are listed in **Supplemental Table 8**. The native pEU vector was used for the vector control, termed "Mock". Protein reconstitution with liposomes was performed using the WEPRO7240 Expression Kit and Asolectin Liposomes (Cell Free Sciences, Japan) according to a previous study[19]. The presence of the synthesized proteoliposomes was verified by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie brilliant blue (CBB) staining (**Supplementary Figure 8**). The amount of the expressed proteins was determined using a bovine serum albumin (BSA) standard, and 1 μg of expressed protein-containing liposomes was used in the enzymatic assays. The average liquid volume from enzymatic assays involving $GPAT1^{WT}$ and $GPAT1^{H230A}$ was utilized for the mock sample's enzymatic analysis.

Evaluation of GPAT1 enzyme activity

GPAT1 enzymatic activity was determined as in the previous study[32]. The assay was performed for 1 h at 37 °C in 100 μL solution containing 1 μg of protein-containing proteoliposomes, 75 mM Tris-HCl (pH 7.5), 4 mM $MgCl_2$, 1 mg/mL BSA (essentially fatty acid-free), 500 μM CoA, 2.5 mM ATP, 8 mM NaF, 800 μM glycerol 3-phosphate, and 10 μM of a fatty acid. In this study, three fatty acids, palmitic acid (uniformly ^{13}C-labeled, U-^{13}C), DHA, and FA n-3-32:6, were examined. A solution of 1% ethanol was used as the vehicle control. The reaction mixture was incubated for 10 min at 37 °C. The SPE method using MonoSpin C18 SPE column (GL Sciences Inc., Tokyo, Japan) was used for lipid extraction. First, 100 μL of 1 M ammonium acetate and 200 μL of MeOH containing 17:1 CoA (0.5 μM) and LPA 17:1 (0.5 mM) as internal standards were added to the reaction mixture. After the SPE column was activated with 200 μL MeOH, 200 μL H_2O, and 200 μL of 1 M ammonium acetate, 360 μL of the sample was applied. After the column was washed with 200 μL H_2O and 200 μL hexane, the targeted lipid fractions containing LPAs and acyl CoAs were retrieved using 200 μL MeOH. The solvent was dried and used for the LPA and acyl-CoA analyses.

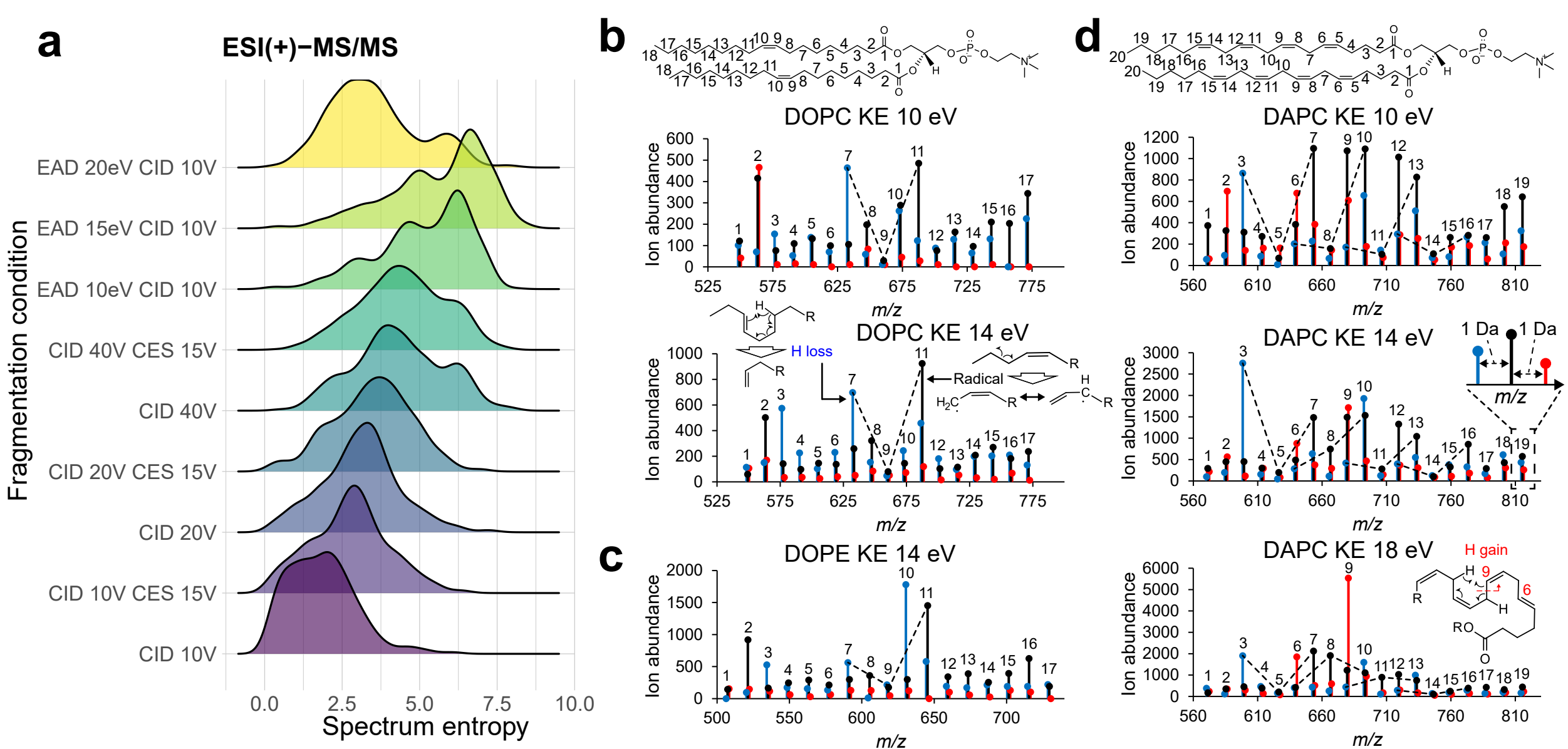
LC-MS/MS analysis for LPA and acyl CoA profiling for the extract from GPAM assay

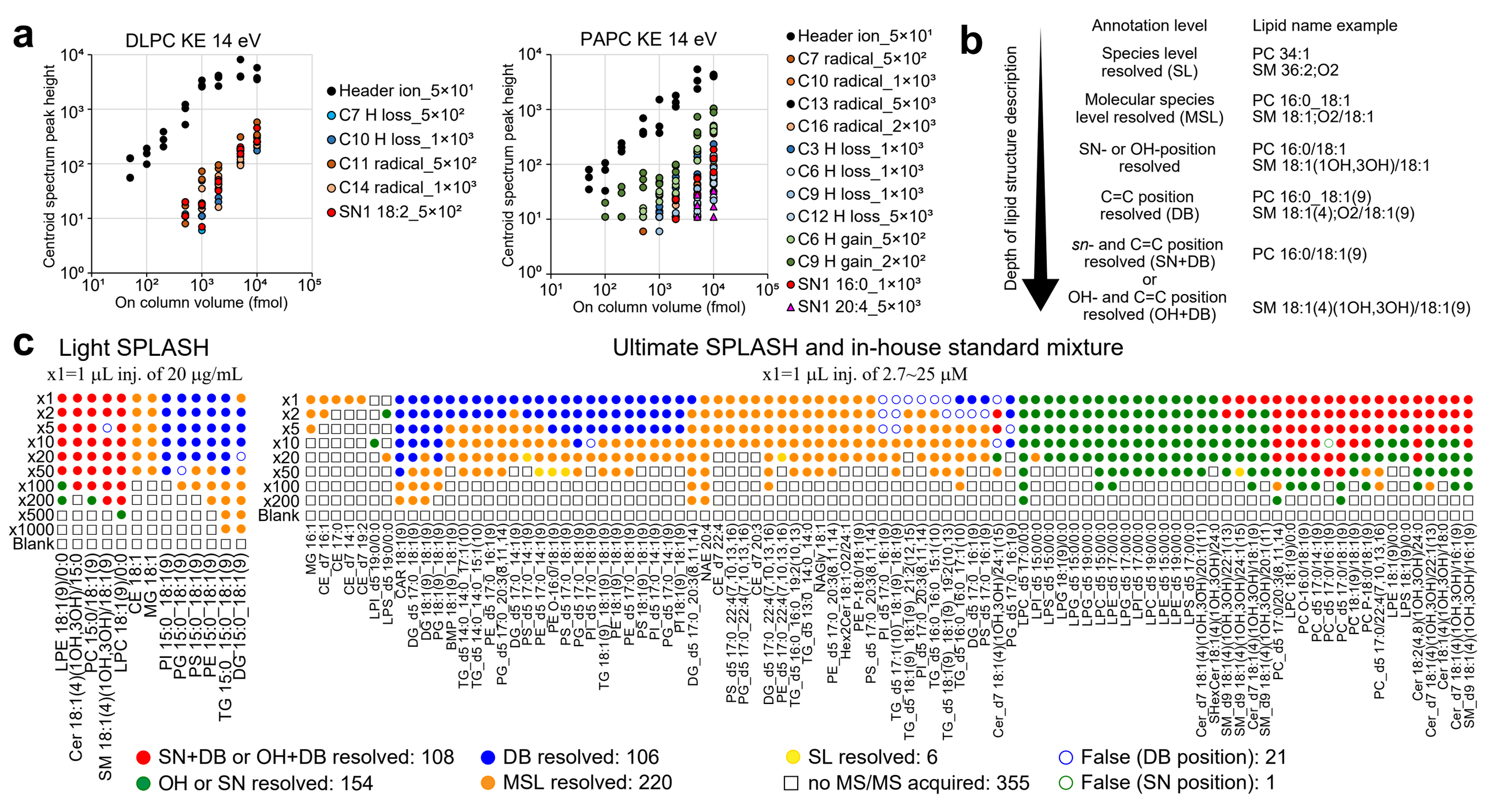
The dried sample was dissolved in 90 μL MeOH, 30 μL of which was transferred to an LC-MS vial for acyl-CoA profiling. The derivatization and LC-MS/MS protocol described above for LPA analysis were performed using the remaining solvent. The detail of LPA analysis is also described in the lipidomics minimal reporting checklist (**Supplementary Note 5**). The acyl CoA separation was carried out with the column of L-column3 C8 (3 μm, 2.0 × 100 mm metal free, CERI, Japan), and the mobile phases of (A) MeOH:H_2O (1:4, v/v) with 0.05% NH_3 and (B) MeOH:ACN (1:4, v/v) with 0.05% NH_3. The injection volume, flow rate, sample rack temperature, and column oven temperature were set to 5 μL, 250 μL/min, 4 °C, and 40 °C, respectively. The gradient conditions were 0.1% (B) (1.2 min), 0.1–100% (B) (4.8 min), 100% (B) (4 min), 100–0.1% (B) (0.1 min), and 0.1% (B) (4.9 min). Agilent 1290 Bio UHPLC coupled with 6546 QTOF system was used for the LC-MS analysis. The MS settings were as follows: gas temperature, 325 °C; gas flow, 12 L/min; nebulizer (psig), 55; sheath gas temperature, 300 °C; sheath gas flow, 11 L/min; vcap, 3500 V; nozzle voltage, 1000V; fragmentor, 175 V; skimmer 65V; octupole RF Vpp, 750 V; MS1 and MS2 ranges, *m/z* 90-1250; isolation width, narrow (~1.3 *m/z*); and collision energy, 20 eV. The MS/MS spectra were acquired by targeted MS/MS scanning mode. The detail of acyl-CoA analysis is also described in the lipidomics minimal reporting checklist (**Supplementary Note 6**)

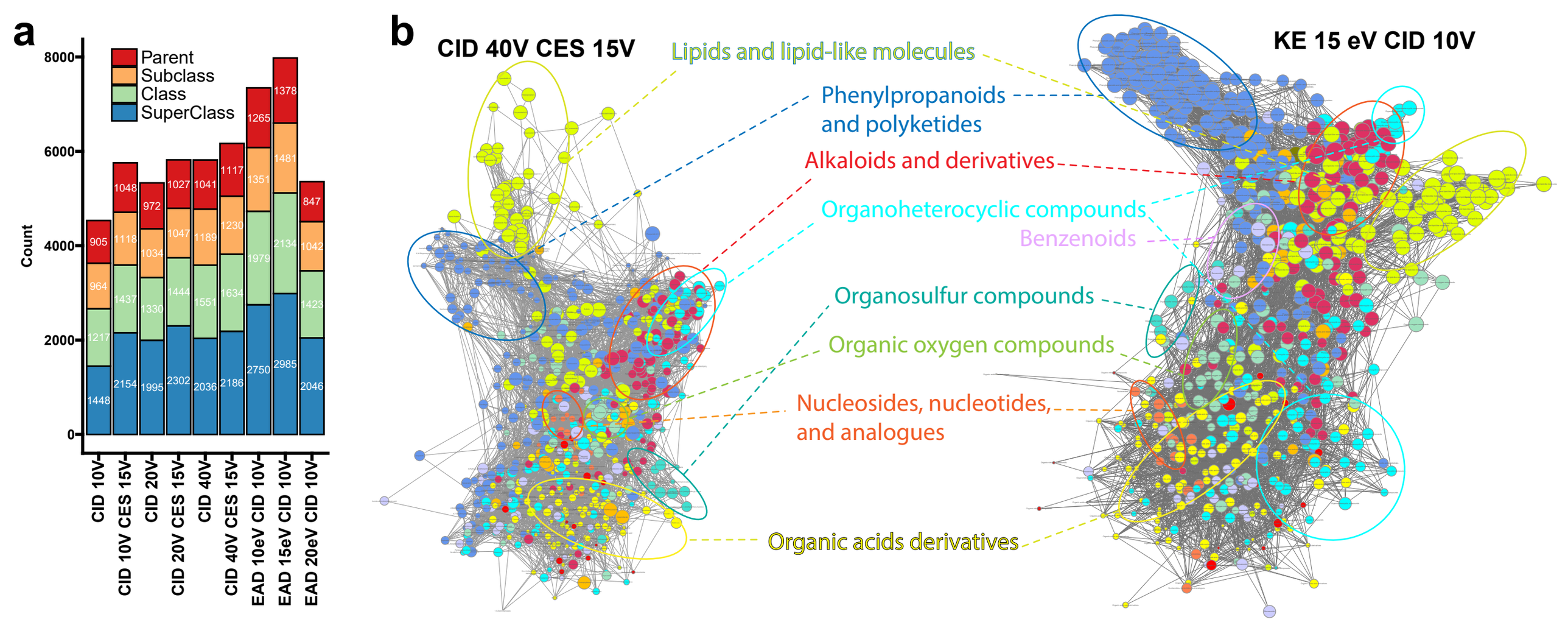
References

1. Rakusanova, S., Fiehn, O. & Cajka, T. Toward building mass spectrometry-based metabolomics and lipidomics atlases for biological and clinical research. *Trac-Trend Anal Chem* **158**(2023).
2. Liebisch, G., *et al.* Update on LIPID MAPS classification, nomenclature, and shorthand notation for MS-derived lipid structures. *J Lipid Res* **61**, 1539-1555 (2020).
3. Baba, T., Campbell, J.L., Le Blanc, J.C.Y., Baker, P.R.S. & Ikeda, K. Quantitative structural multiclass lipidomics using differential mobility: electron impact excitation of ions from organics (EIEIO) mass spectrometry. *J Lipid Res* **59**, 910-919 (2018).
4. Shi, H.X., *et al.* Visible-Light Paterno-Buchi Reaction for Lipidomic Profiling at Detailed Structure Levels. *Anal Chem* **95**, 5117-5125 (2023).
5. Brodbelt, J.S., Morrison, L.J. & Santos, I. Ultraviolet Photodissociation Mass Spectrometry for Analysis of Biological Molecules. *Chem Rev* **120**, 3328-3380 (2020).
6. Menzel, J.P., *et al.* Ozone-enabled fatty acid discovery reveals unexpected diversity in the human lipidome. *Nat Commun* **14**(2023).
7. Uchino, H., Tsugawa, H., Takahashi, H. & Arita, M. Computational mass spectrometry accelerates C = C position-resolved untargeted lipidomics using oxygen attachment dissociation. *Commun Chem* **5**(2022).
8. Djambazova, K.V., *et al.* Resolving the Complexity of Spatial Lipidomics Using MALDI TIMS Imaging Mass Spectrometry. *Anal Chem* **92**, 13290-13297 (2020).
9. Ni, Z.X., *et al.* Guiding the choice of informatics software and tools for lipidomics research applications. *Nat Methods* **20**, 193-204 (2023).
10. Tsugawa, H., *et al.* A lipidome atlas in MS-DIAL 4. *Nat Biotechnol* **38**, 1159-+ (2020).
11. Li, Y., *et al.* Spectral entropy outperforms MS/MS dot product similarity for small-molecule compound identification. *Nat Methods* **18**, 1524-1531 (2021).
12. Watrous, J., *et al.* Mass spectral molecular networking of living microbial colonies. *P Natl Acad Sci USA* **109**, E1743-E1752 (2012).
13. Demarque, D.P., Crotti, A.E.M., Vessecchi, R., Lopes, J.L.C. & Lopes, N.P. Fragmentation reactions using electrospray ionization mass spectrometry: an important tool for the structural elucidation and characterization of synthetic and natural products. *Nat Prod Rep* **33**, 432-455 (2016).
14. Gross, M.L. Charge-Remote Fragmentations - Method, Mechanism and Applications. *Int J Mass Spectrom* **118**, 137-165 (1992).
15. Yeboah, G.K., Lobanova, E.S., Brush, R.S. & Agbaga, M.P. Very long chain fatty acid-containing lipids: a decade of novel insights from the study of ELOVL4. *Journal of Lipid Research* **62**(2021).
16. Aveldano, M.I. Phospholipid species containing long and very long polyenoic fatty acids remain with rhodopsin after hexane extraction of photoreceptor membranes. *Biochemistry* **27**, 1229-1239 (1988).
17. Kuroha, S., *et al.* Long chain acyl-CoA synthetase 6 facilitates the local distribution of di-docosahexaenoic acid- and ultra-long-chain-PUFA-containing phospholipids in the retina to support normal visual function in mice. *Faseb J* **37**(2023).
18. Yu, J., *et al.* Update on glycerol-3-phosphate acyltransferases: the roles in the development of insulin resistance. *Nutr Diabetes* **8**, 34 (2018).
19. Takeda, H., *et al.* Production of monoclonal antibodies against GPCR using cell-free synthesized GPCR antigen and biotinylated liposome-based interaction assay. *Sci Rep-Uk* **5**(2015).
20. Tsugawa, H., *et al.* MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat Methods* **12**, 523-+ (2015).

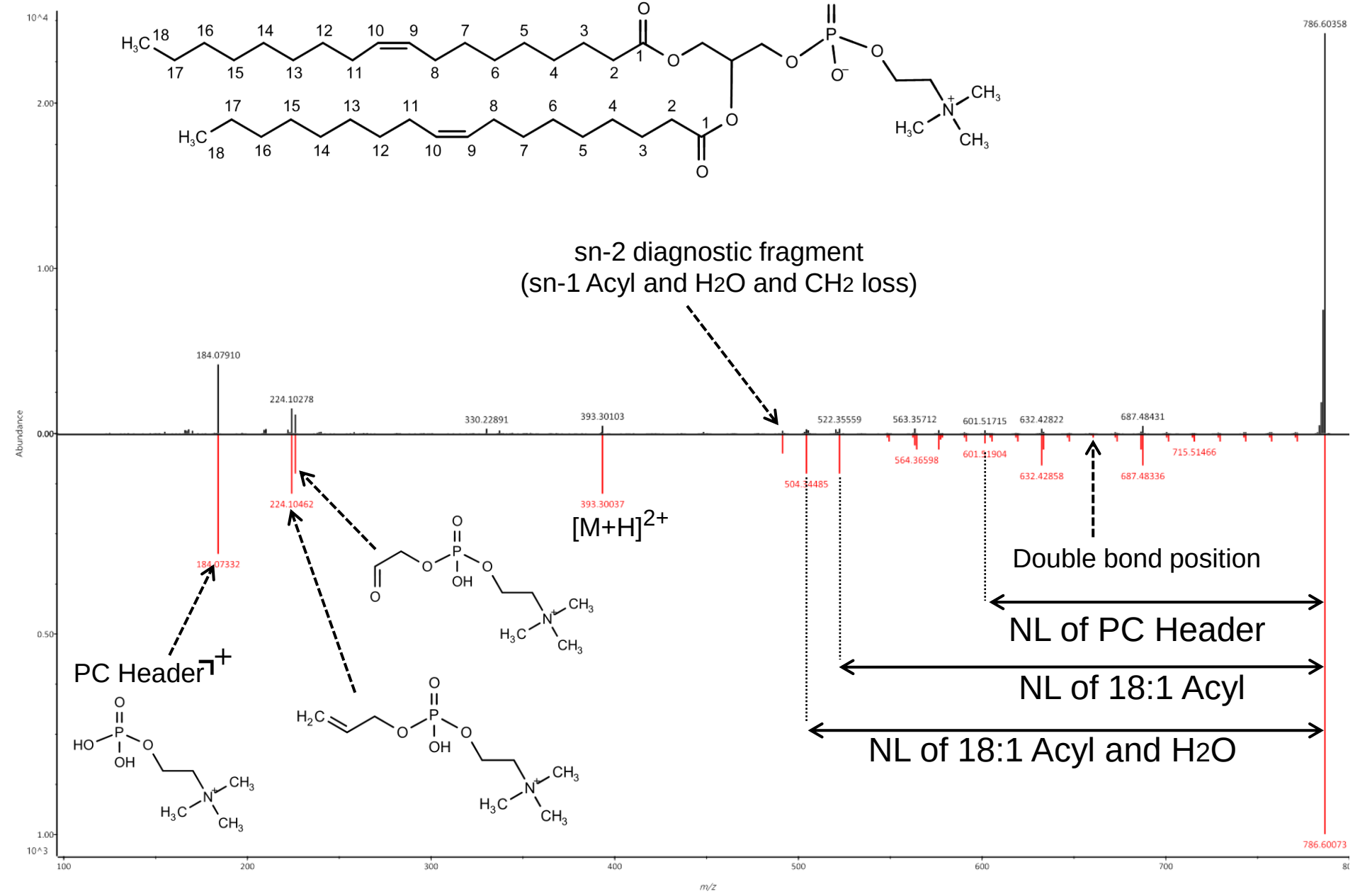
21. Kessner, D., Chambers, M., Burke, R., Agusand, D. & Mallick, P. ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics* **24**, 2534-2536 (2008).
22. Meier, F., *et al.* diaPASEF: parallel accumulation-serial fragmentation combined with data-independent acquisition. *Nat Methods* **17**, 1229-+ (2020).
23. Elapavalore, A., *et al.* Adding open spectral data to MassBank and PubChem using open source tools to support non-targeted exposomics of mixtures. *Environ Sci-Proc Imp* **25**, 1788-1801 (2023).
24. Wang, F., *et al.* CFM-ID 4.0: More Accurate ESI-MS/MS Spectral Prediction and Compound Identification. *Anal Chem* **93**, 11692-11700 (2021).
25. Tsugawa, H., *et al.* A cheminformatics approach to characterize metabolomes in stable-isotope-labeled organisms (vol 16, pg 295, 2019). *Nat Methods* **16**, 446-446 (2019).
26. Feunang, Y.D., *et al.* ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. *J Cheminformatics* **8**(2016).
27. McDonald, J.G., *et al.* Introducing the Lipidomics Minimal Reporting Checklist. *Nat Metab* **4**, 1086-1088 (2022).
28. Morowvat, M.H., Rasoul-Amini, S. & Ghasemi, Y. as a "new" organism for biodiesel production. *Bioresource Technol* **101**, 2059-2062 (2010).
29. Tanaka, T., *et al.* Oil Accumulation by the Oleaginous Diatom as Revealed by the Genome and Transcriptome. *Plant Cell* **27**, 162-176 (2015).
30. Tsugawa, H., *et al.* A lipidome landscape of aging in mice. (Research Square, 2022).
31. Lee, J.W., Nishiumi, S., Yoshida, M., Fukusaki, E. & Bamba, T. Simultaneous profiling of polar lipids by supercritical fluid chromatography/tandem mass spectrometry with methylation. *J Chromatogr A* **1279**, 98-107 (2013).
32. Wendel, A.A., Cooper, D.E., Ilkayeva, O.R., Muoio, D.M. & Coleman, R.A. Glycerol-3-phosphate acyltransferase (GPAT)-1, but not GPAT4, incorporates newly synthesized fatty acids into triacylglycerol and diminishes fatty acid oxidation. *J Biol Chem* **288**, 27299-27306 (2013).


a
ESI(+)−MS/MS
Fragmentation condition
EAD 20eV CID 10V
EAD 15eV CID 10V
EAD 10eV CID 10V
CID 40V CES 15V
CID 40V
CID 20V CES 15V
CID 20V
CID 10V CES 15V
CID 10V
0.0
2.5
5.0
7.5
10.0
Spectrum entropy
b
DOPC KE 10 eV
DOPC KE 14 eV
H loss
Radical
Ion abundance
m/z
c
DOPE KE 14 eV
d
DAPC KE 10 eV
DAPC KE 14 eV
1 Da
DAPC KE 18 eV
H gain


a
DLPC KE 14 eV
Centroid spectrum peak height
On column volume (fmol)
Header ion_5×10¹
C7 H loss_5×10²
C10 H loss_1×10³
C11 radical_5×10²
C14 radical_1×10³
SN1 18:2_5×10²
PAPC KE 14 eV
Header ion_5×10¹
C7 radical_5×10²
C10 radical_1×10³
C13 radical_5×10³
C16 radical_2×10³
C3 H loss_1×10³
C6 H loss_1×10³
C9 H loss_1×10³
C12 H loss_5×10³
C6 H gain_5×10²
C9 H gain_2×10²
SN1 16:0_1×10³
SN1 20:4_5×10³
b
Depth of lipid structure description
Annotation level
Lipid name example
Species level resolved (SL)
PC 34:1
SM 36:2;O2
Molecular species level resolved (MSL)
PC 16:0_18:1
SM 18:1;O2/18:1
SN- or OH-position resolved
PC 16:0/18:1
SM 18:1(1OH,3OH)/18:1
C=C position resolved (DB)
PC 16:0_18:1(9)
SM 18:1(4);O2/18:1(9)
sn- and C=C position resolved (SN+DB)
or
OH- and C=C position resolved (OH+DB)
PC 16:0/18:1(9)
SM 18:1(4)(1OH,3OH)/18:1(9)
c
Light SPLASH
x1=1 μL inj. of 20 μg/mL
Ultimate SPLASH and in-house standard mixture
x1=1 μL inj. of 2.7~25 μM
SN+DB or OH+DB resolved: 108
DB resolved: 106
SL resolved: 6
False (DB position): 21
OH or SN resolved: 154
MSL resolved: 220
no MS/MS acquired: 355
False (SN position): 1

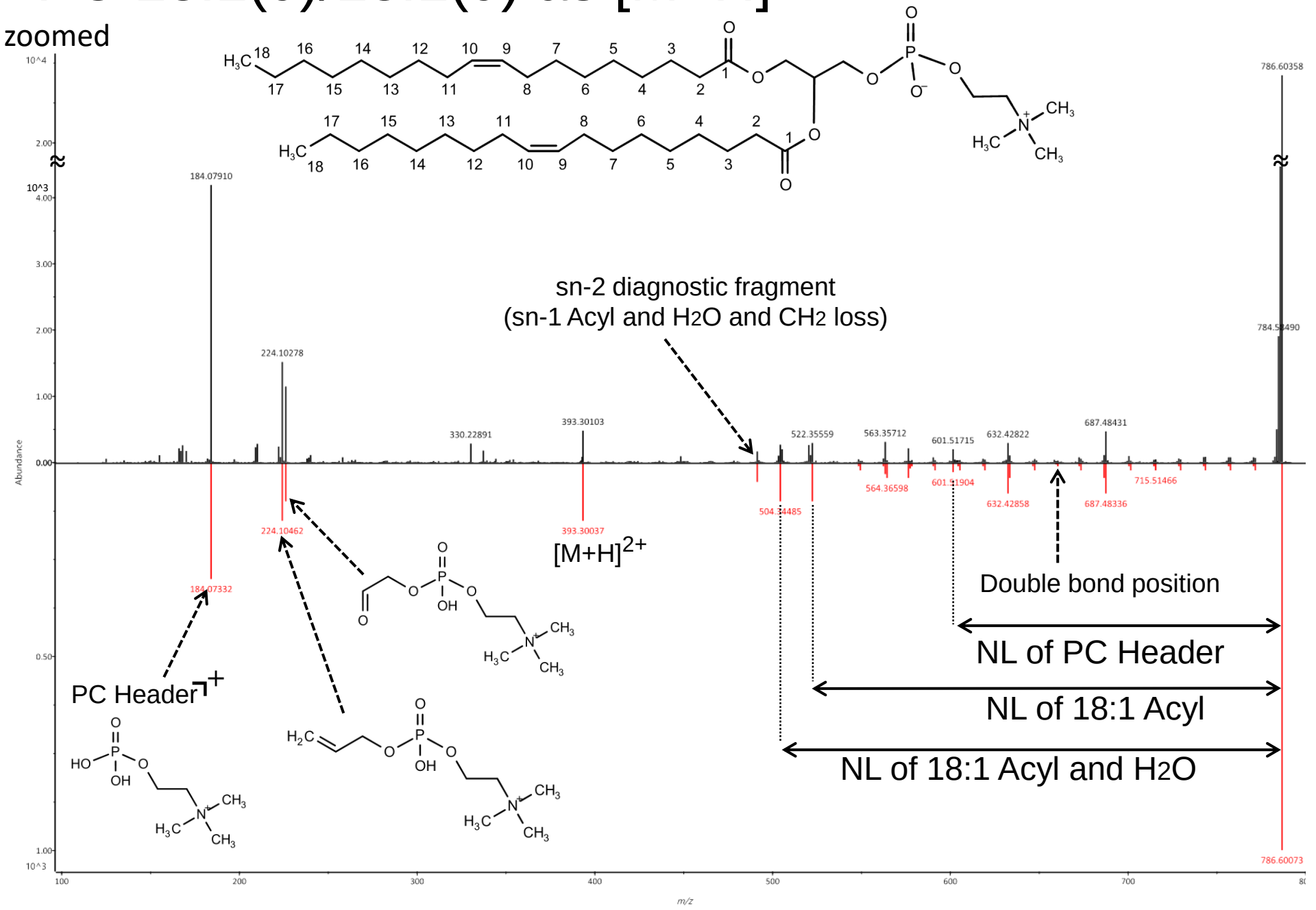

a
Parent
Subclass
Class
SuperClass
Count
8000
6000
4000
2000
0
905
964
1217
1448
1048
1118
1437
2154
972
1034
1330
1995
1027
1047
1444
2302
1041
1189
1551
2036
1117
1230
1634
2186
1265
1351
1979
2750
1378
1481
2134
2985
847
1042
1423
2046
CID 10V
CID 10V CES 15V
CID 20V
CID 20V CES 15V
CID 40V
CID 40V CES 15V
EAD 10eV CID 10V
EAD 15eV CID 10V
EAD 20eV CID 10V
b
CID 40V CES 15V
KE 15 eV CID 10V
Lipids and lipid-like molecules
Phenylpropanoids and polyketides
Alkaloids and derivatives
Organoheterocyclic compounds
Benzenoids
Organosulfur compounds
Organic oxygen compounds
Nucleosides, nucleotides, and analogues
Organic acids derivatives

PC 18:1(9)/18:1(9) as [M+H]+

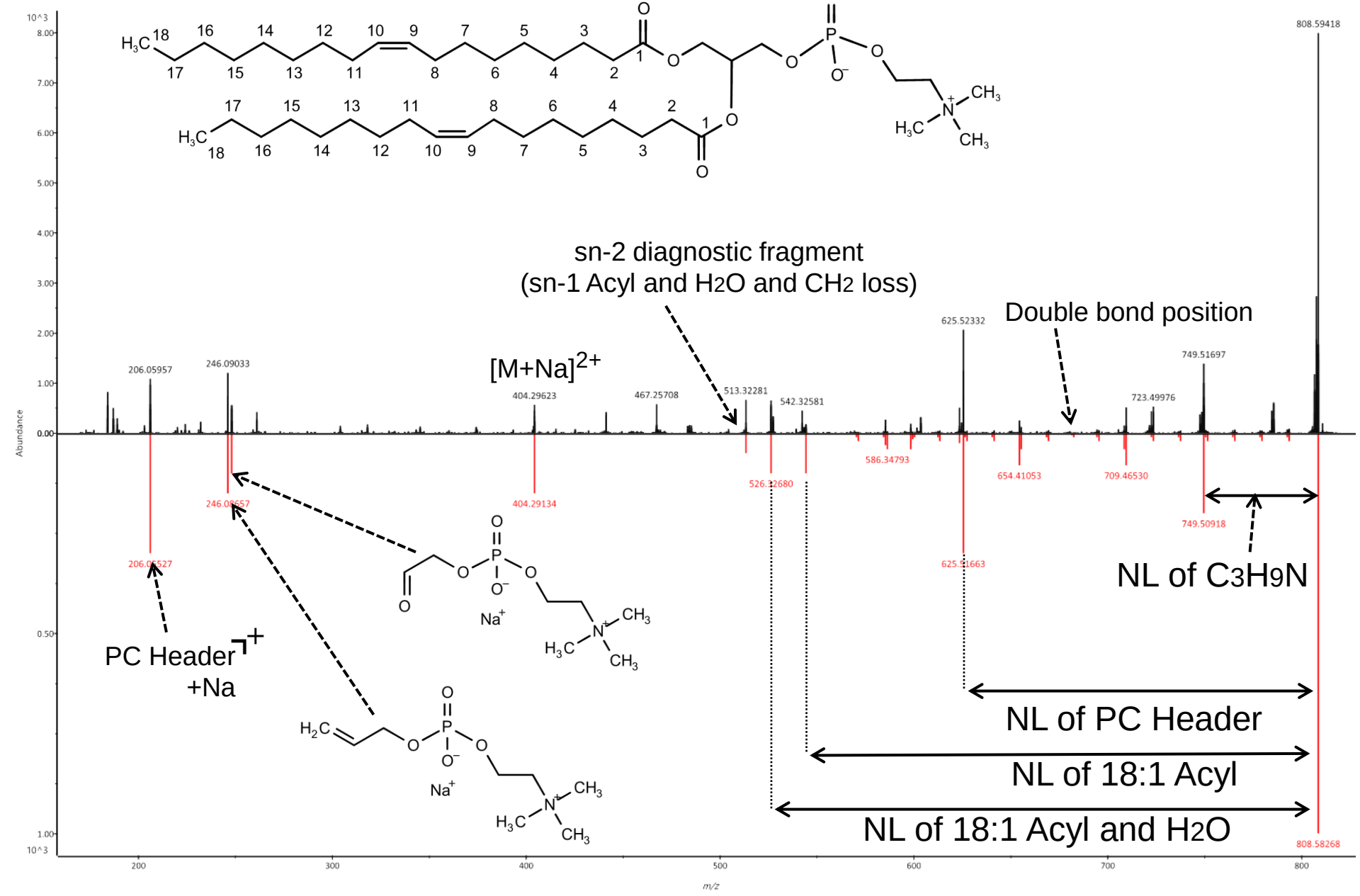


PC 18:1(9)/18:1(9) as [M+H]+

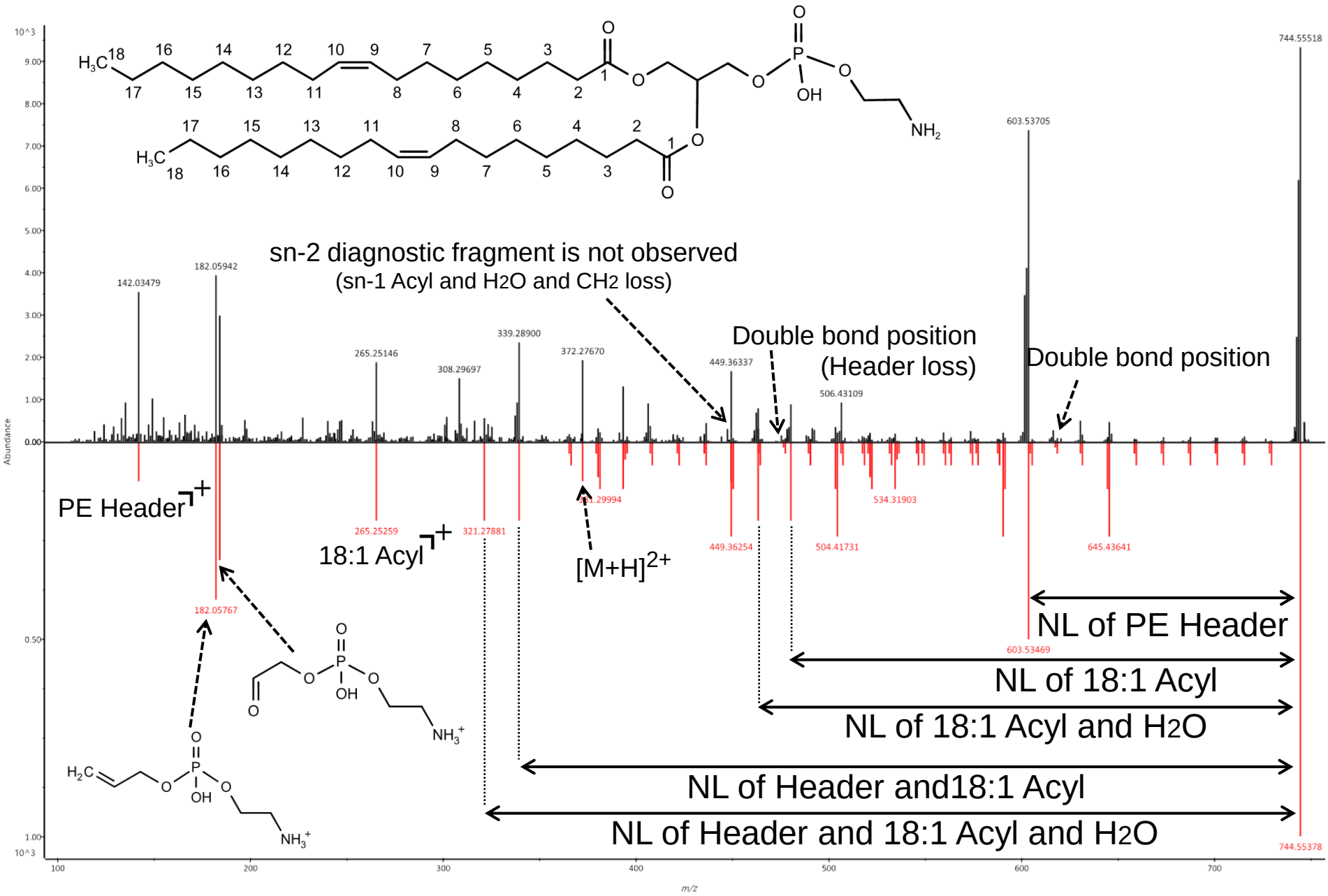
zoomed



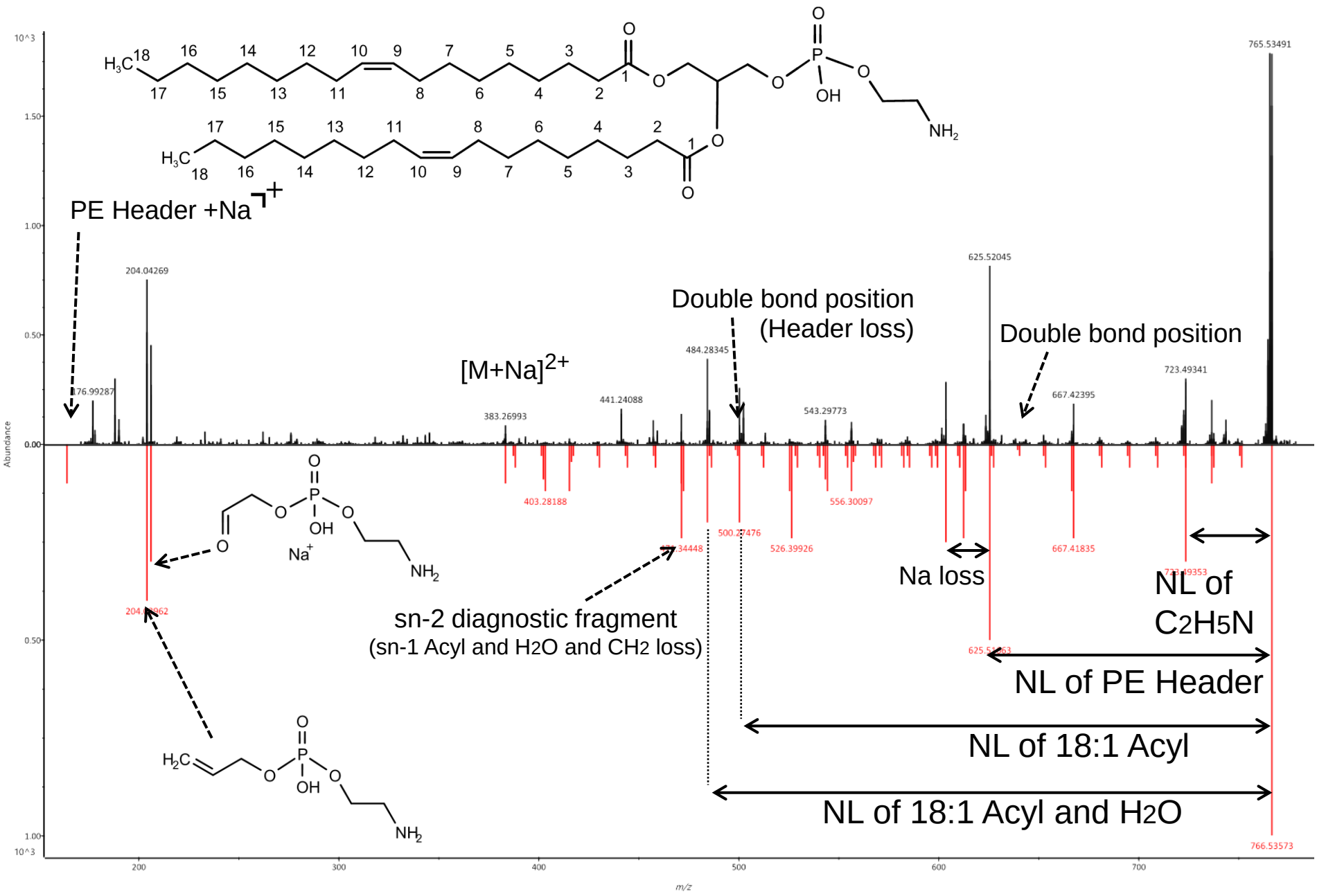
PC 18:1(9)/18:1(9) as [M+Na]+



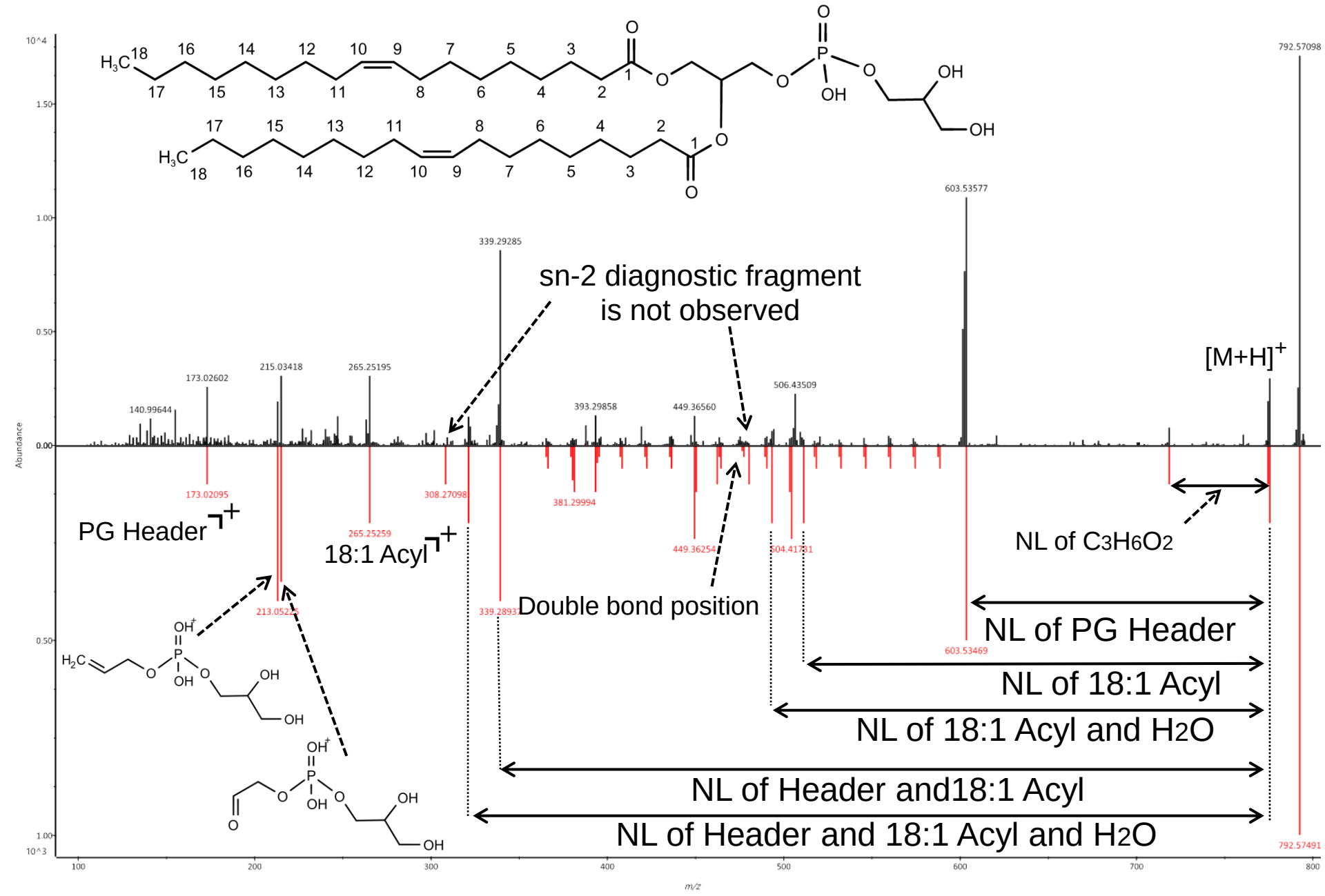
PE 18:1(9)_18:1(9) as [M+H]+



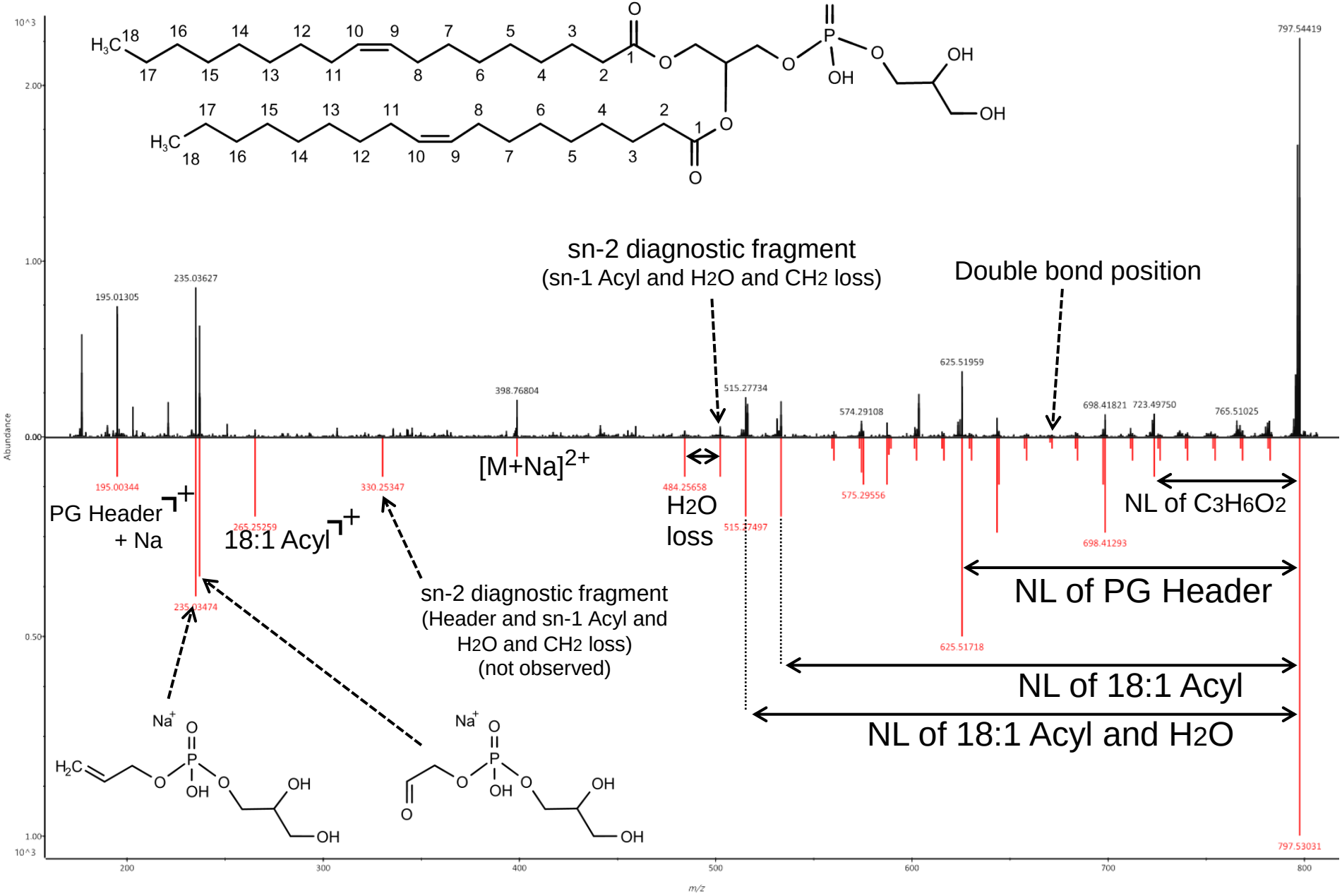
PE 18:1(9)/18:1(9) as [M+Na]+



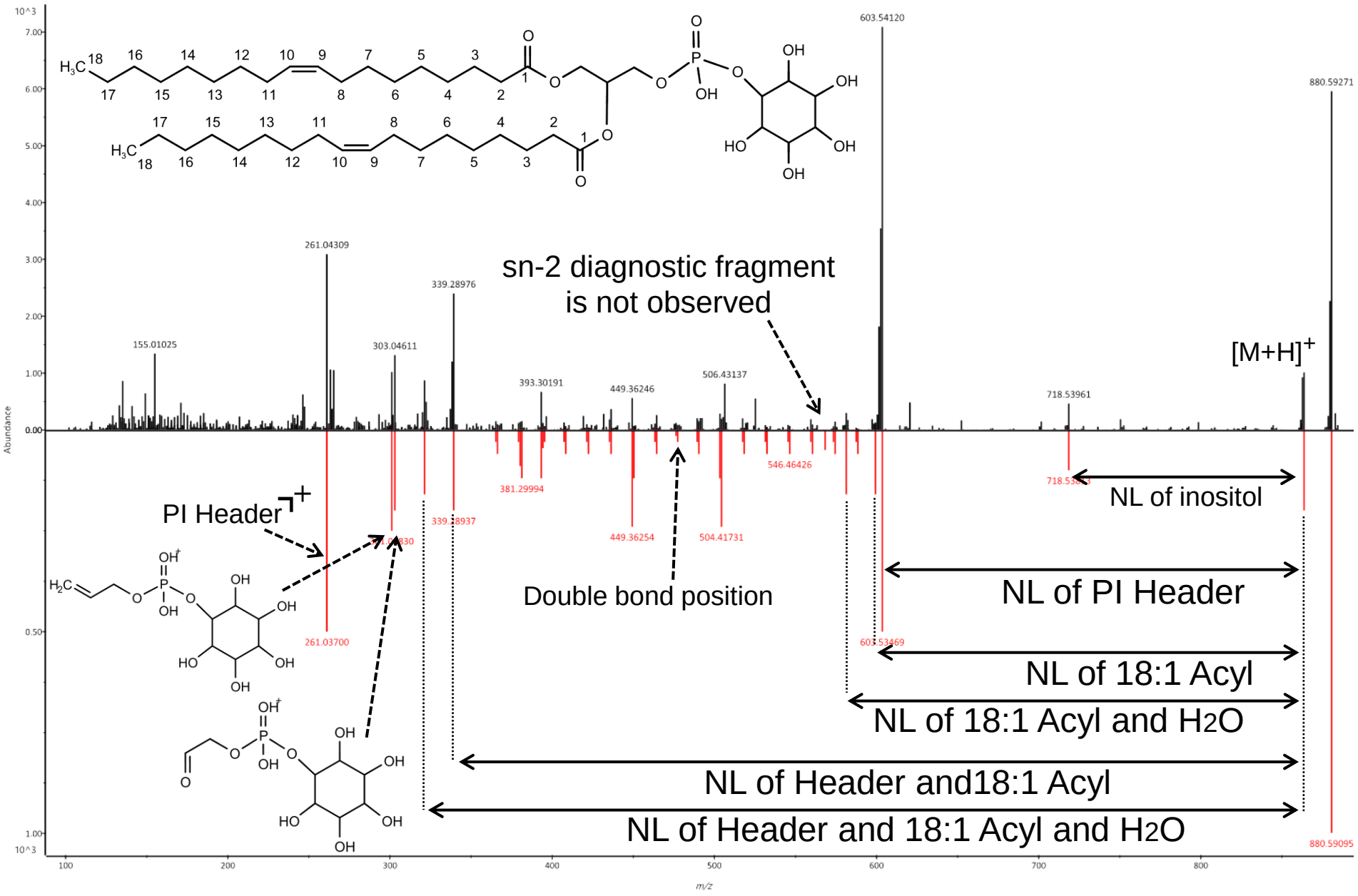
PG 18:1(9)_18:1(9) as [M+NH4]+



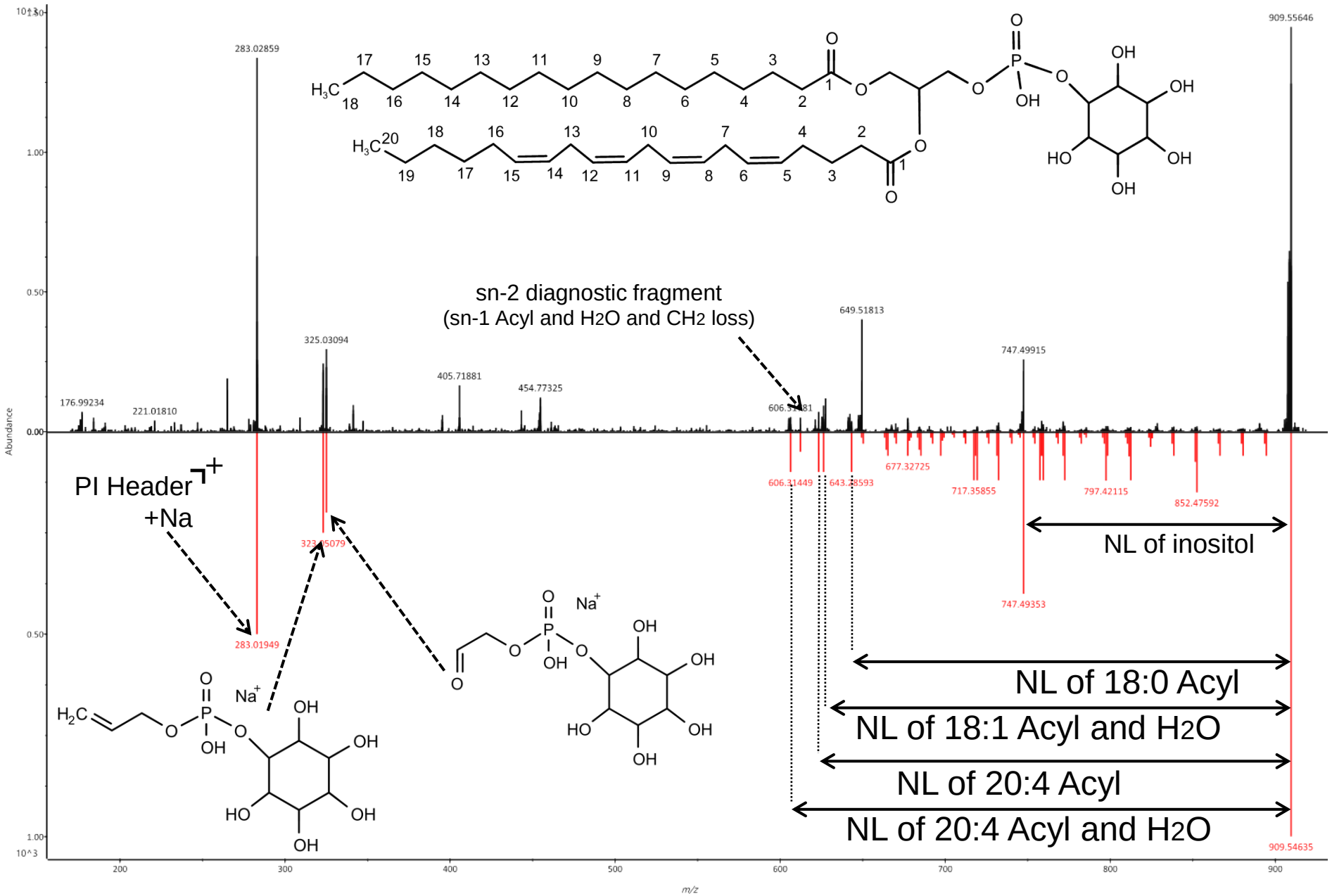
PG 18:1(9)/18:1(9) as [M+Na]+



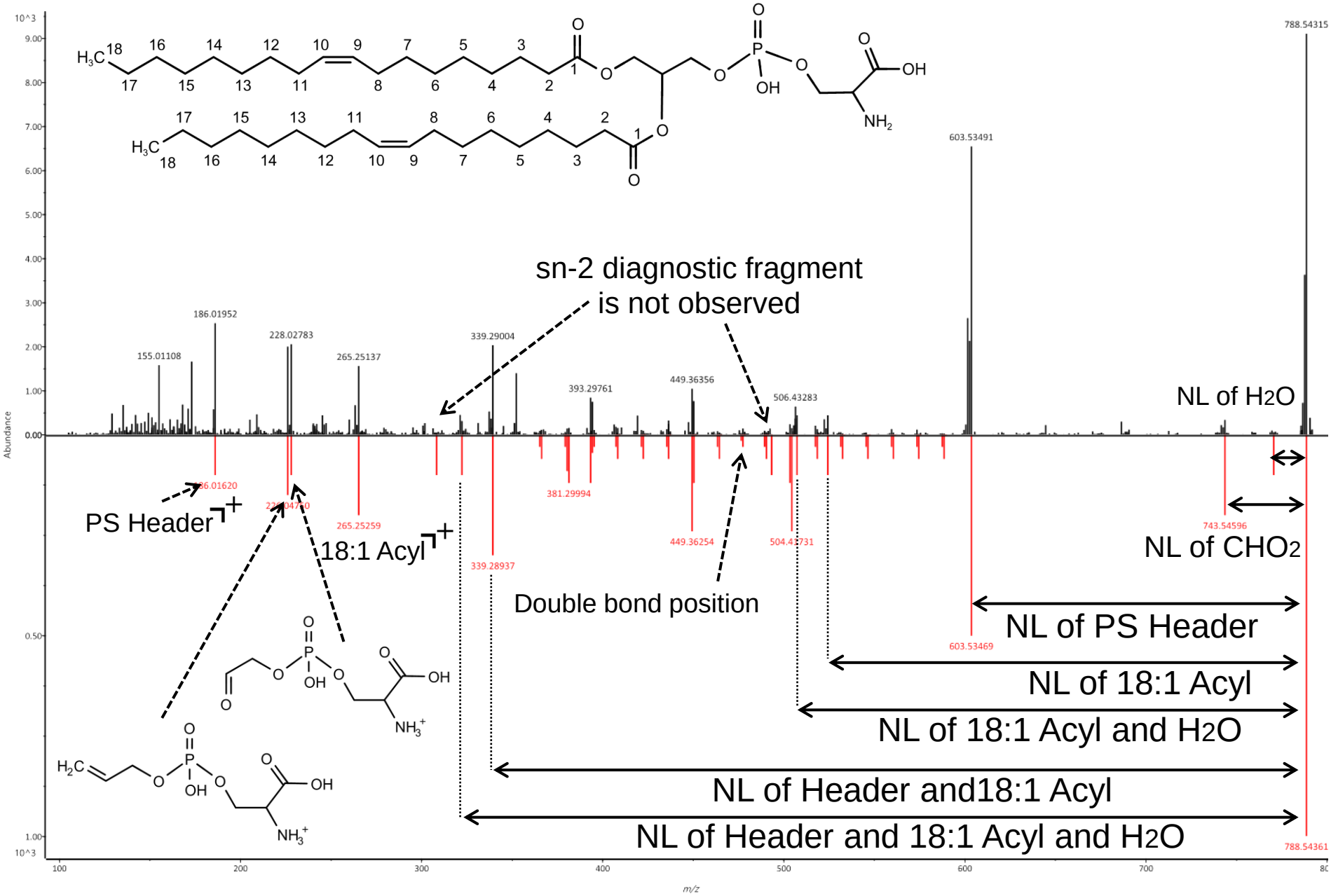
PI 18:1(9)_18:1(9) as [M+NH4]+



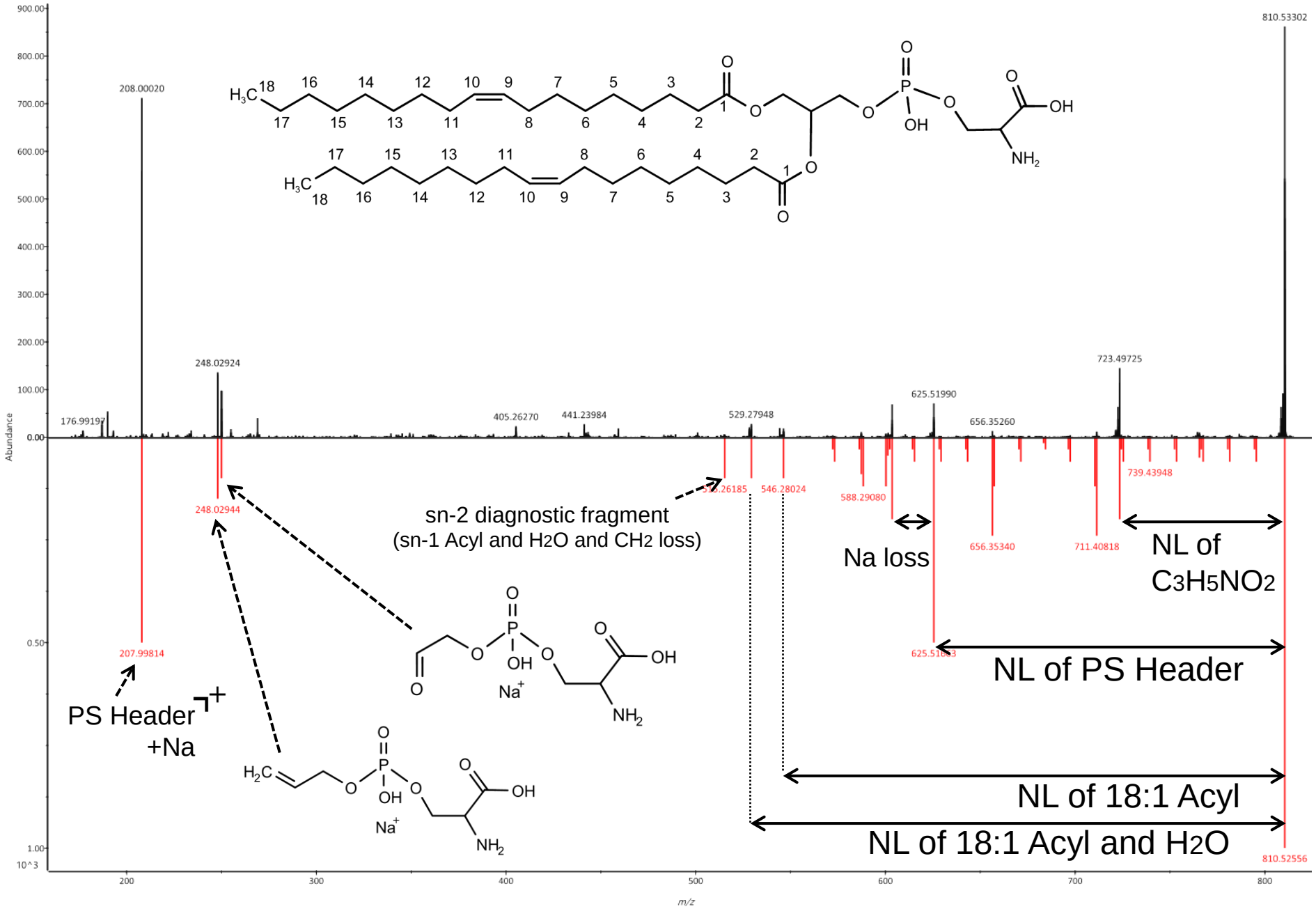
PI 18:0/20:4 as [M+Na]+



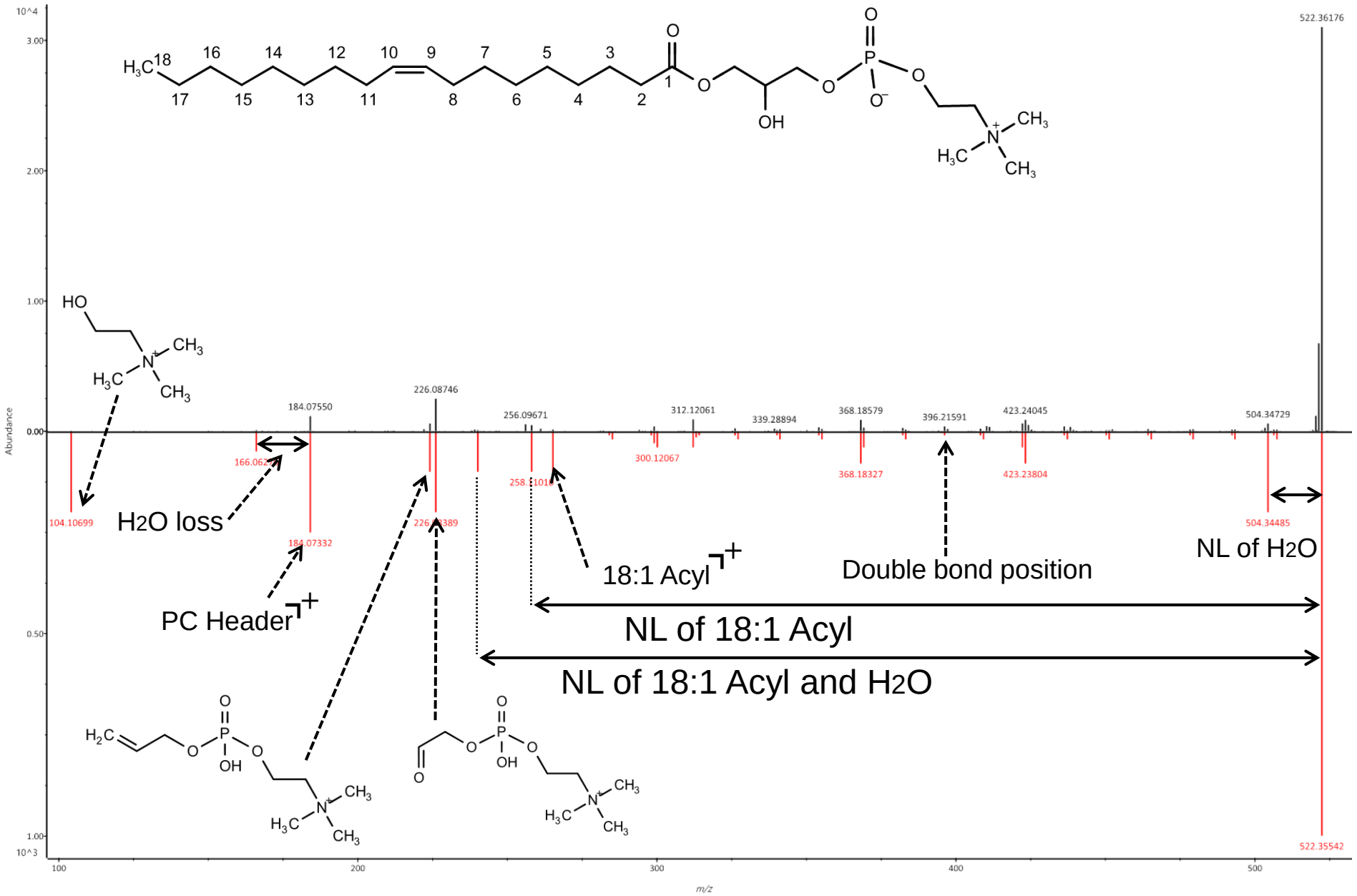
PS 18:1(9)_18:1(9) as [M+H]+



PS 18:1/18:1 as [M+Na]+

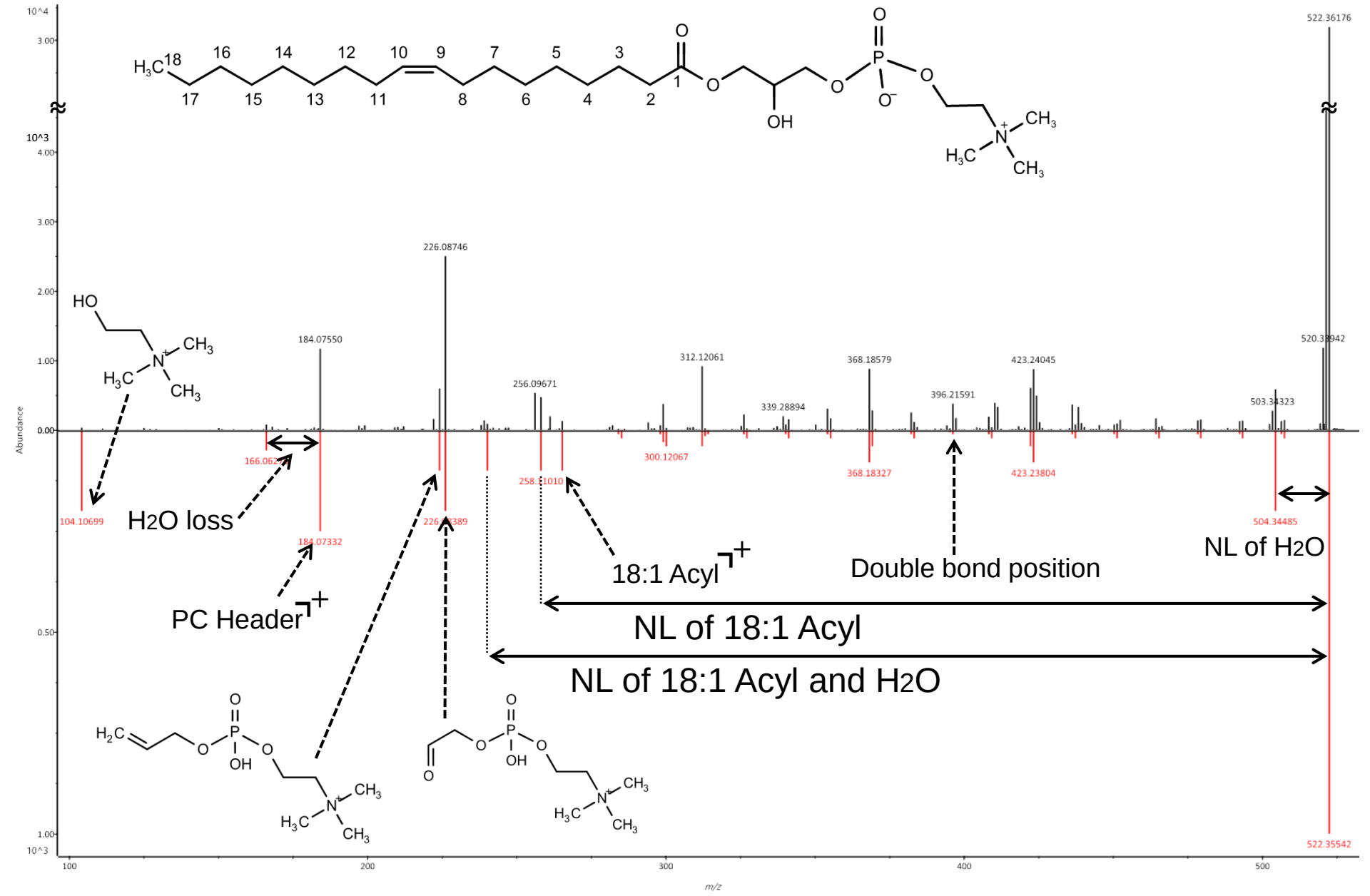


LPC 18:1(9)/0:0 as [M+H]+

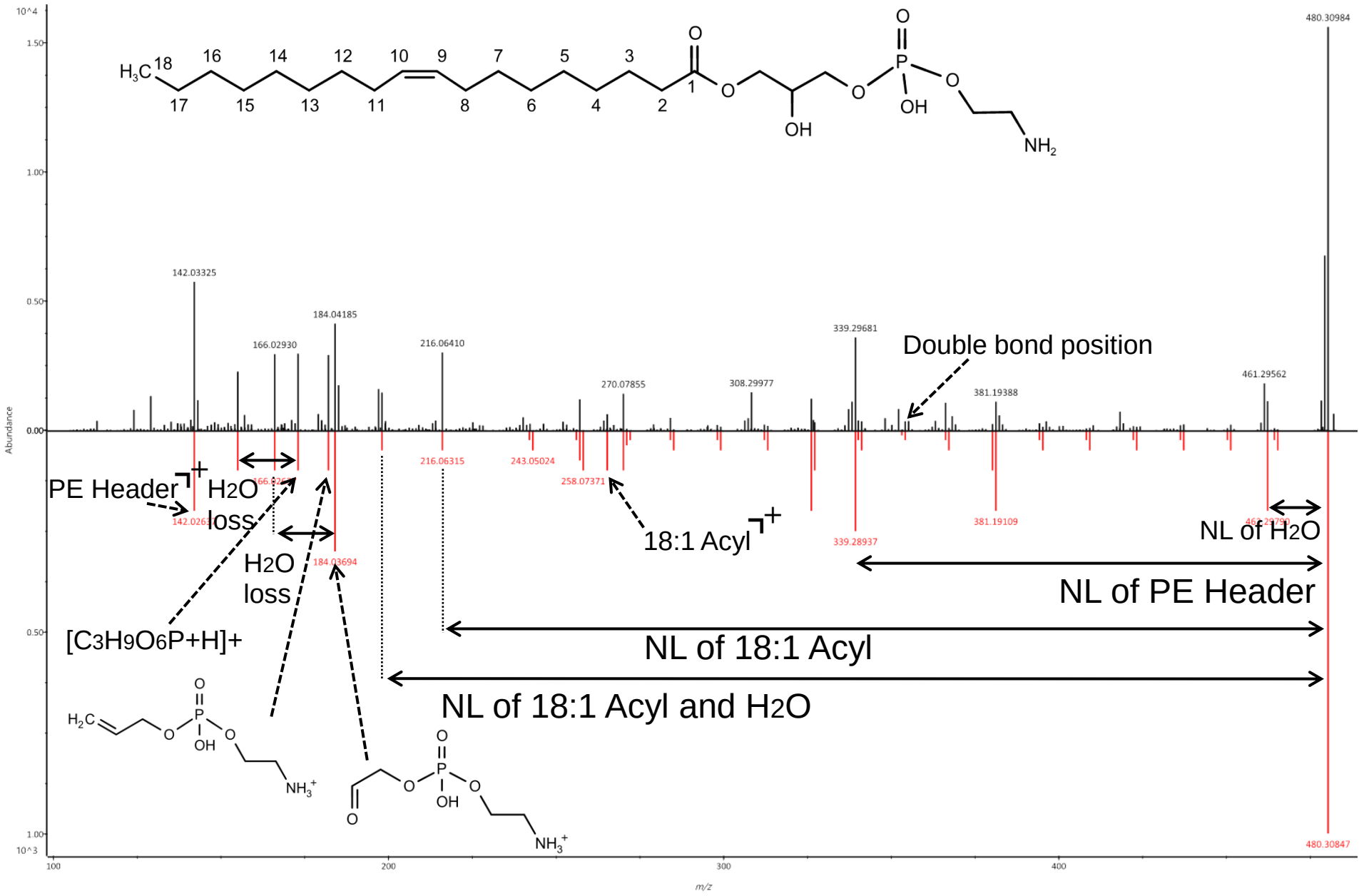


LPC 18:1(9)/0:0 as [M+H]+

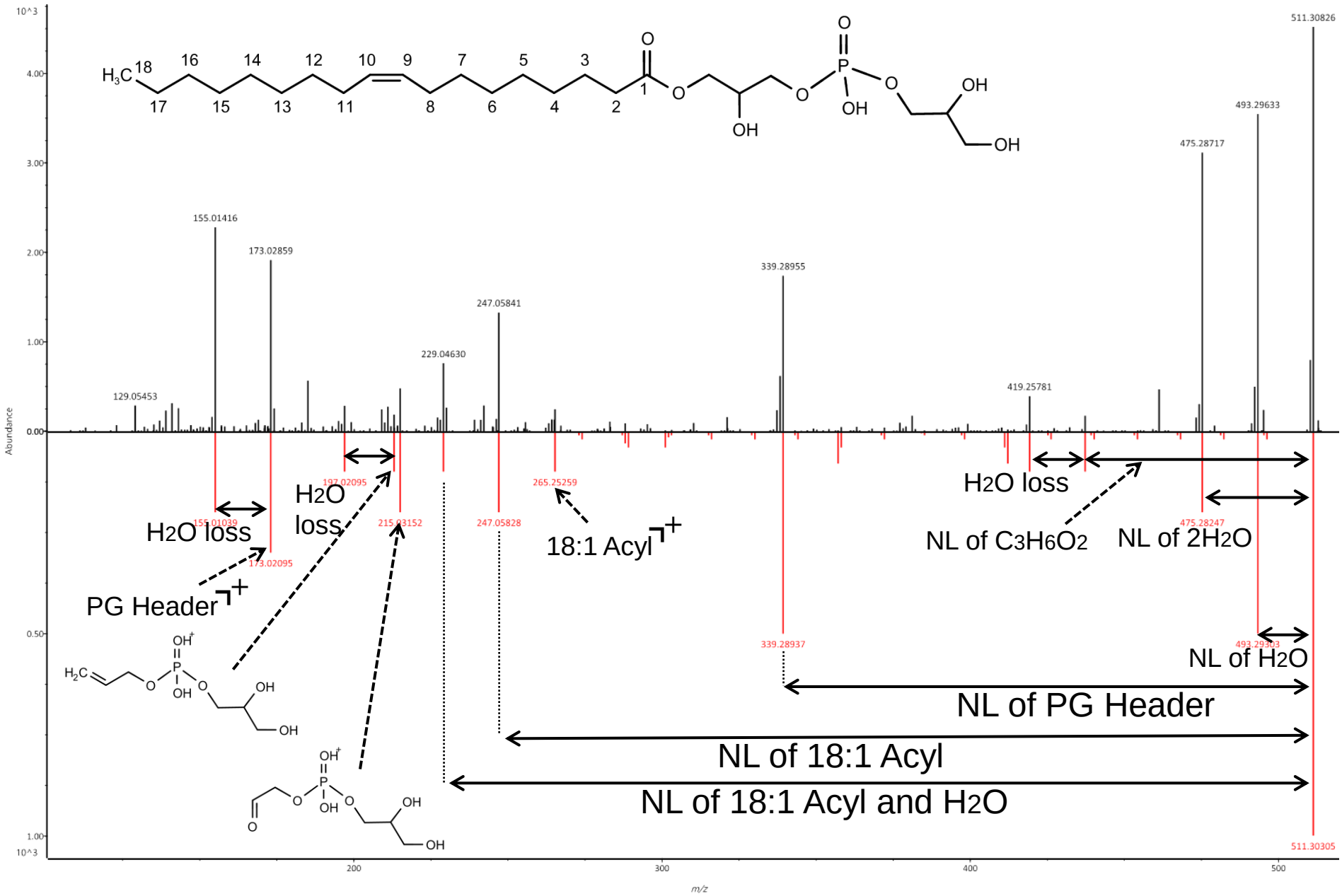
zoomed



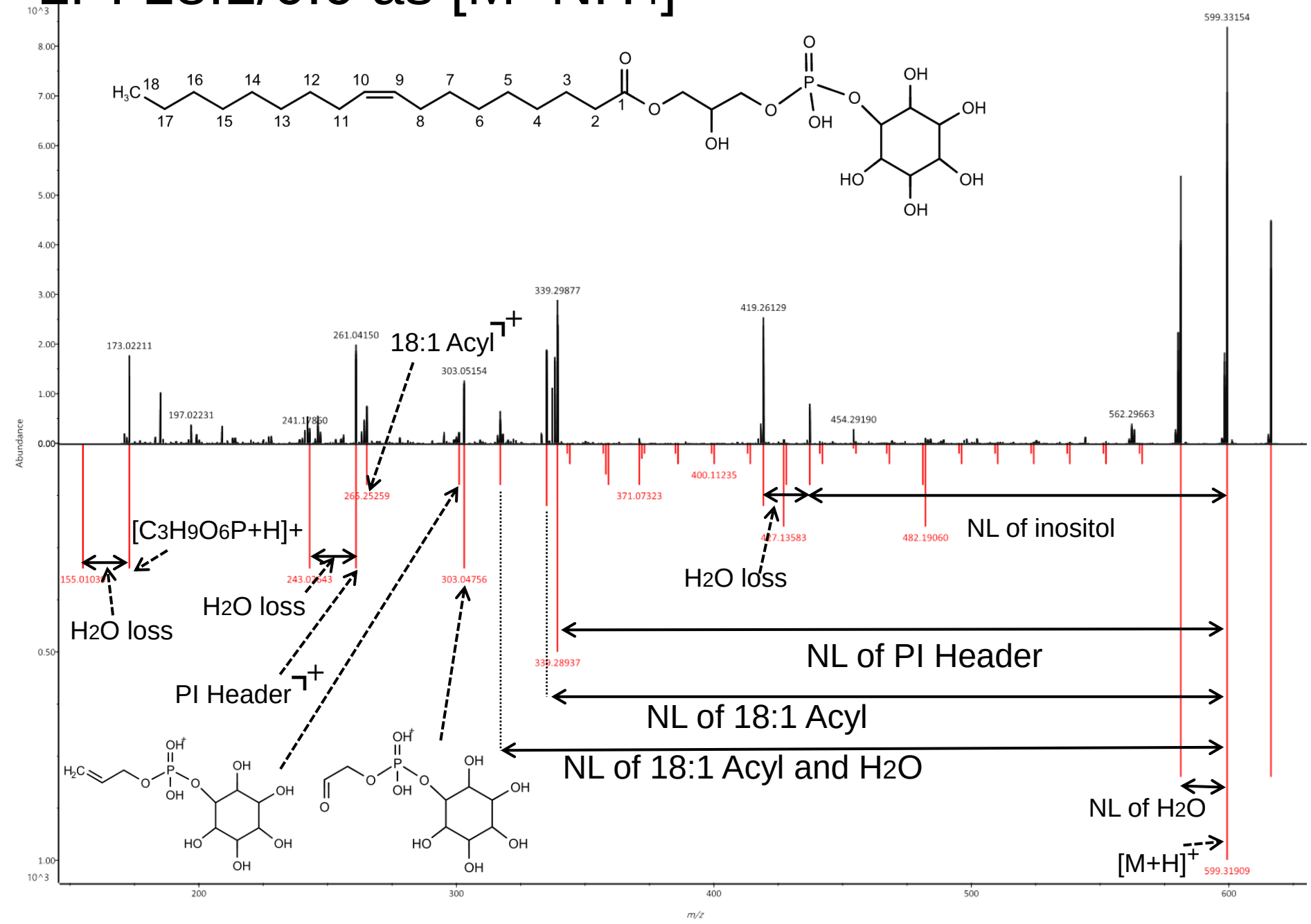
LPE 18:1(9)/0:0 as [M+H]+



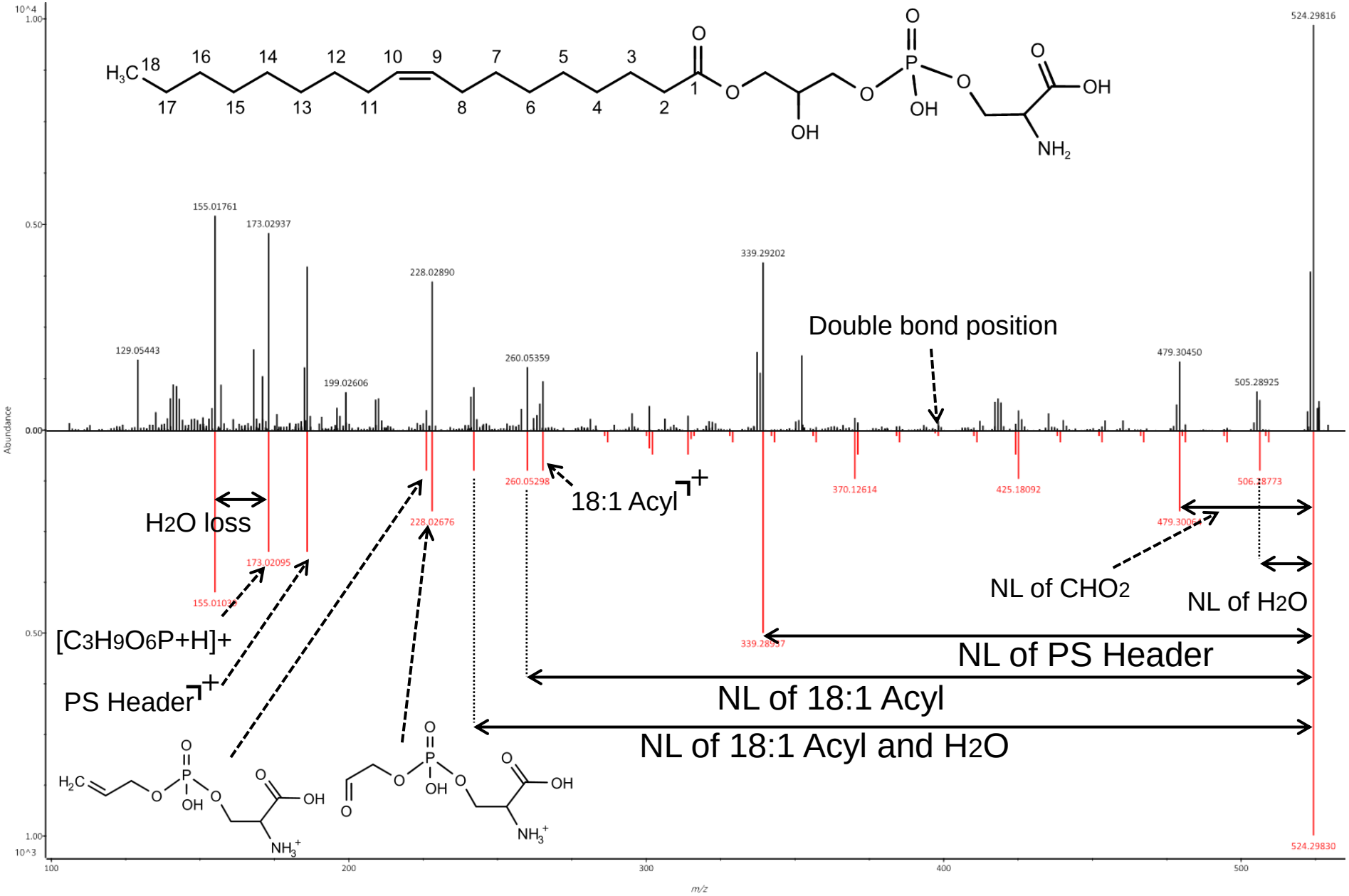
LPG 18:1/0:0 as [M+H]+



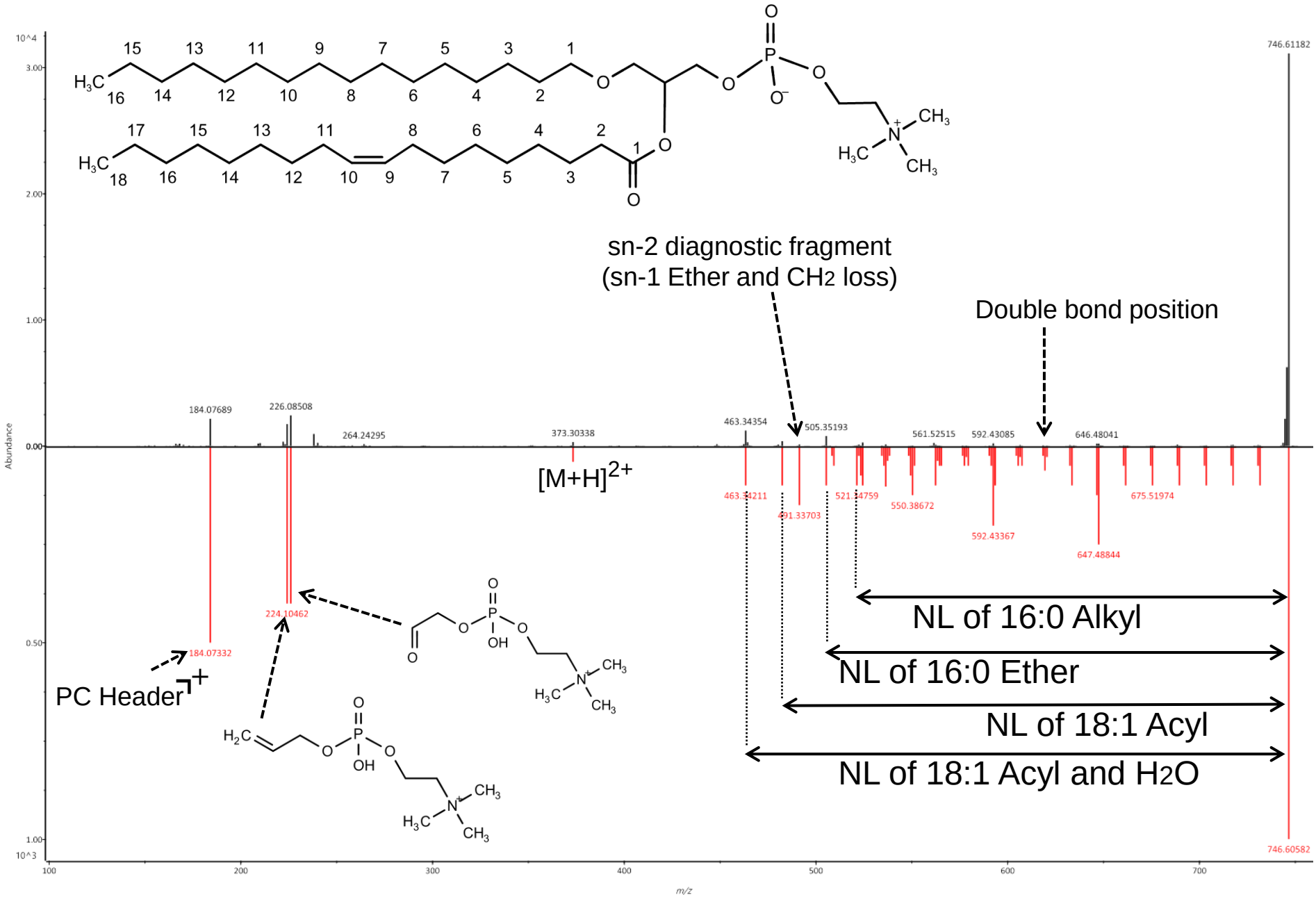
LPI 18:1/0:0 as [M+NH4]+



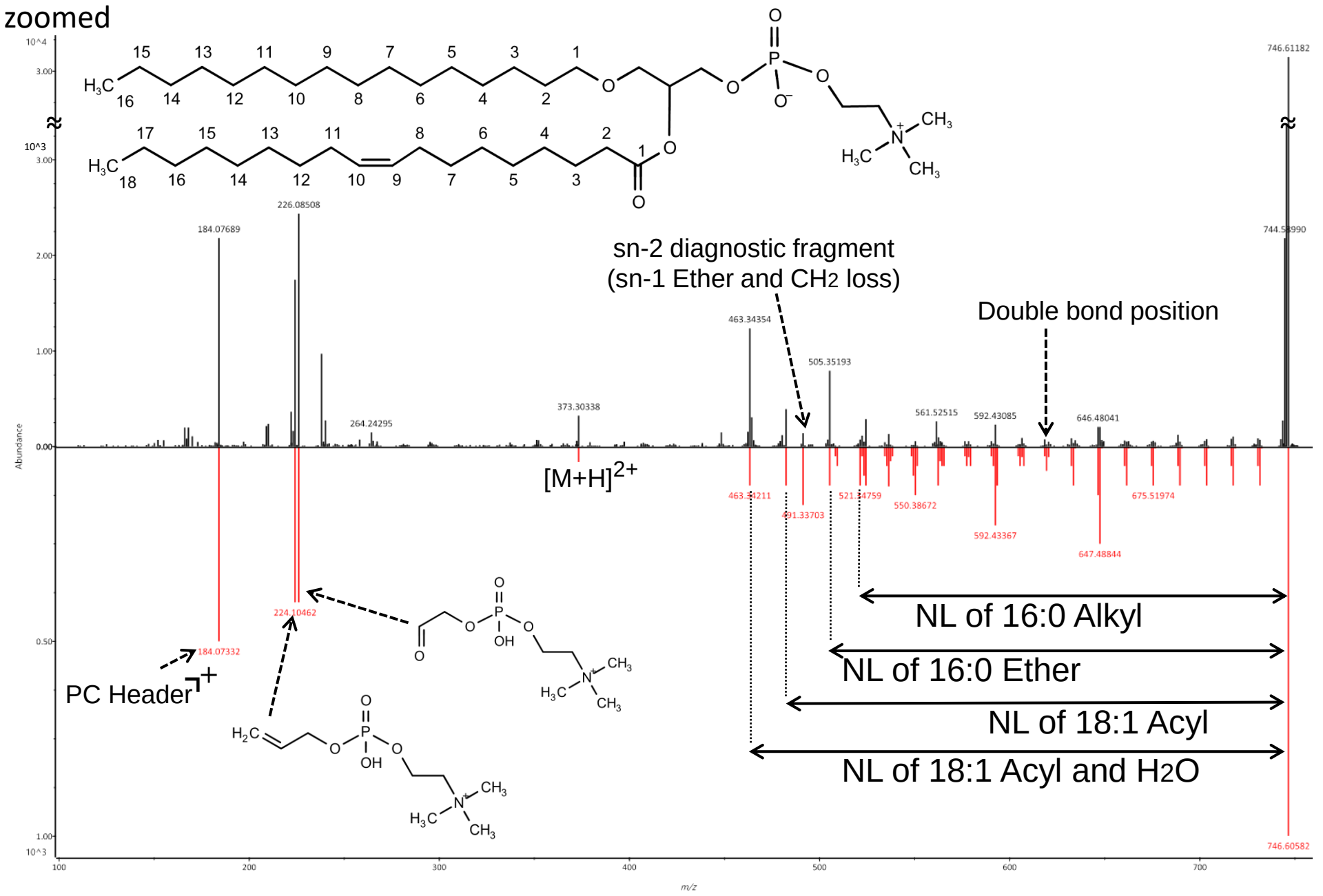
LPS 18:1(9)/0:0 as [M+H]+



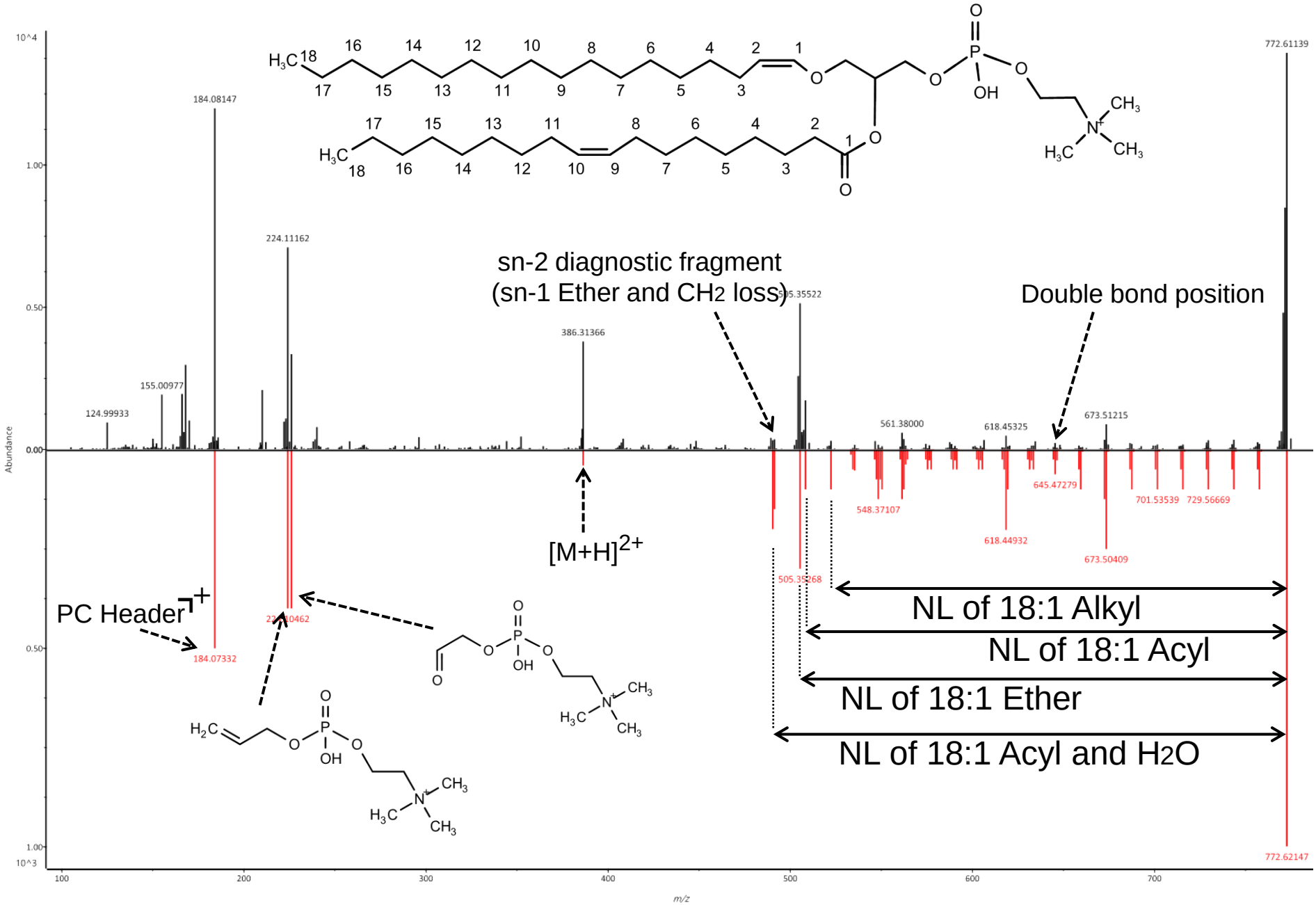
PC O-16:0/18:1(9) as [M+H]+



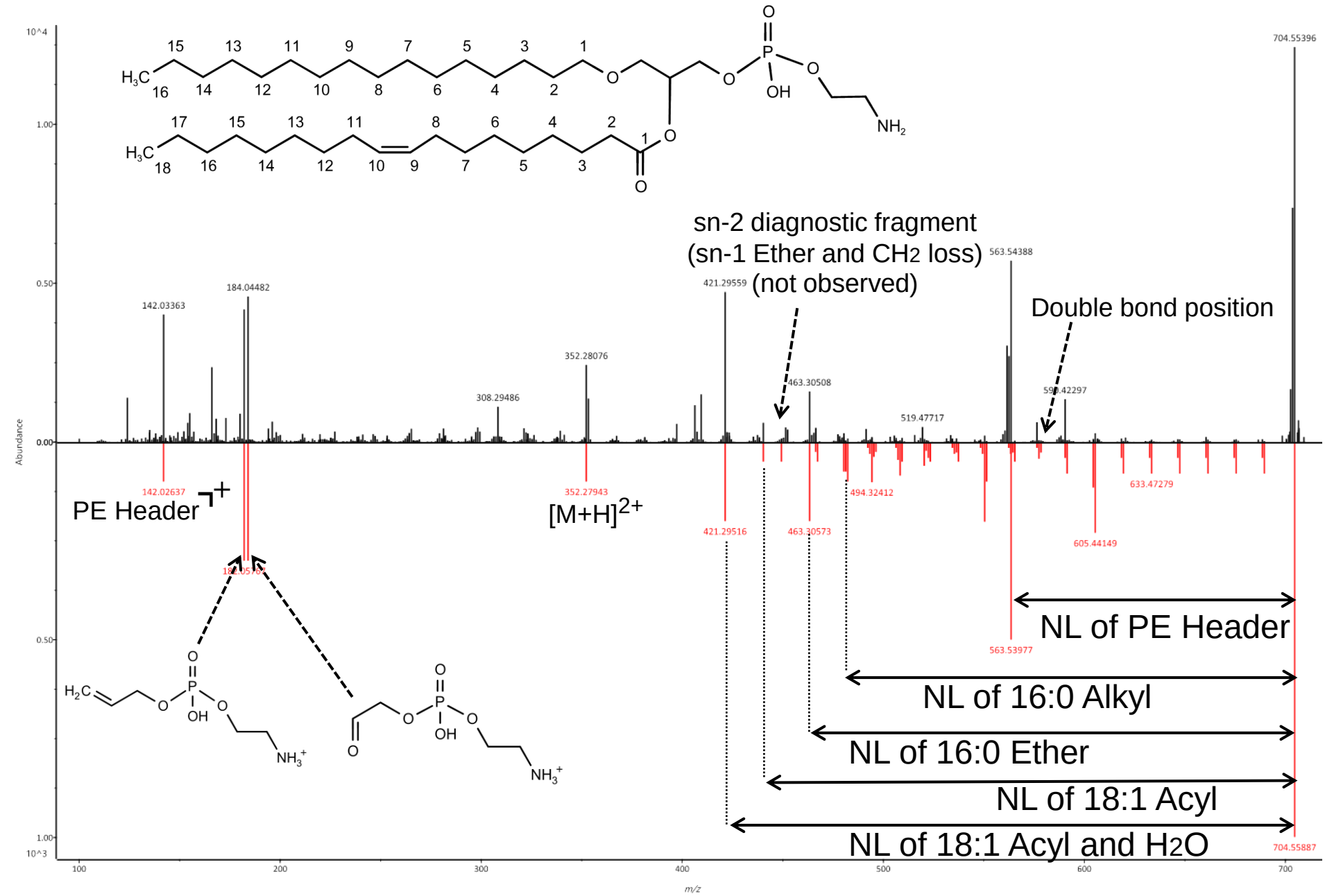
PC O-16:0/18:1(9) as [M+H]+



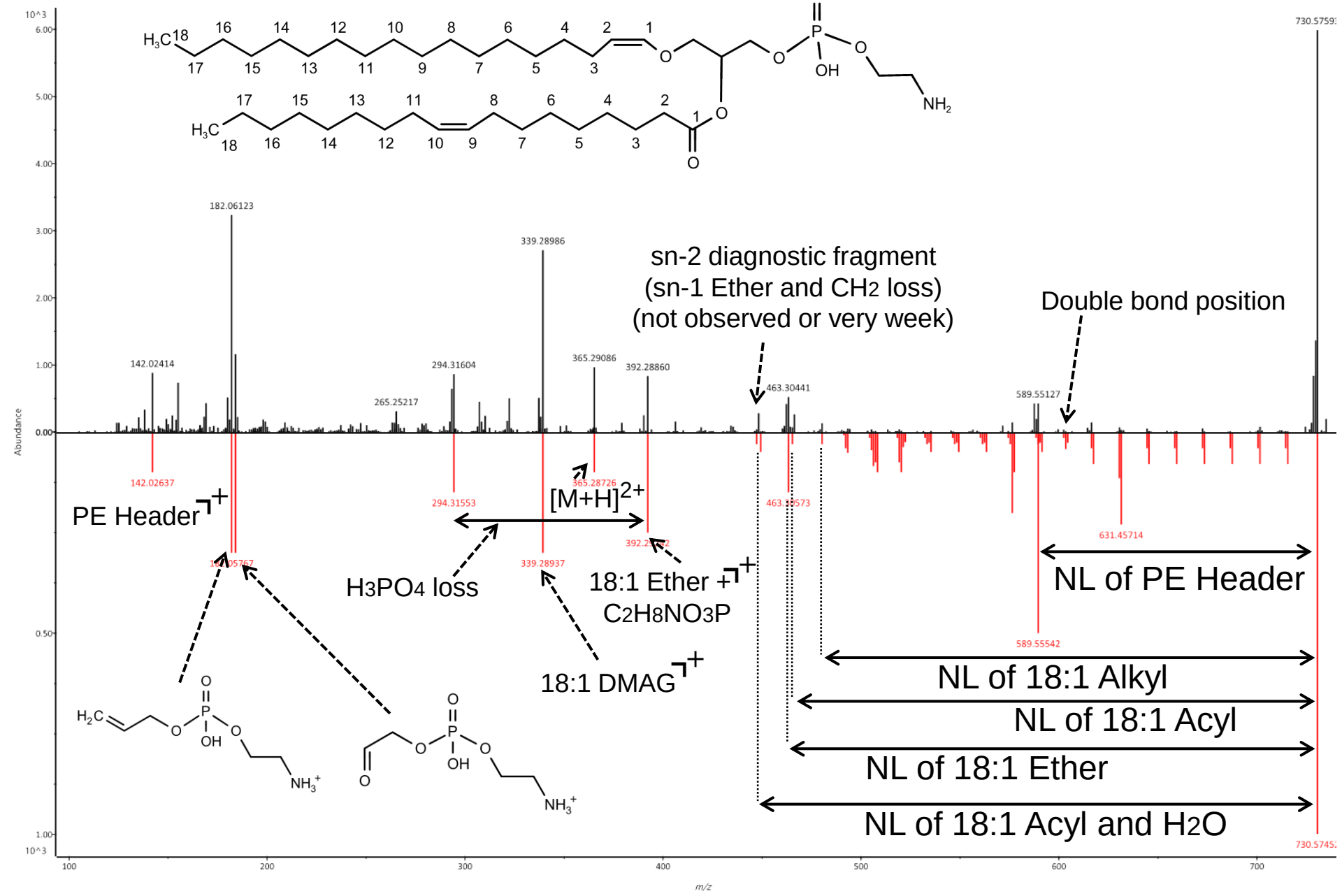
PC P-18:0/18:1(9) as [M+H]+



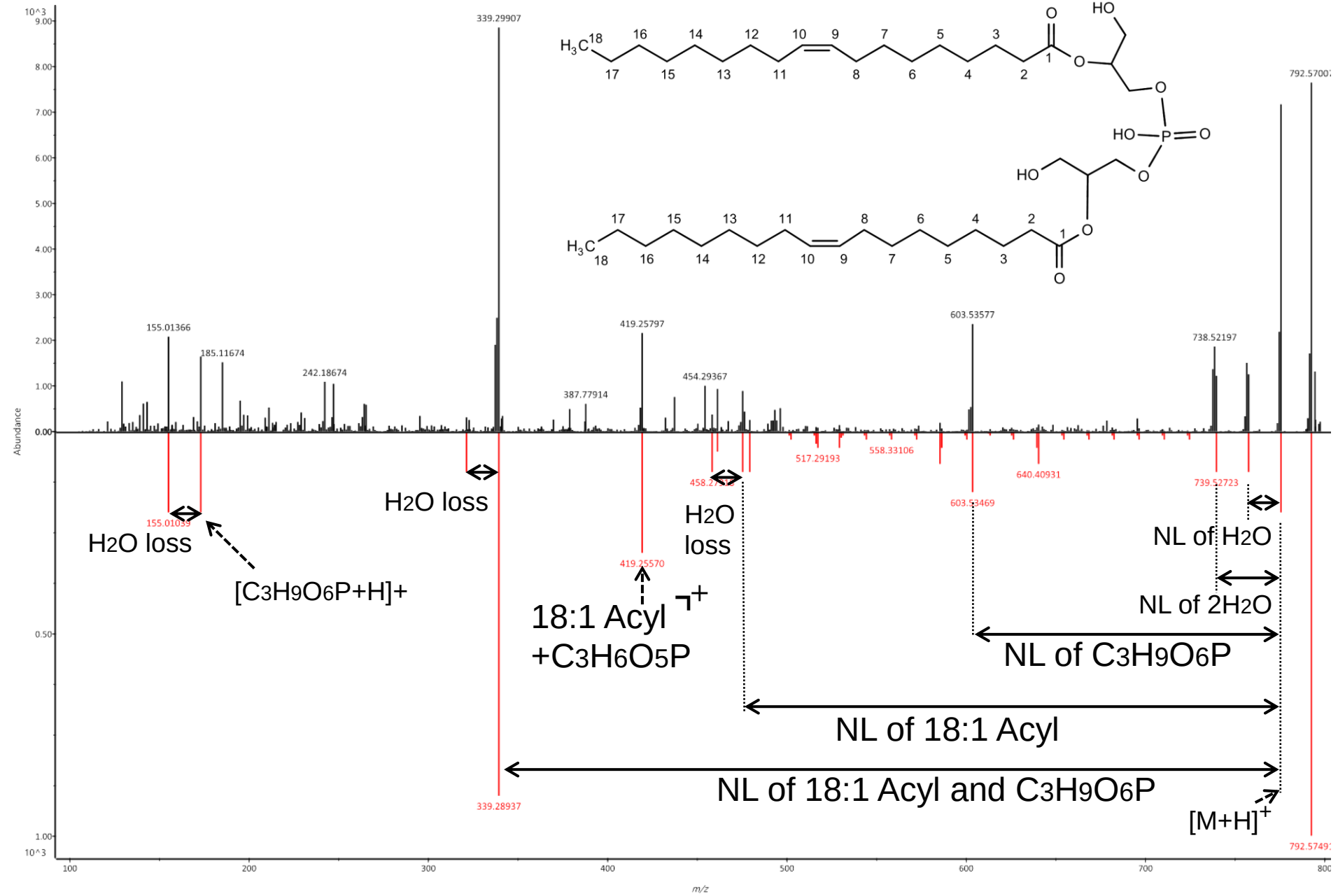
PE O-16:0_18:1(9) as [M+H]+



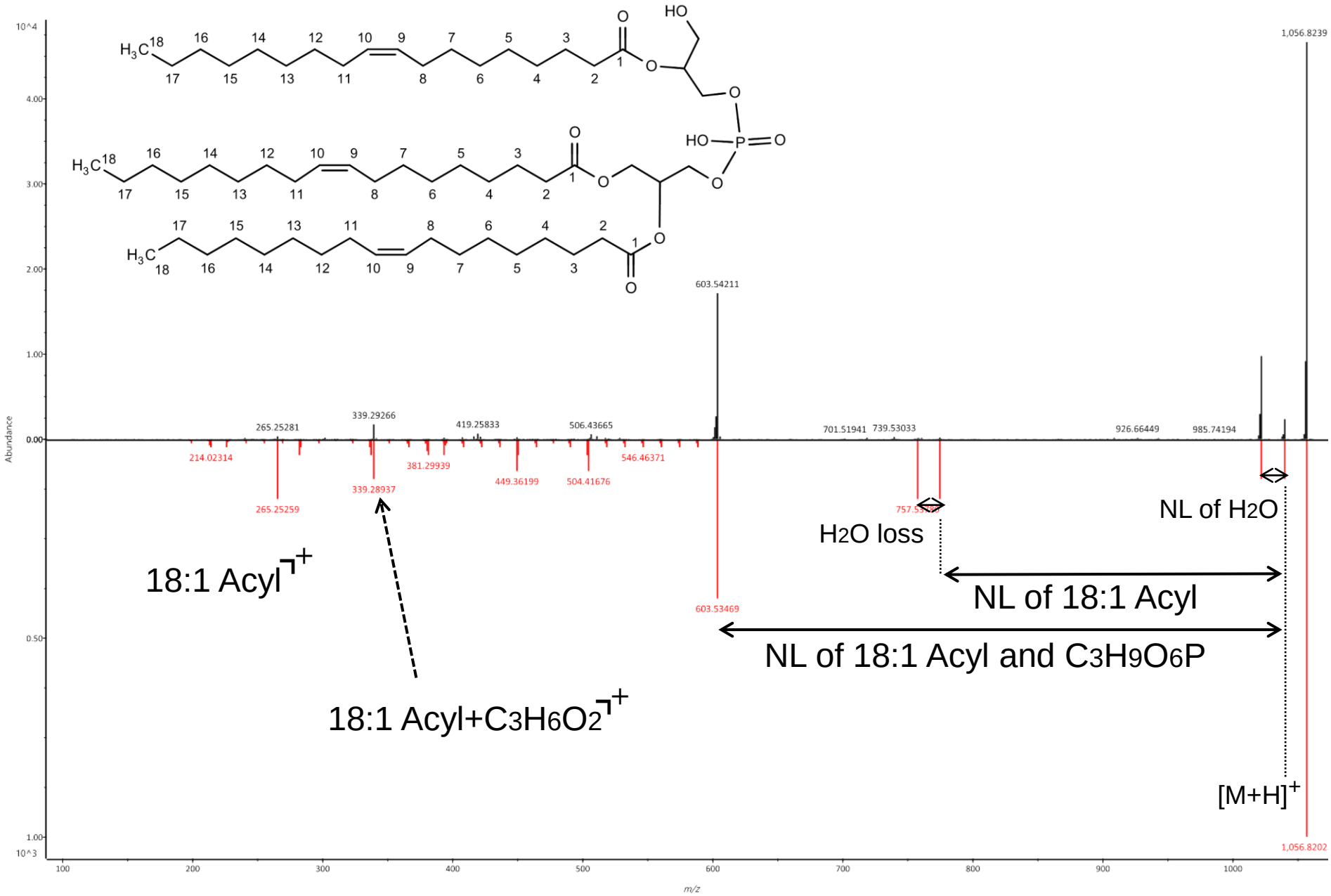
PE P-18:0_18:1(9) as [M+H]+



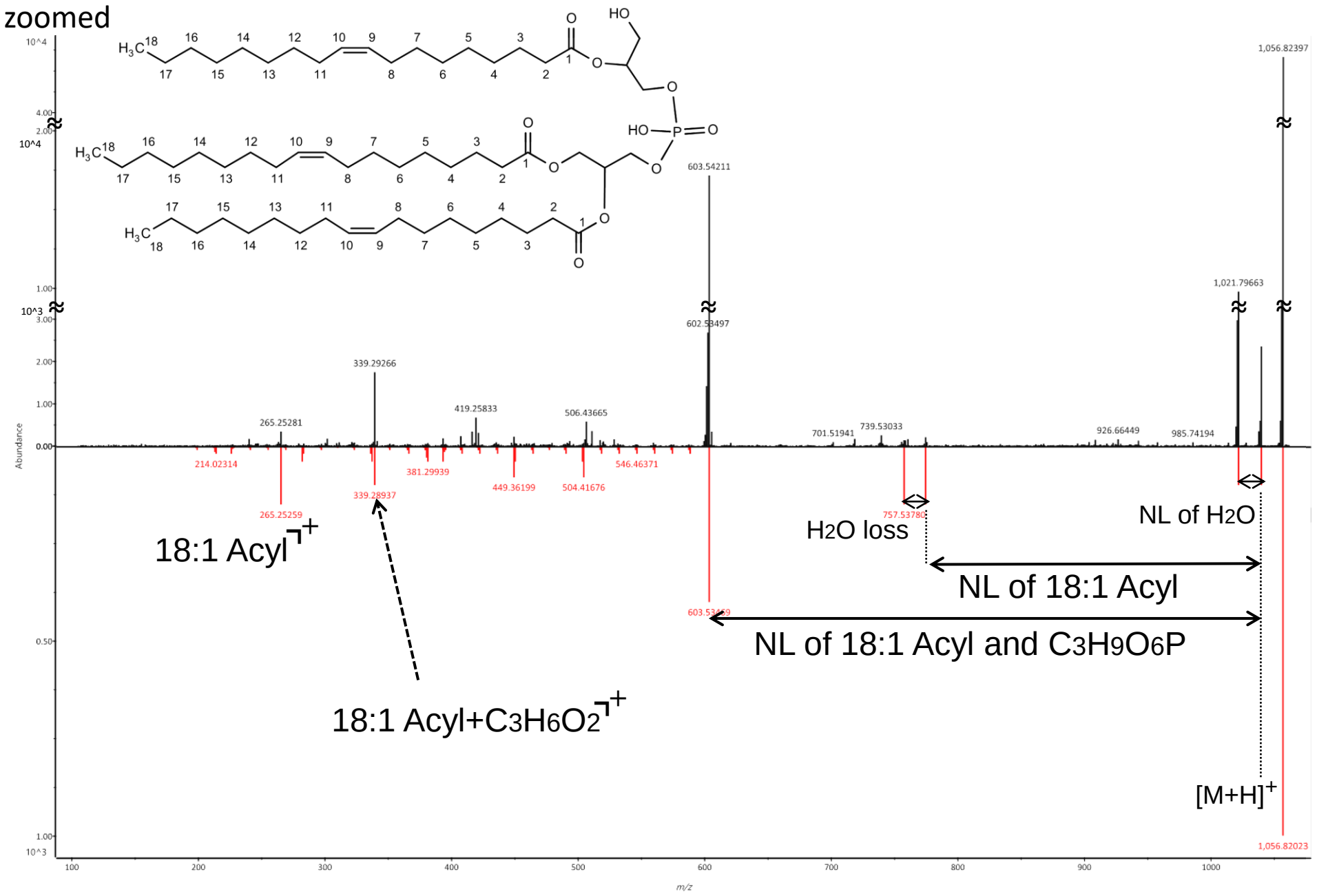
BMP 18:1(9)_18:1(9) as $[M+NH_4]^+$



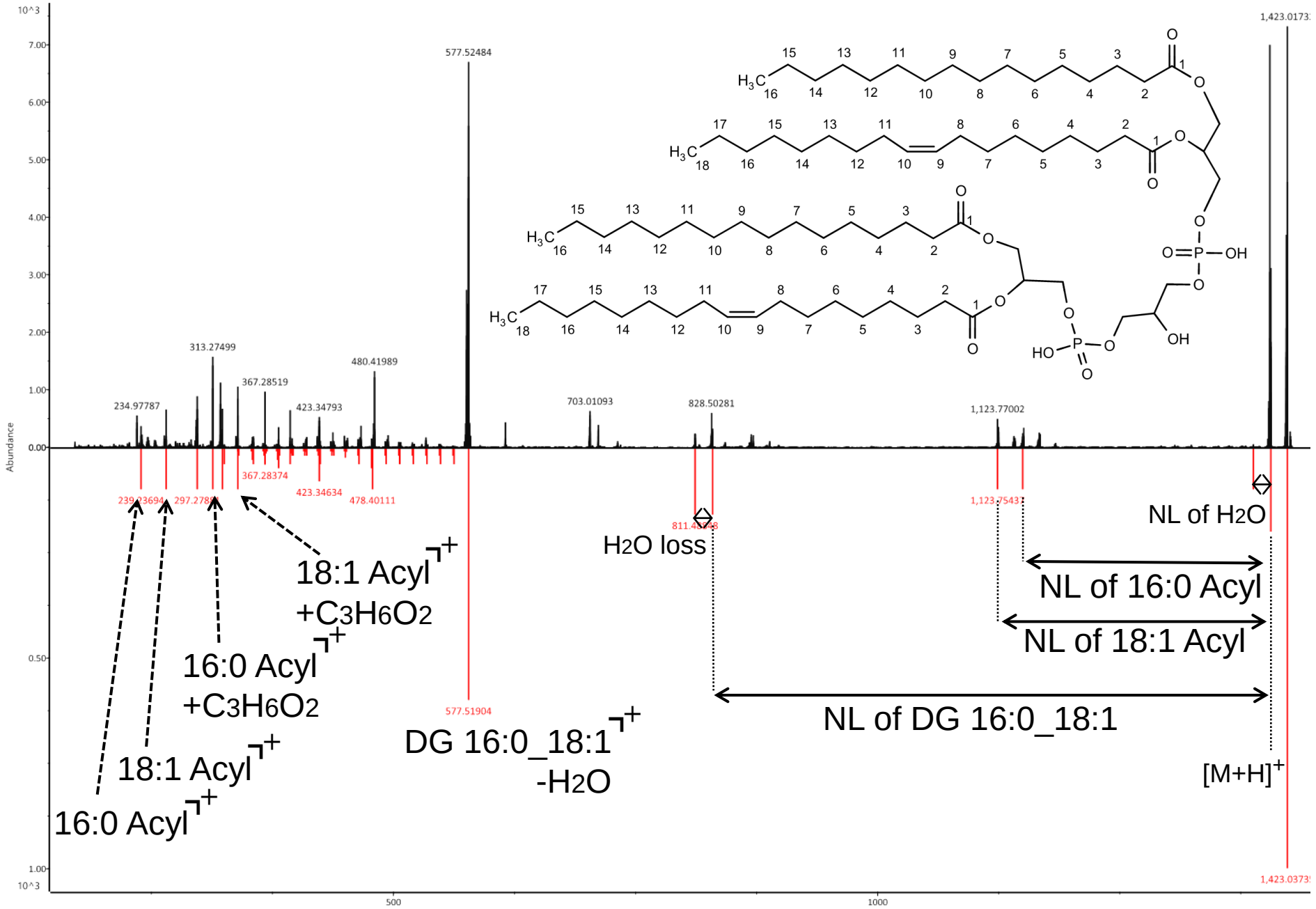
HBMP 18:1_18:1_18:1 as $[M+NH_4]^+$



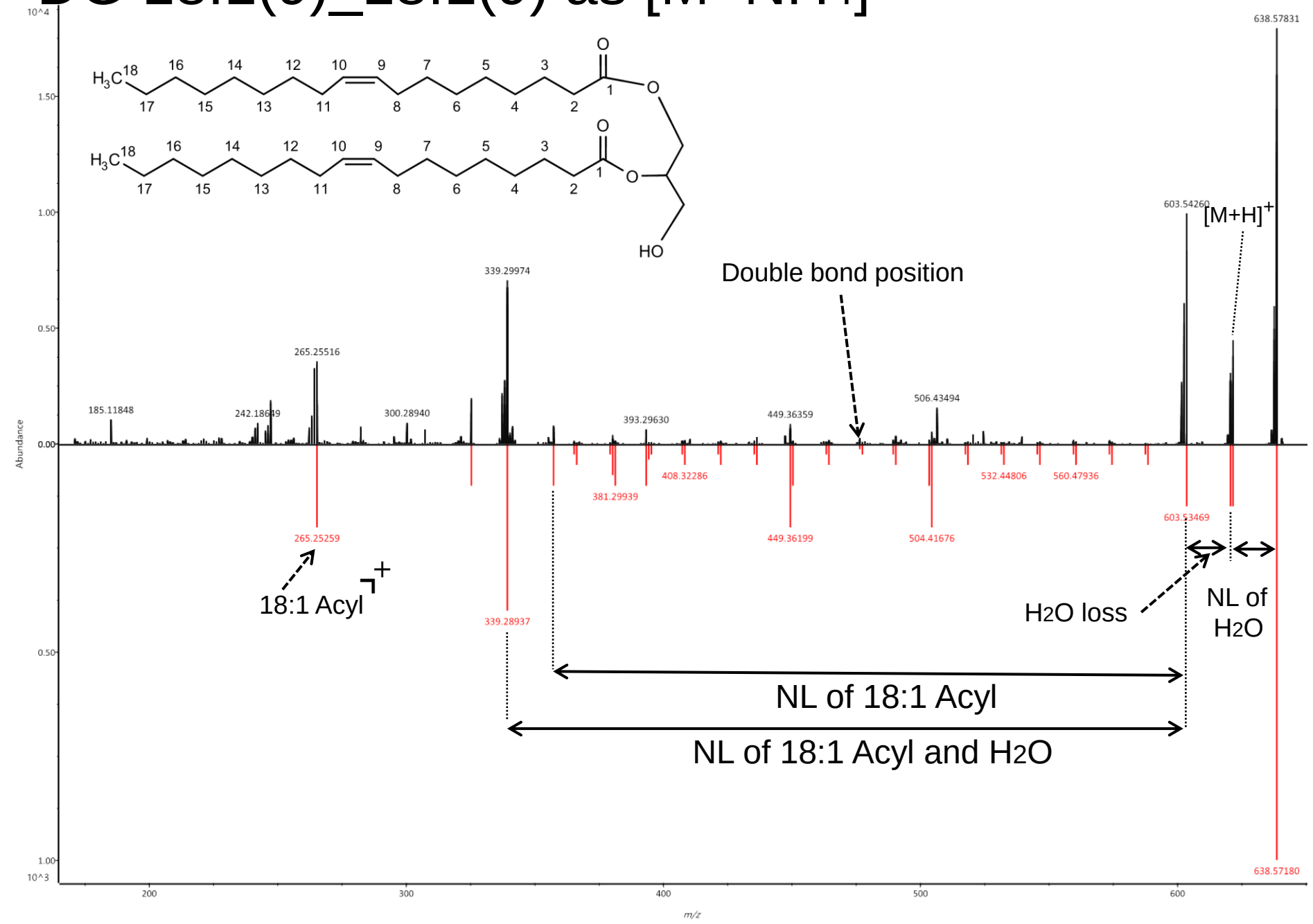
HBMP 18:1_18:1_18:1 as $[M+NH_4]^+$



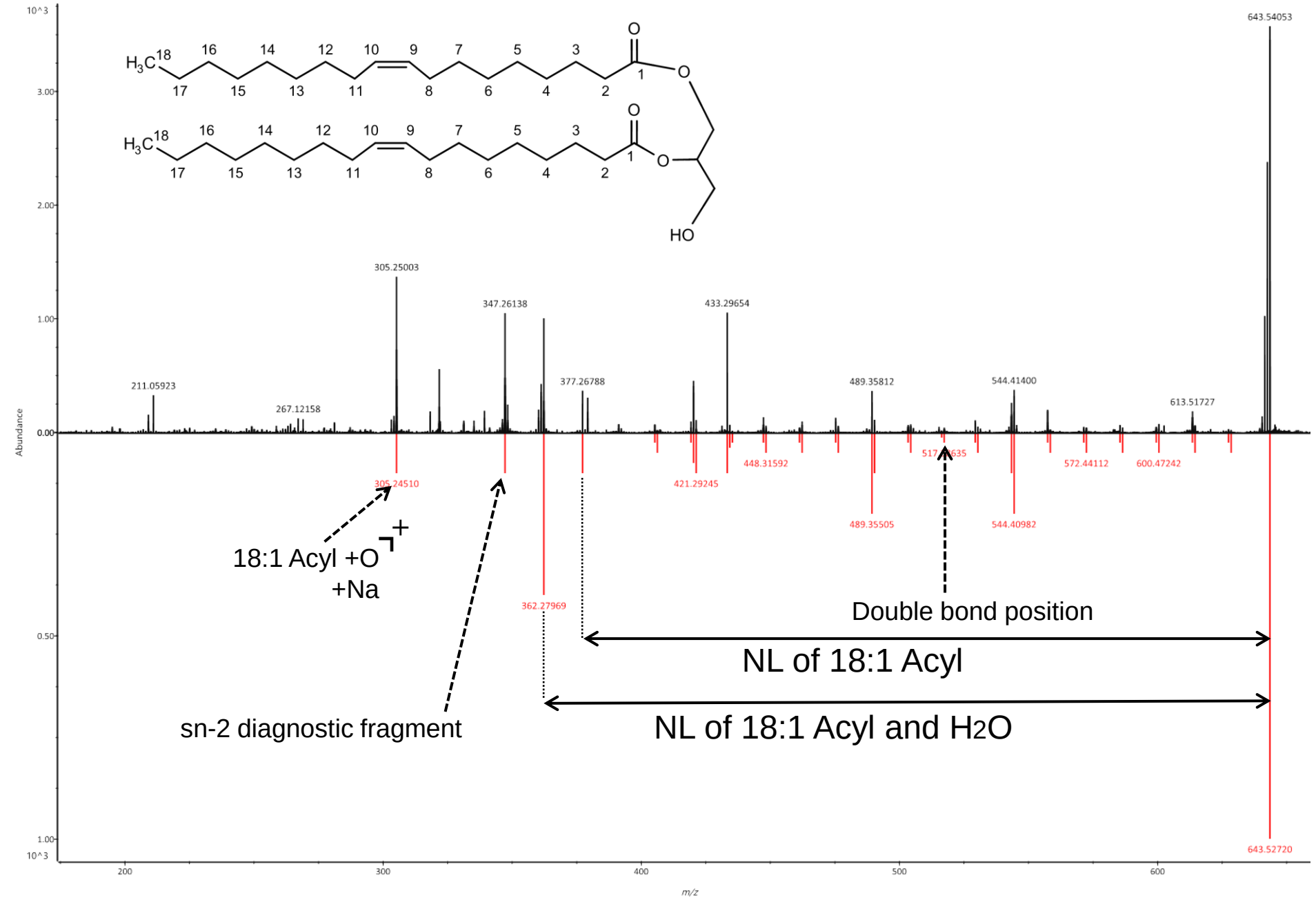
CL 16:0_18:1_16:0_18:1 as $[M+NH_4]^+$



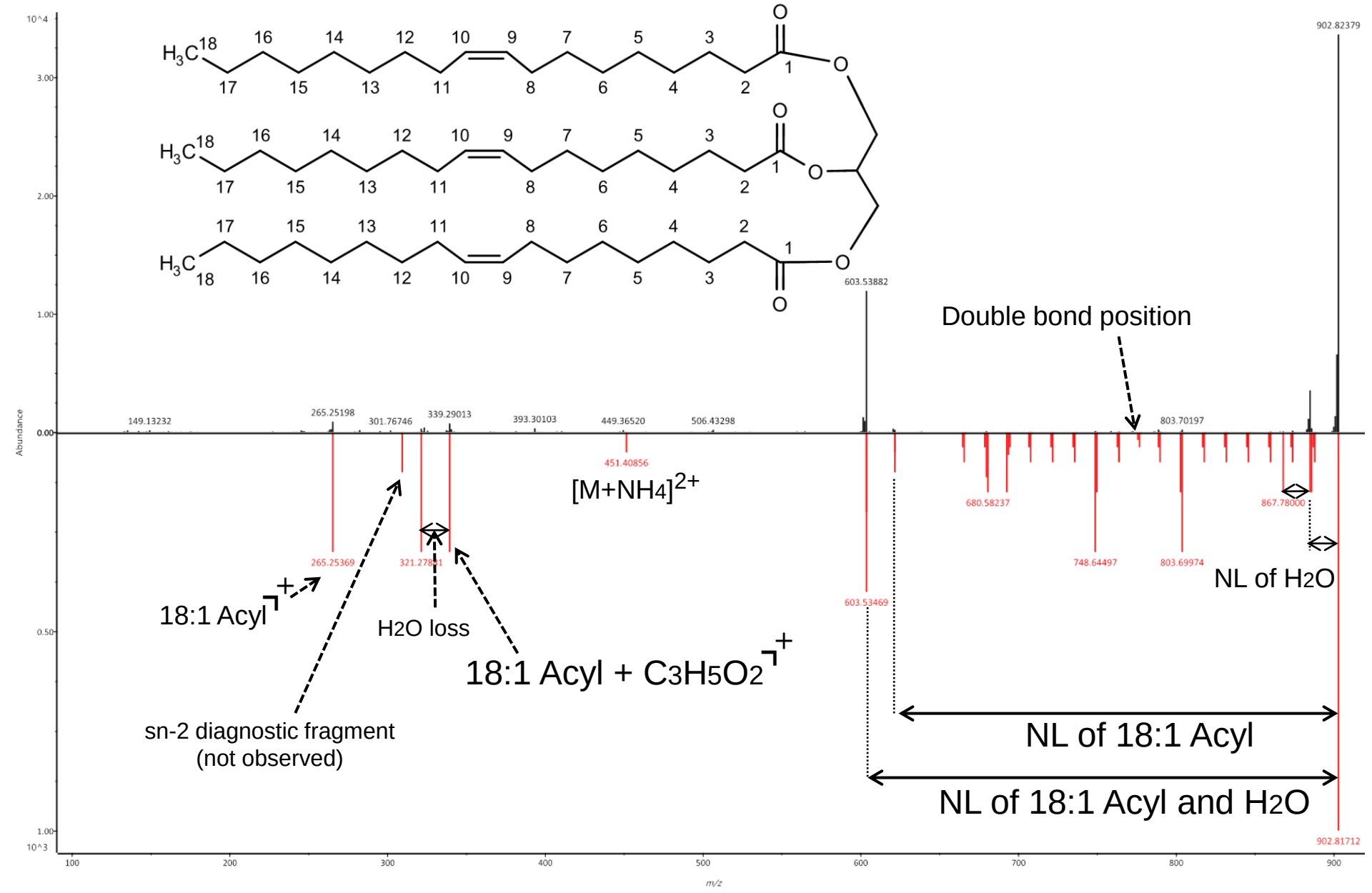
DG 18:1(9)_18:1(9) as [M+NH4]+



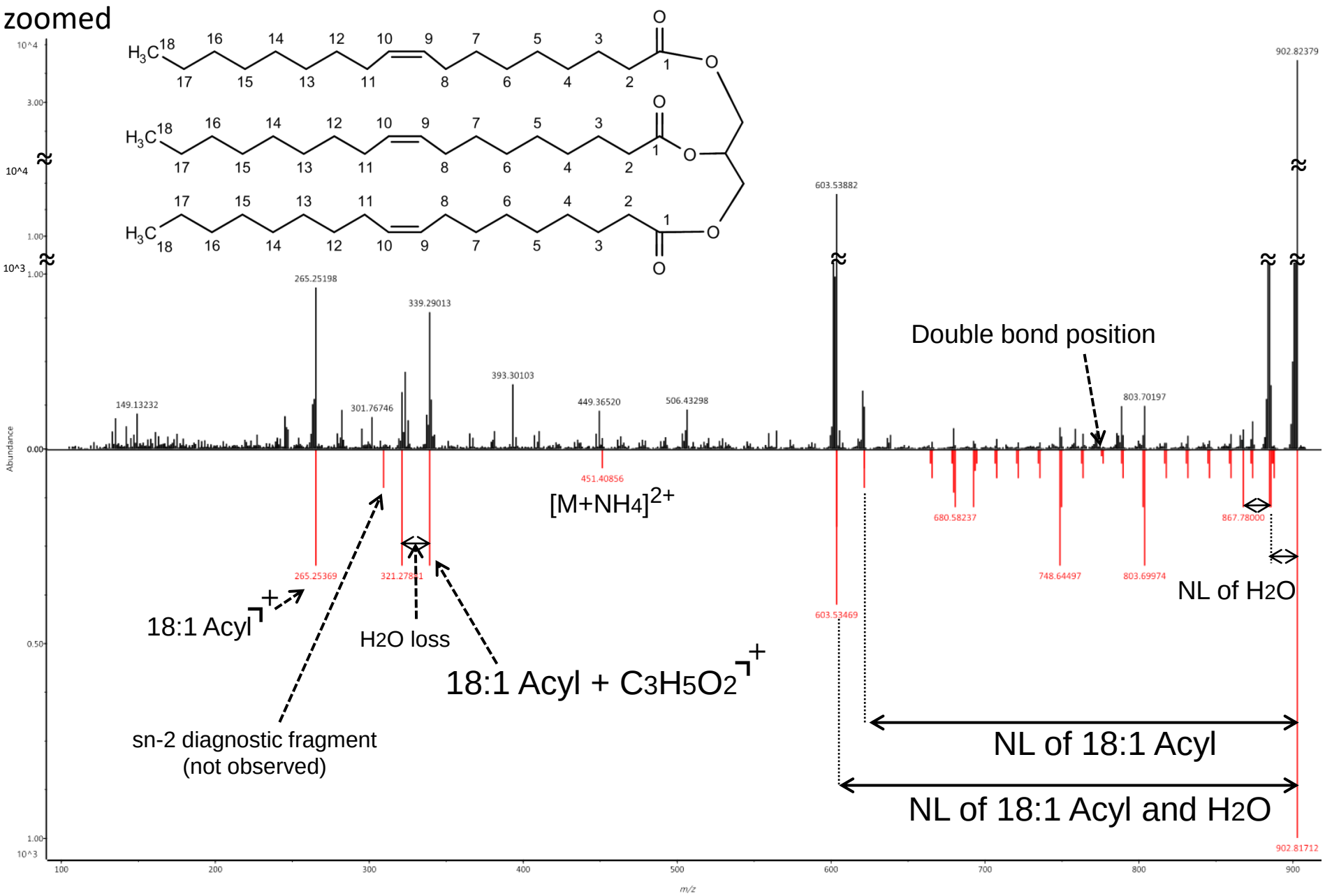
DG 18:1(9)/18:1(9) as [M+Na]+



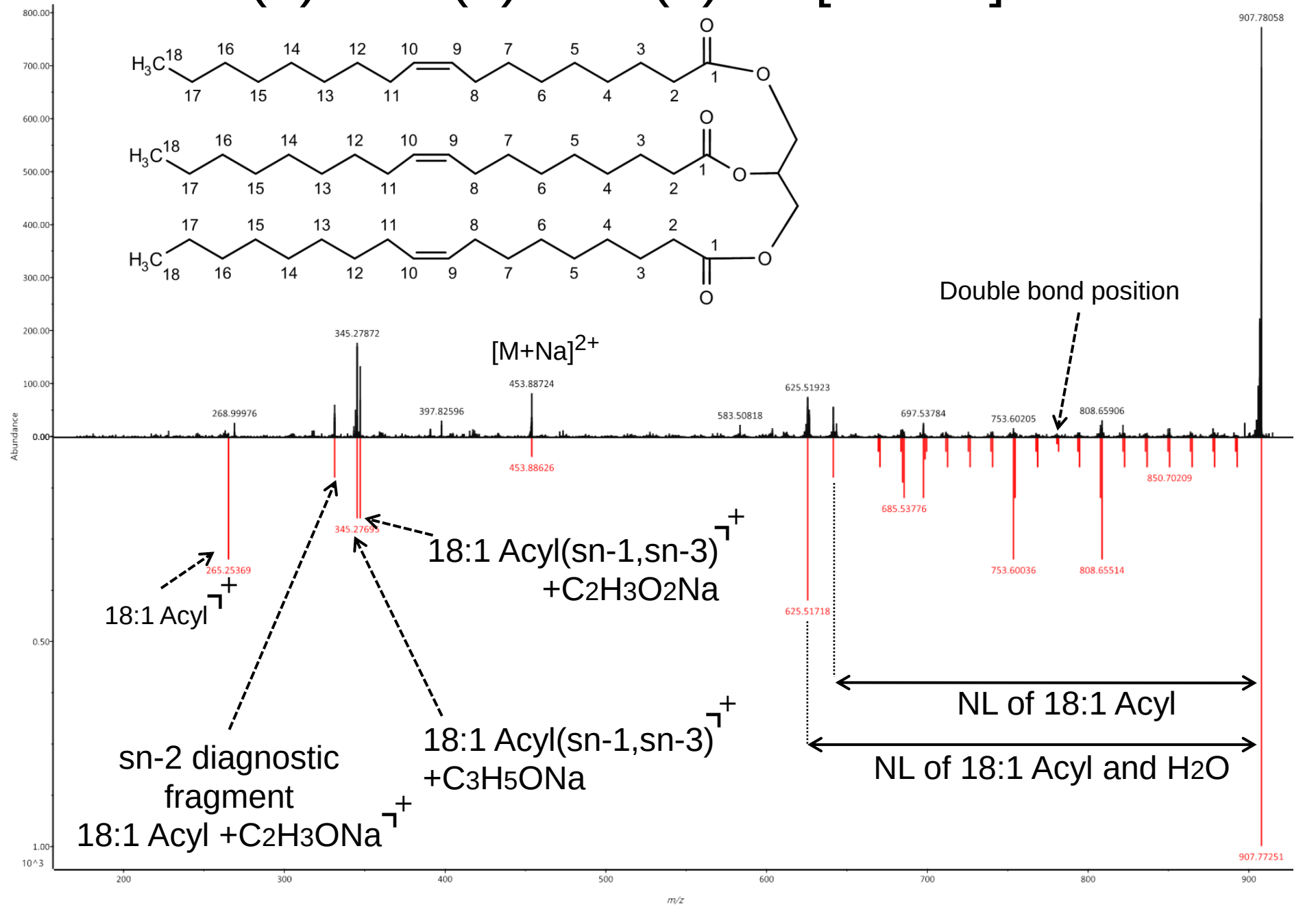
TG 18:1(9)_18:1(9)_18:1(9) as [M+NH4]+



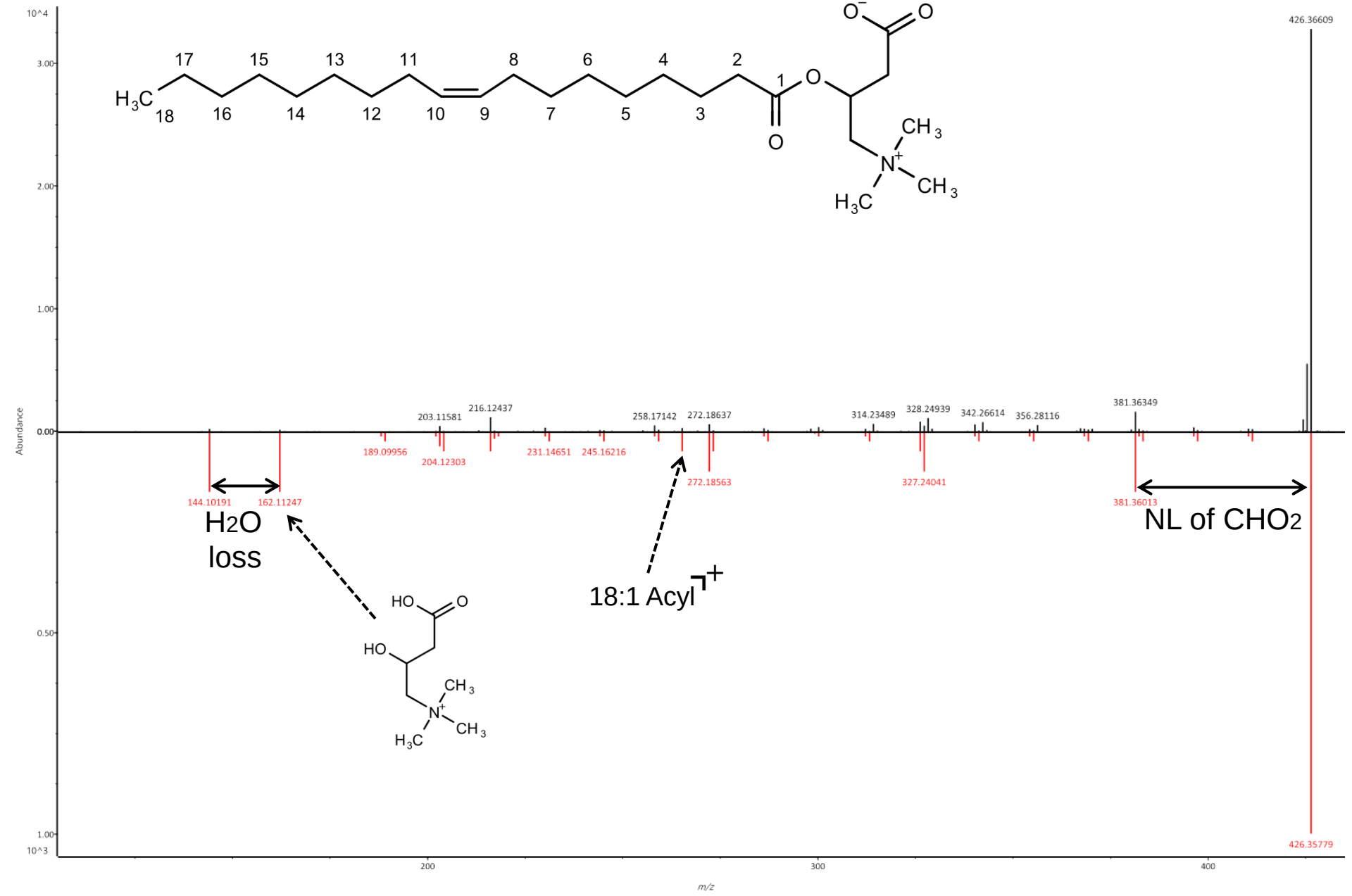
TG 18:1(9)_18:1(9)_18:1(9) as [M+NH4]+



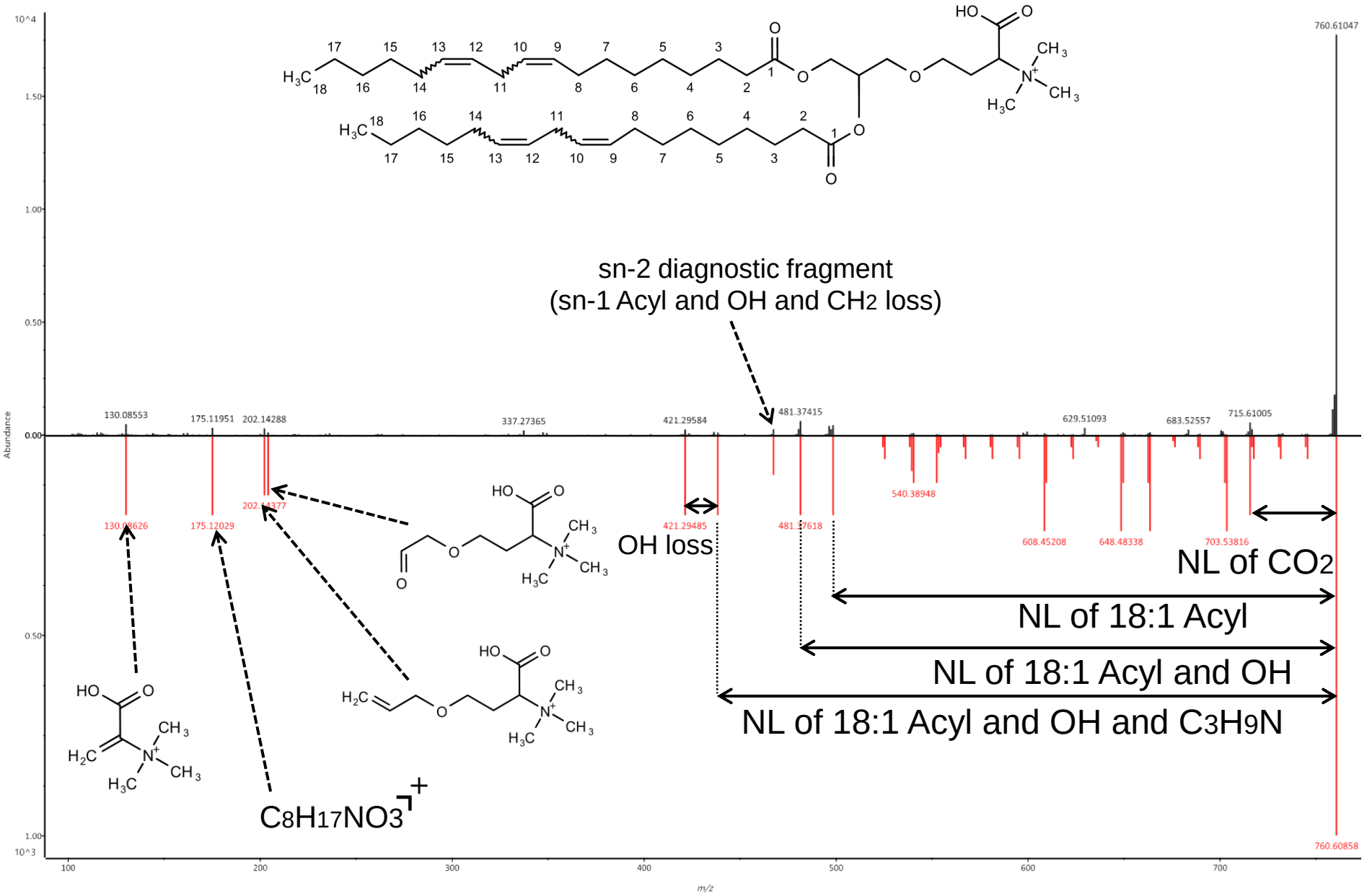
TG 18:1(9)/18:1(9)/18:1(9) as [M+Na]+



CAR 18:1(9) as [M+H]+

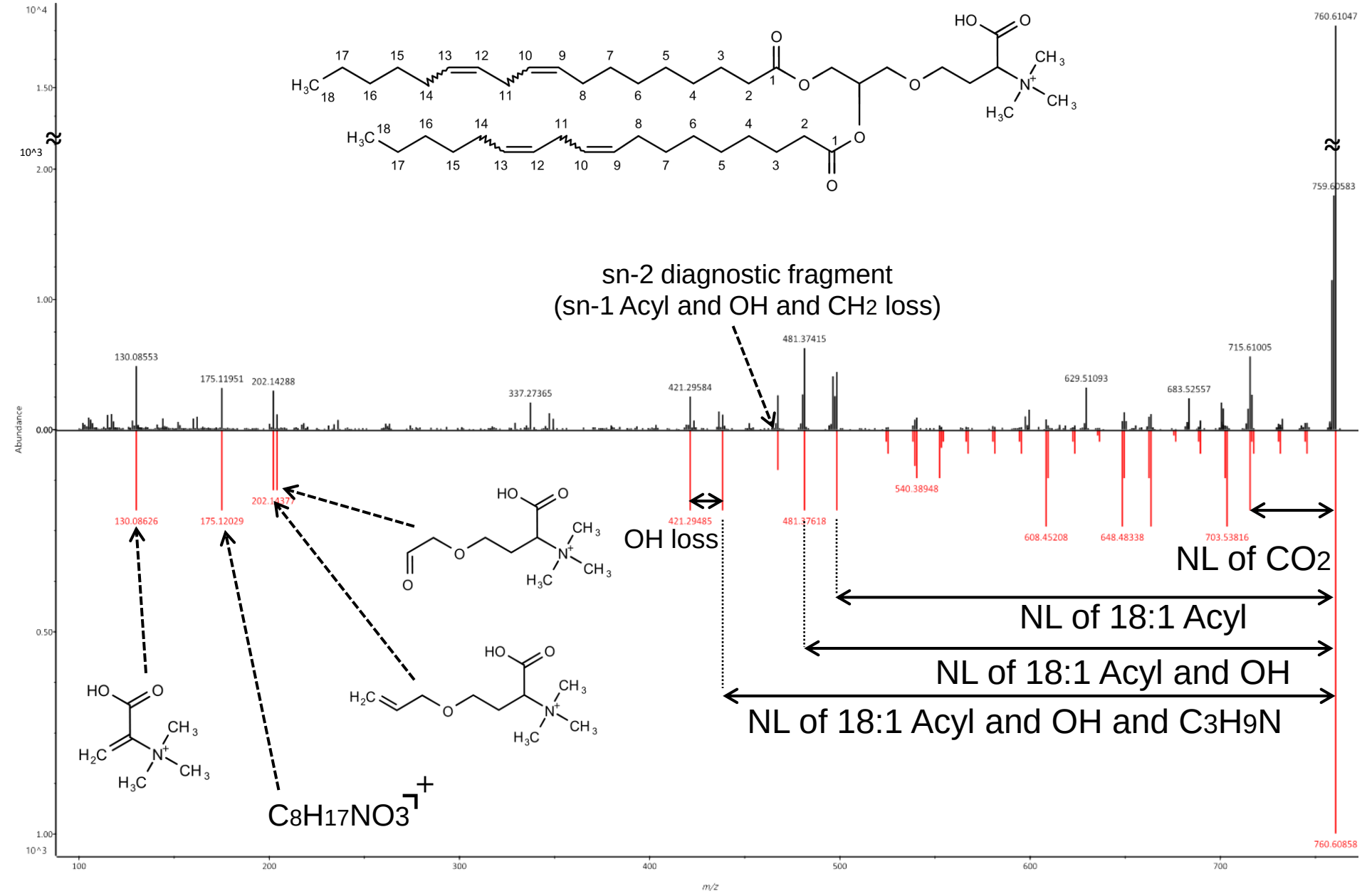


DGTS 18:2(9,12)/18:2(9,12) as [M+H]+

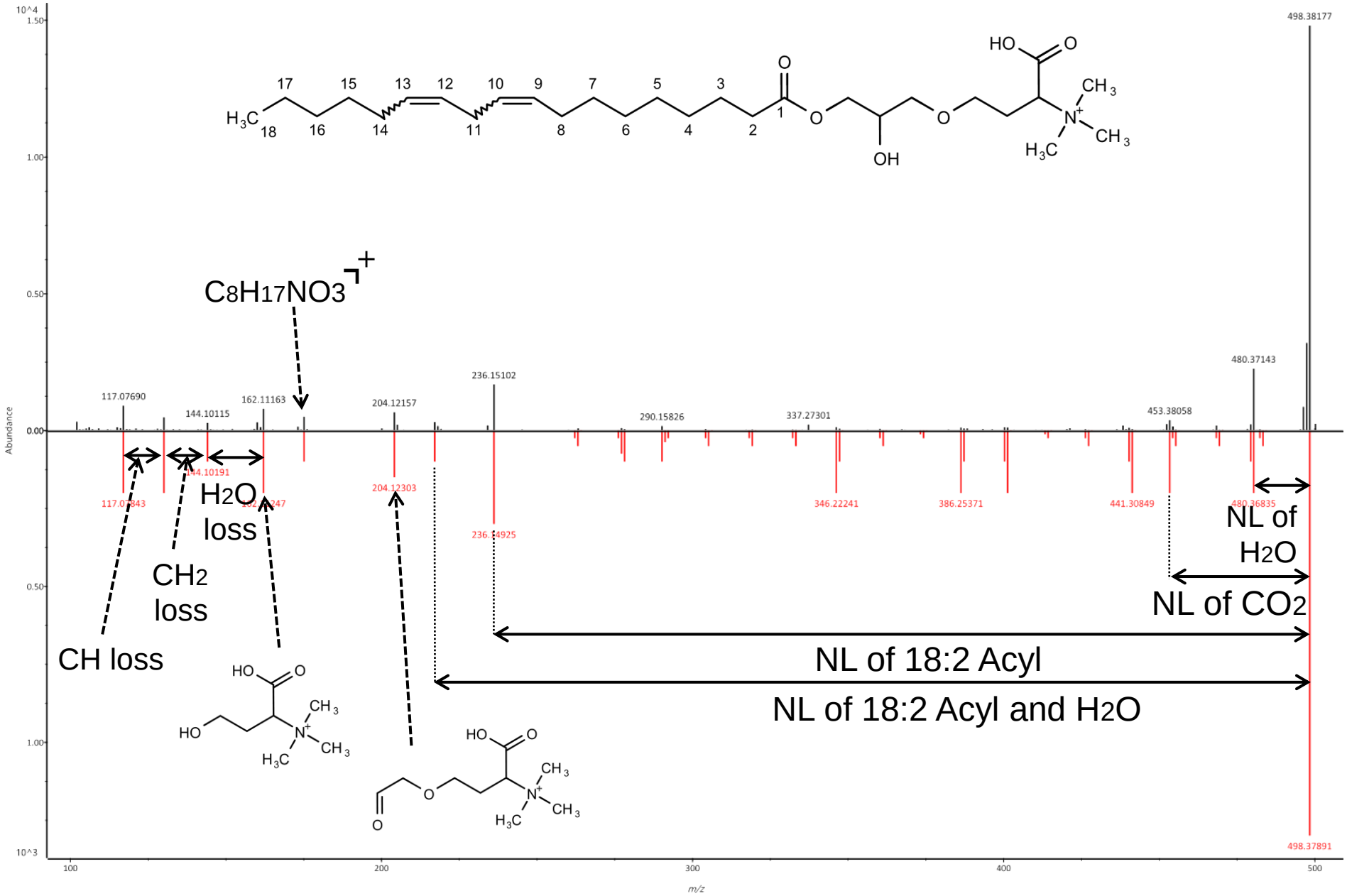


DGTS 18:2(9,12)/18:2(9,12) as [M+H]+

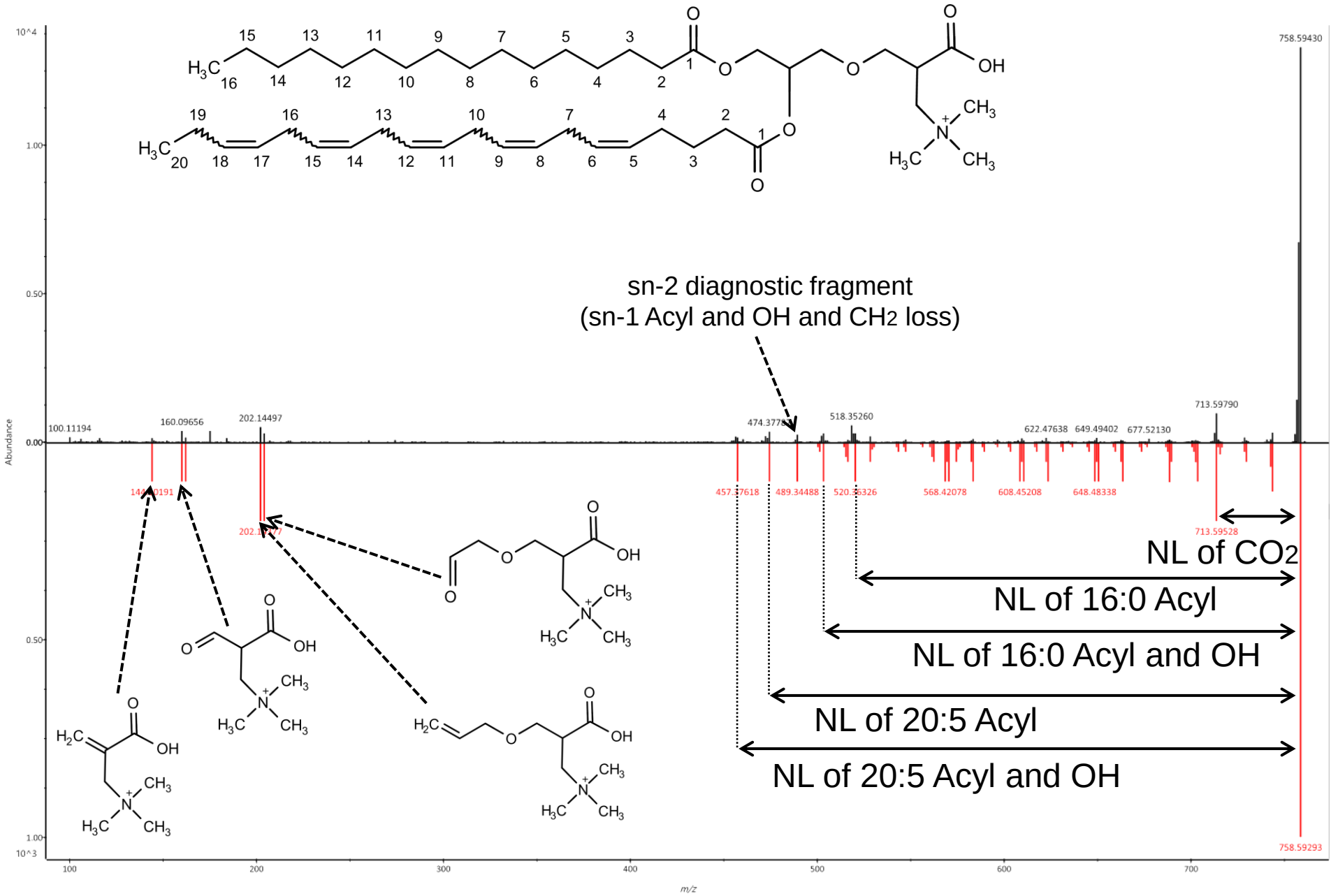
zoomed



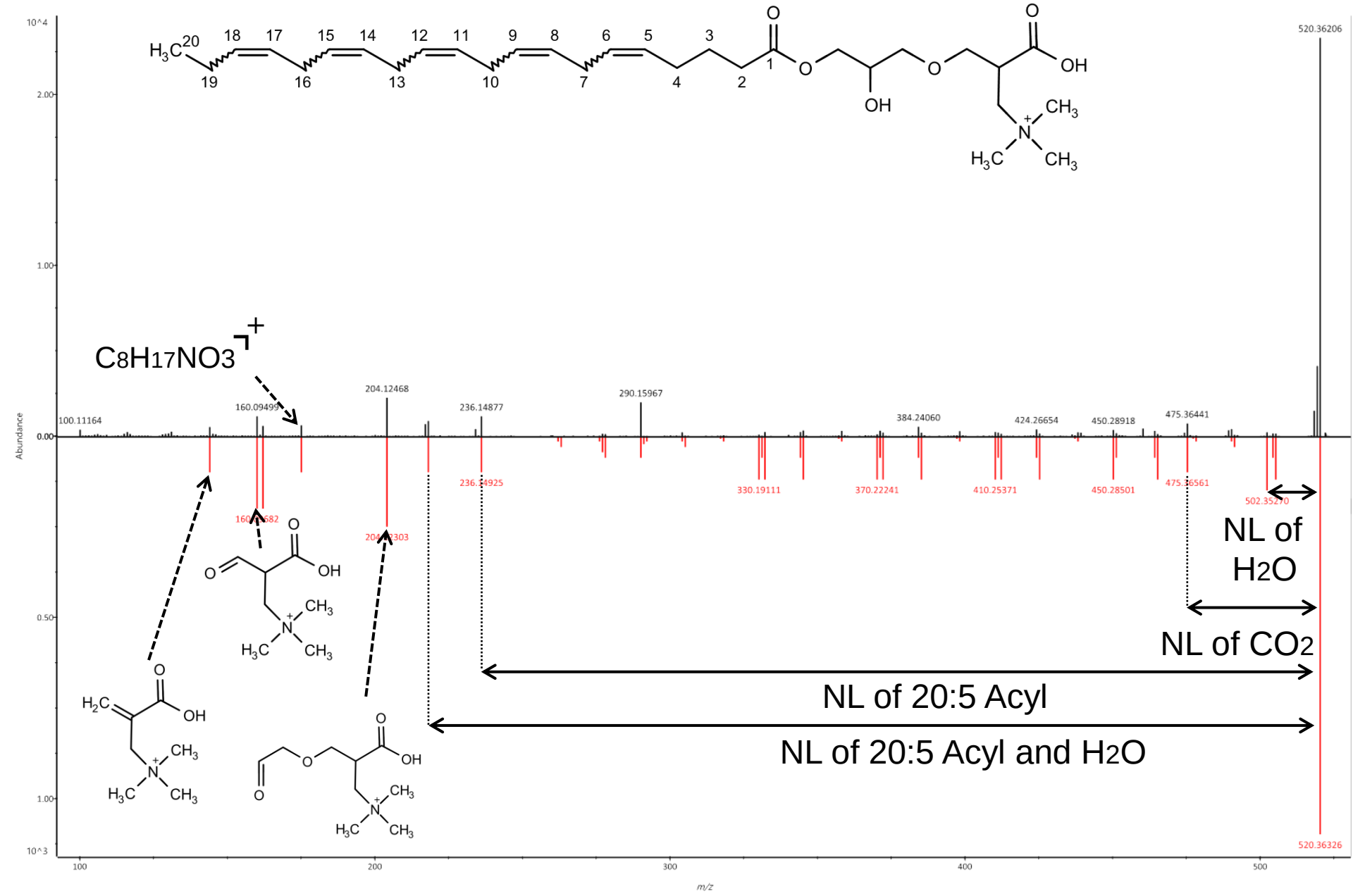
LDGTS 18:2(9,12) as [M+H]+



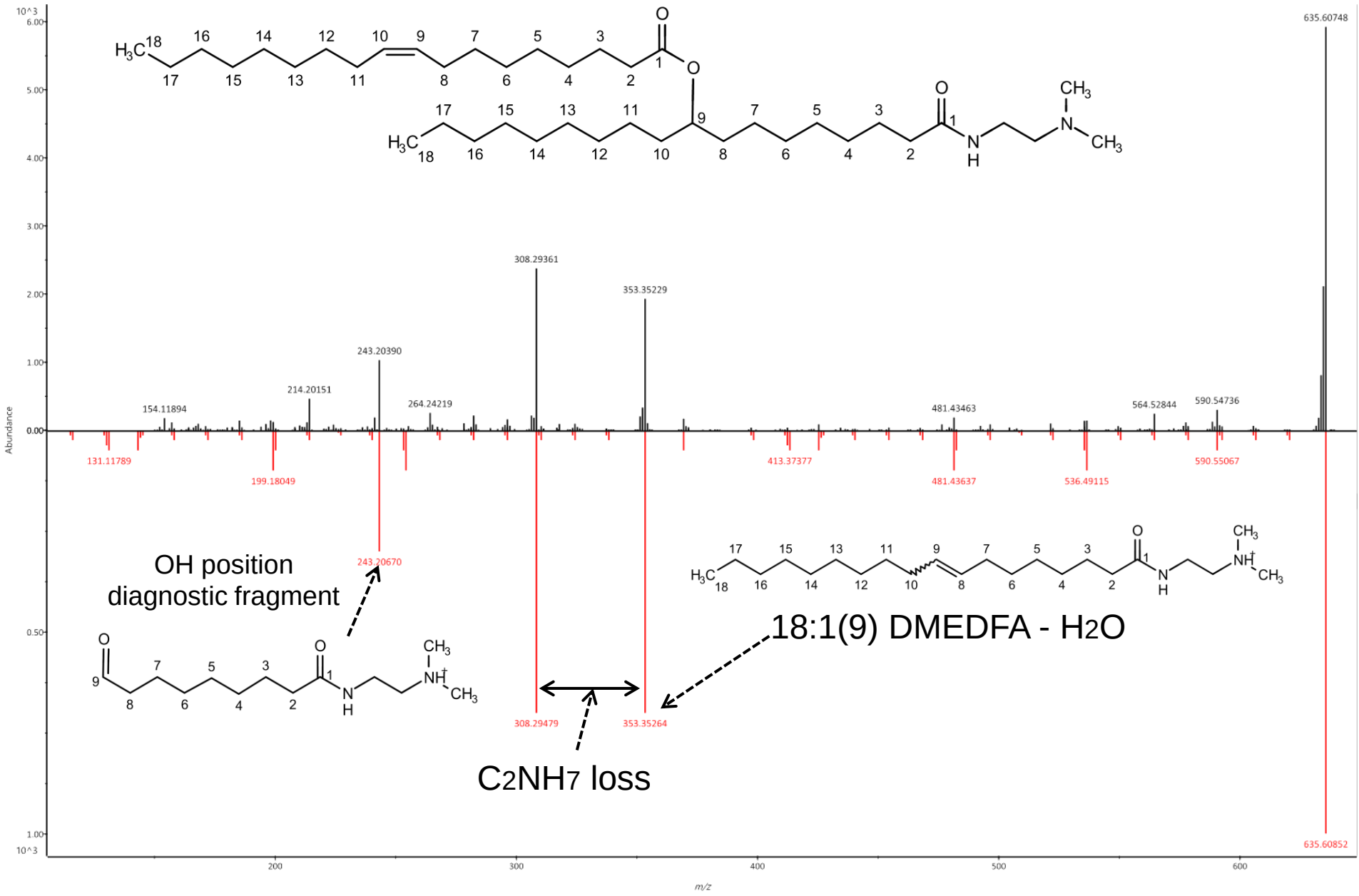
DGTA 16:0/20:5 as [M+H]+



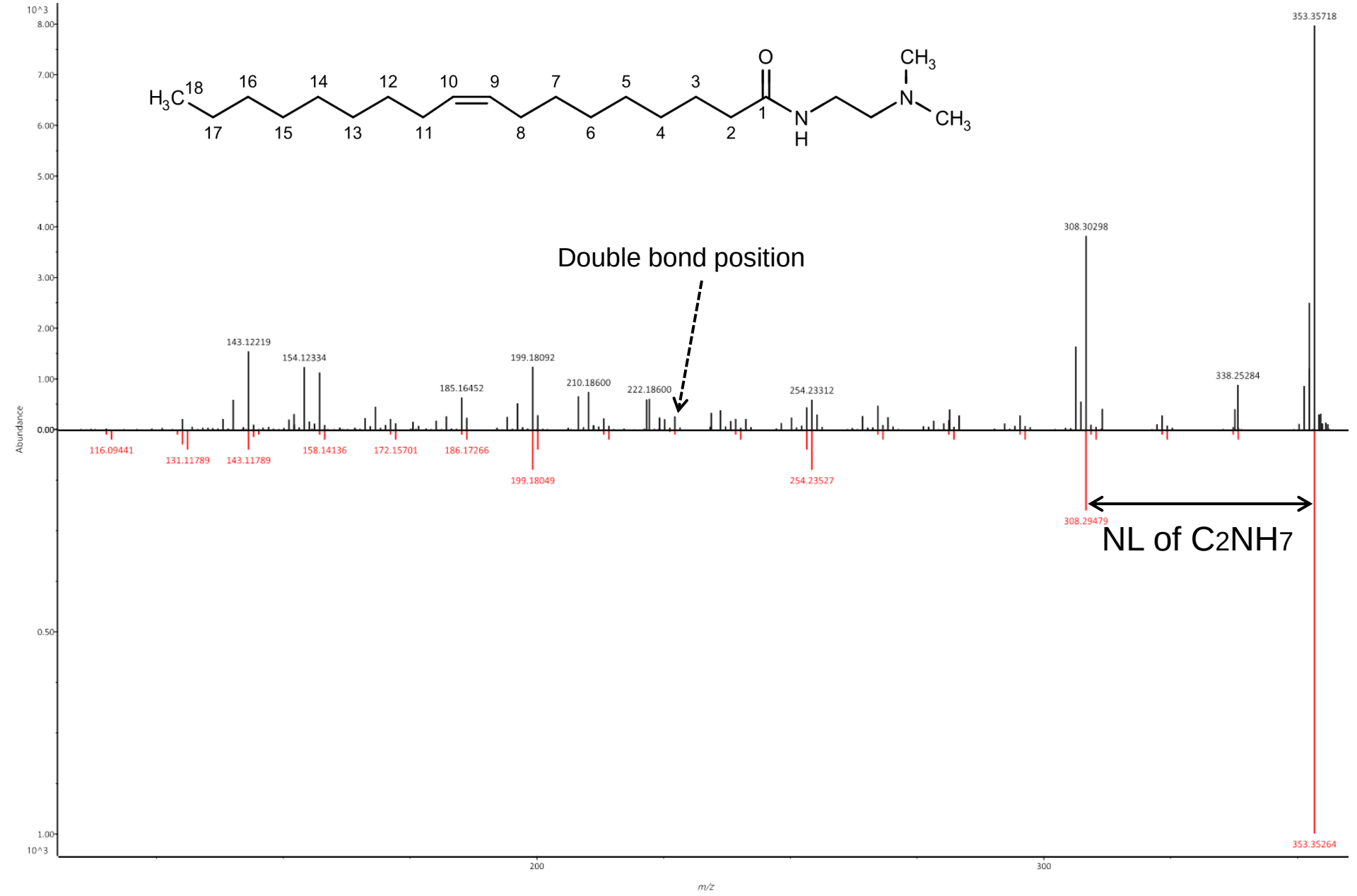
LDGTA 20:5 as [M+H]+



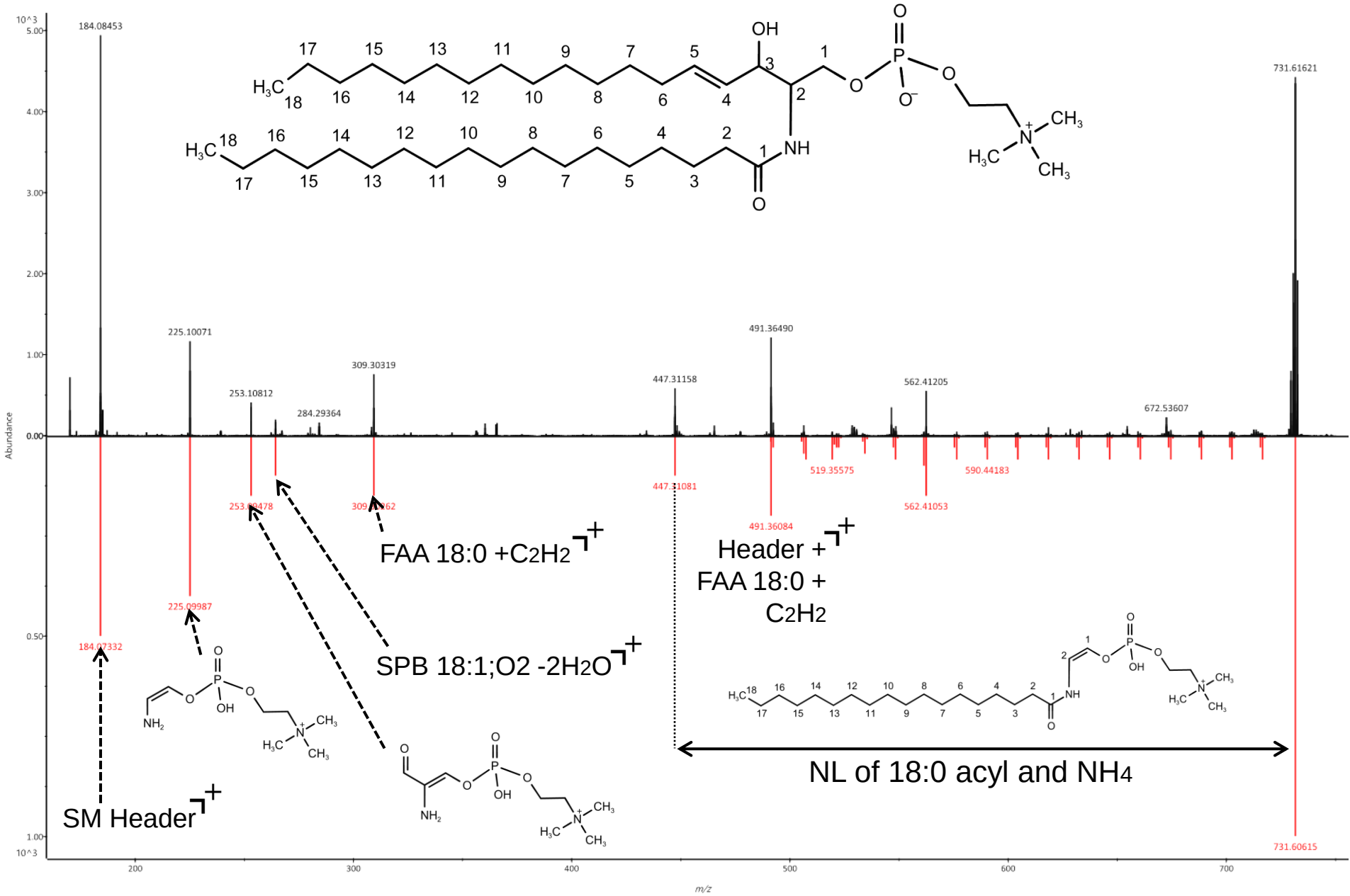
DMEDFAHFA 18:1(9)/18:0(9OH) as [M+H]+



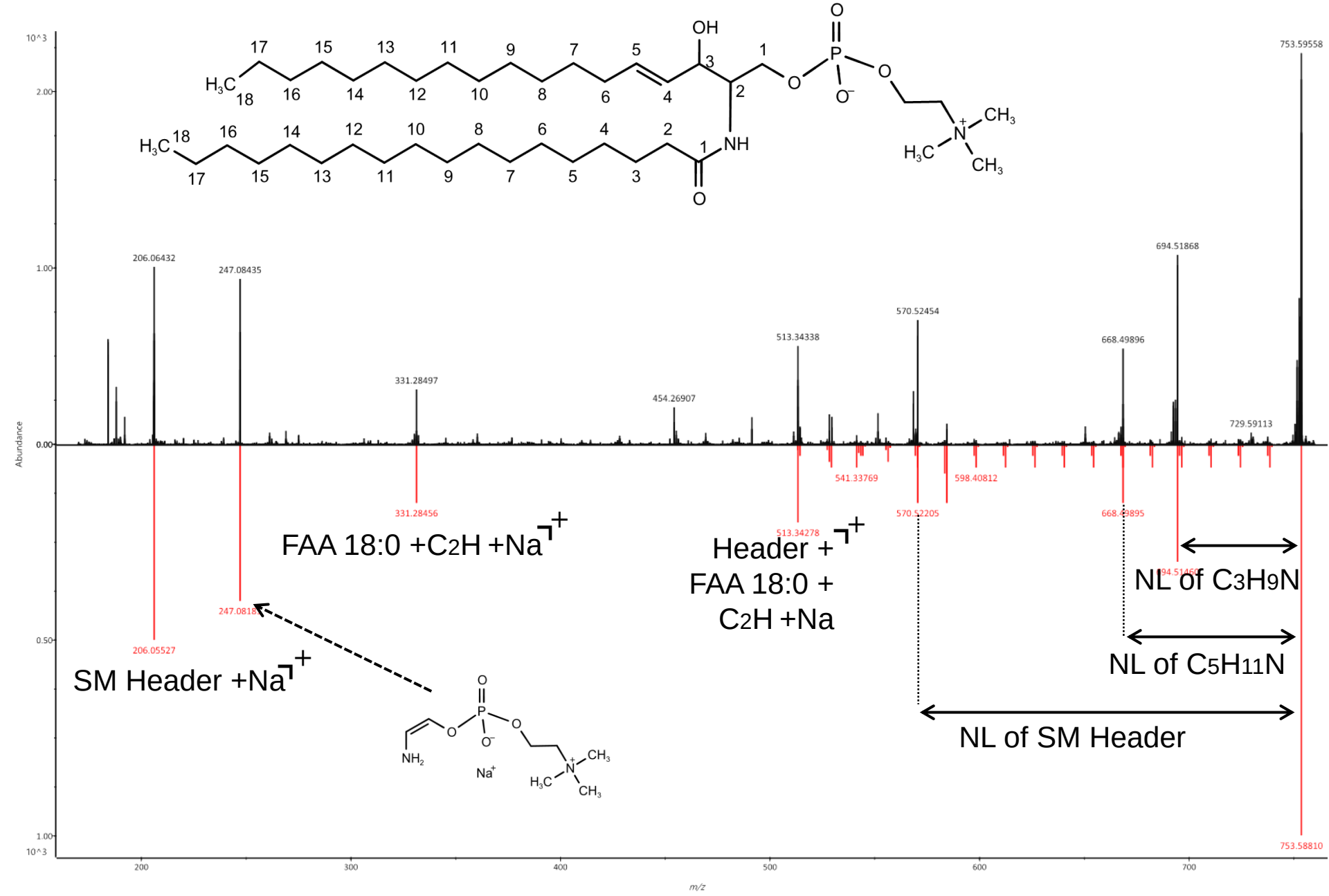
DMEDFA 18:1(9) as [M+H]+



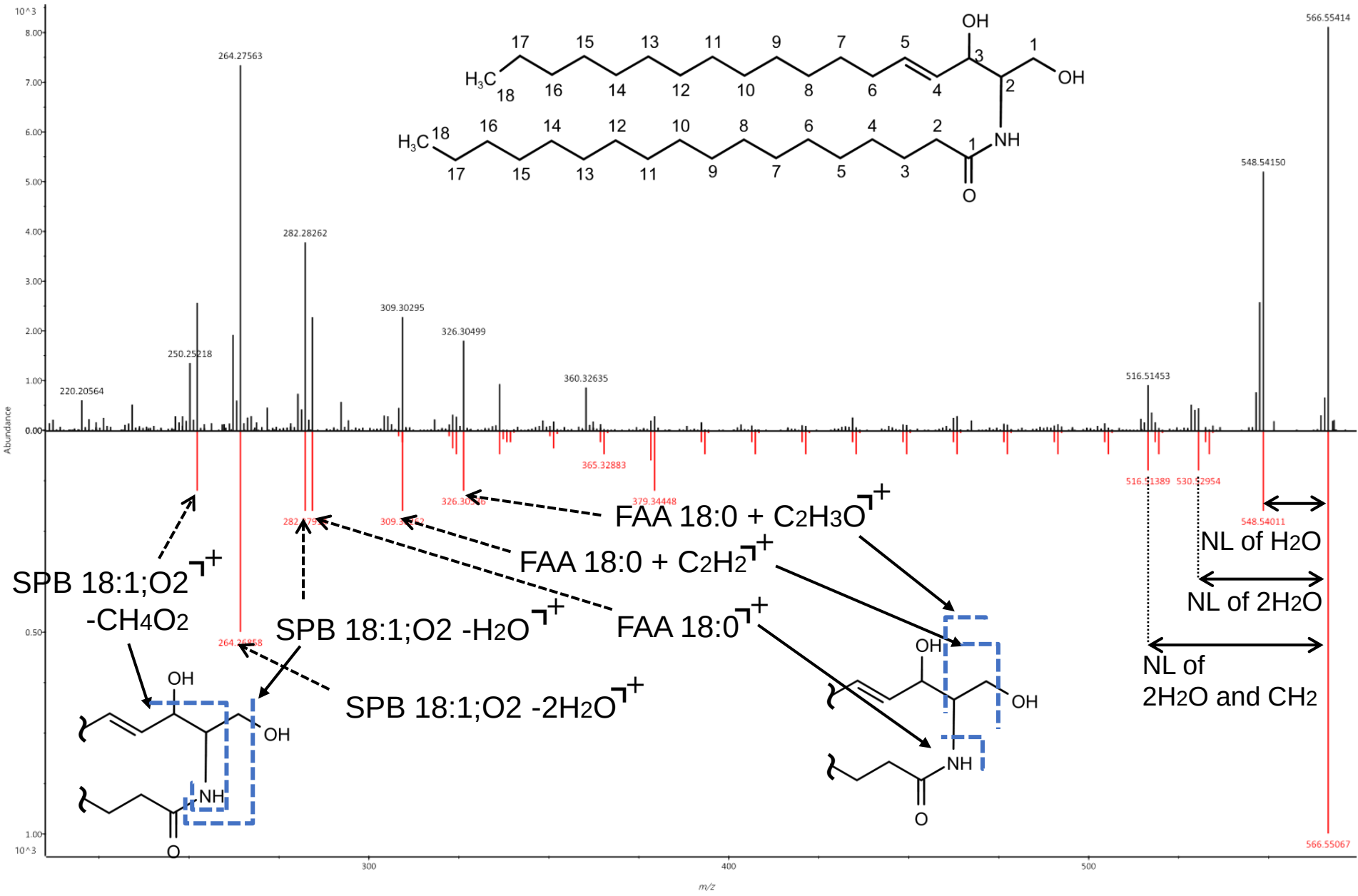
SM 18:1(4)(1OH,3OH)/18:0 as [M+H]+



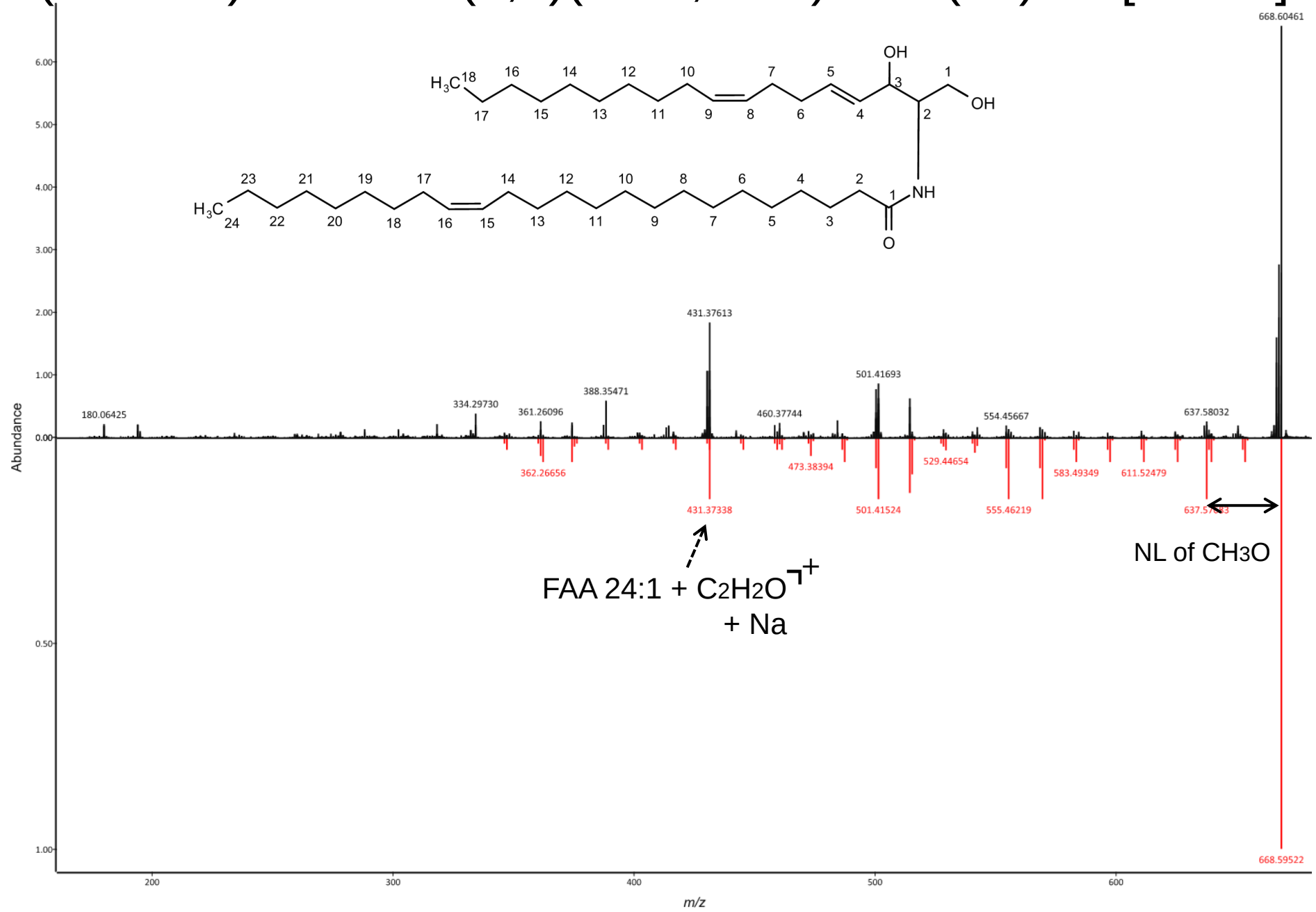
SM 18:1(4)(1OH,3OH)/18:0 as [M+Na]+



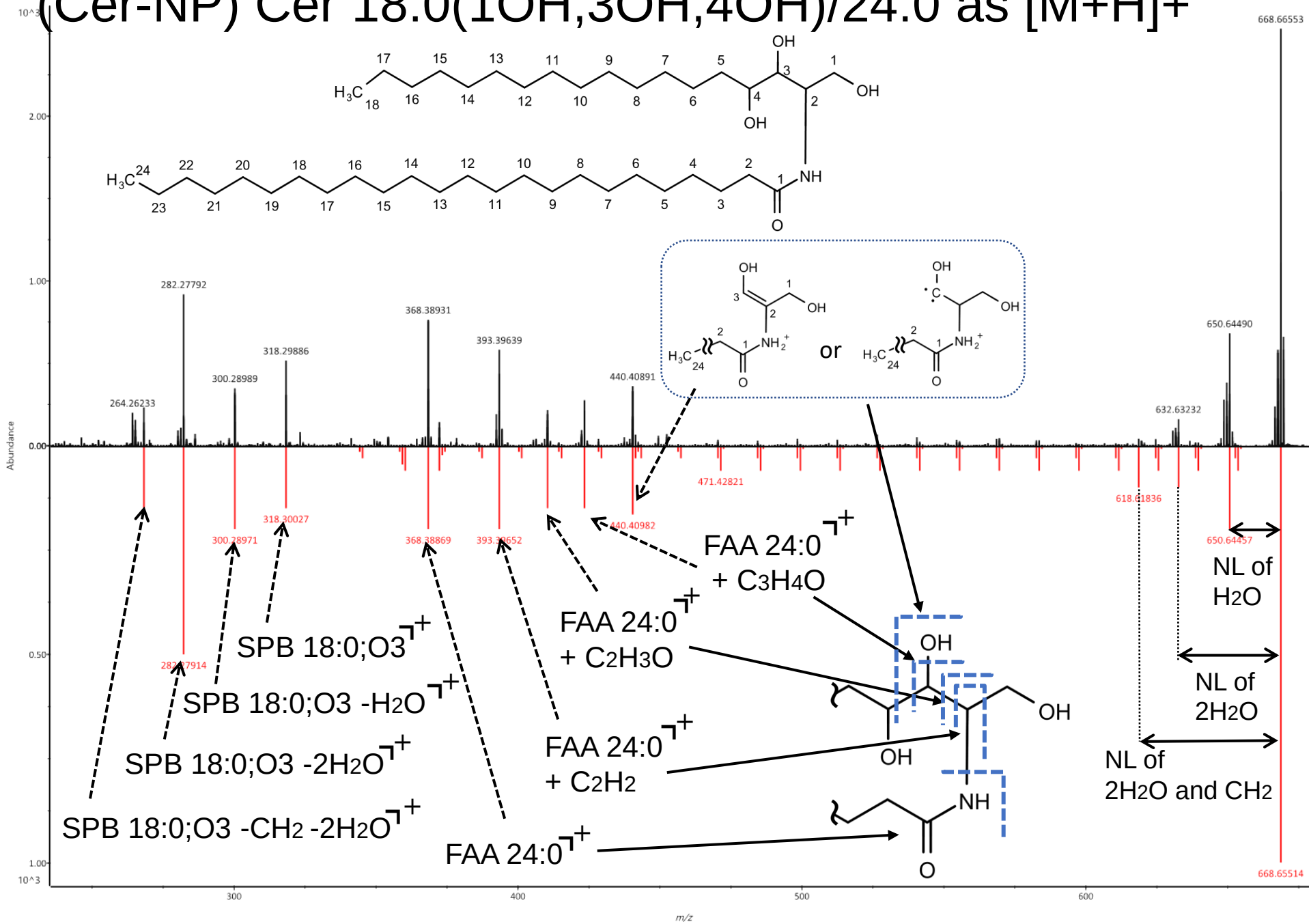
(Cer-NS) Cer 18:1(4)(1OH,3OH)/18:0 as [M+H]+



(Cer-NS) Cer 18:2(4,8)(1OH,3OH)/24:1(15) as [M+Na]+

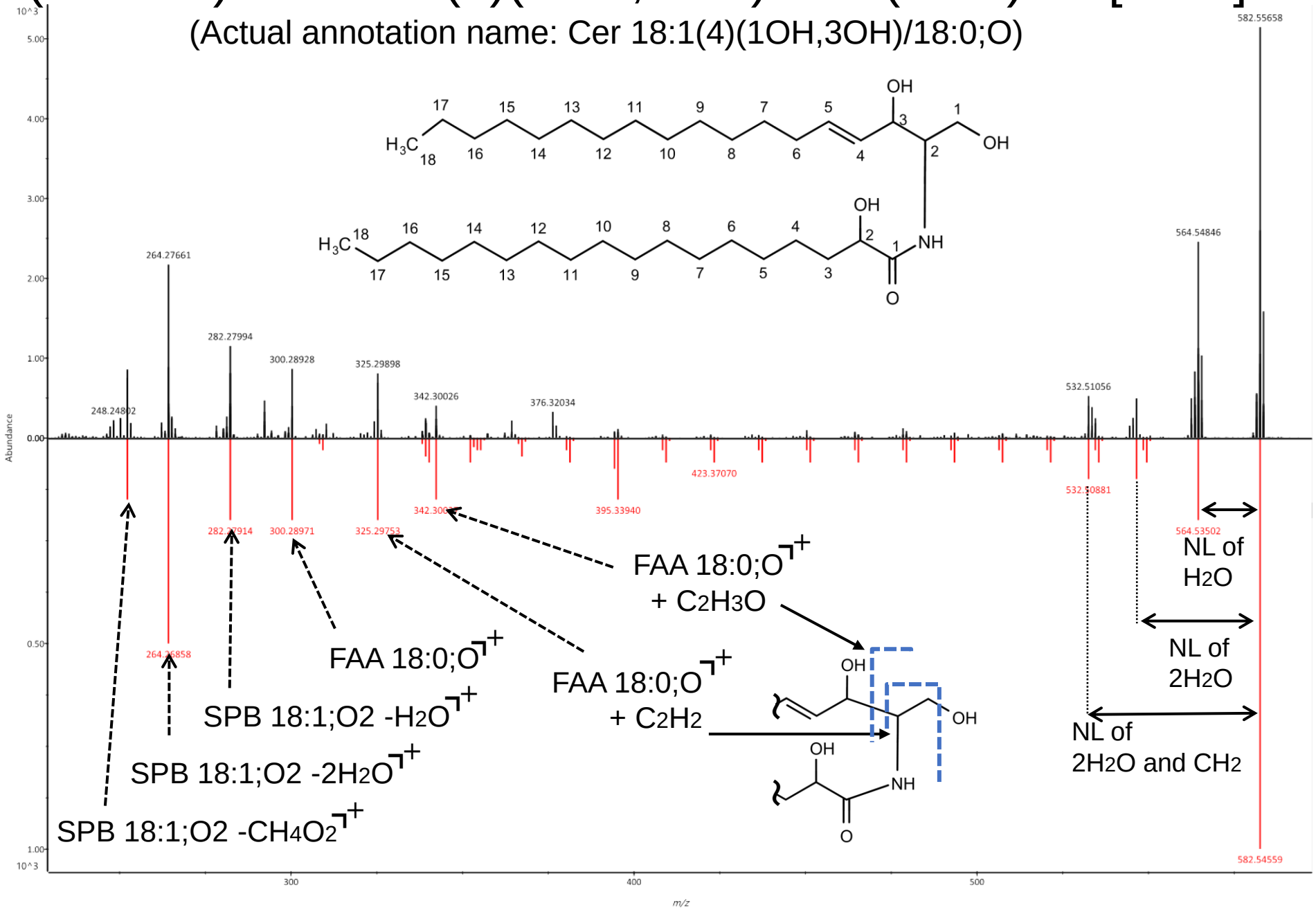


(Cer-NP) Cer 18:0(1OH,3OH,4OH)/24:0 as [M+H]+



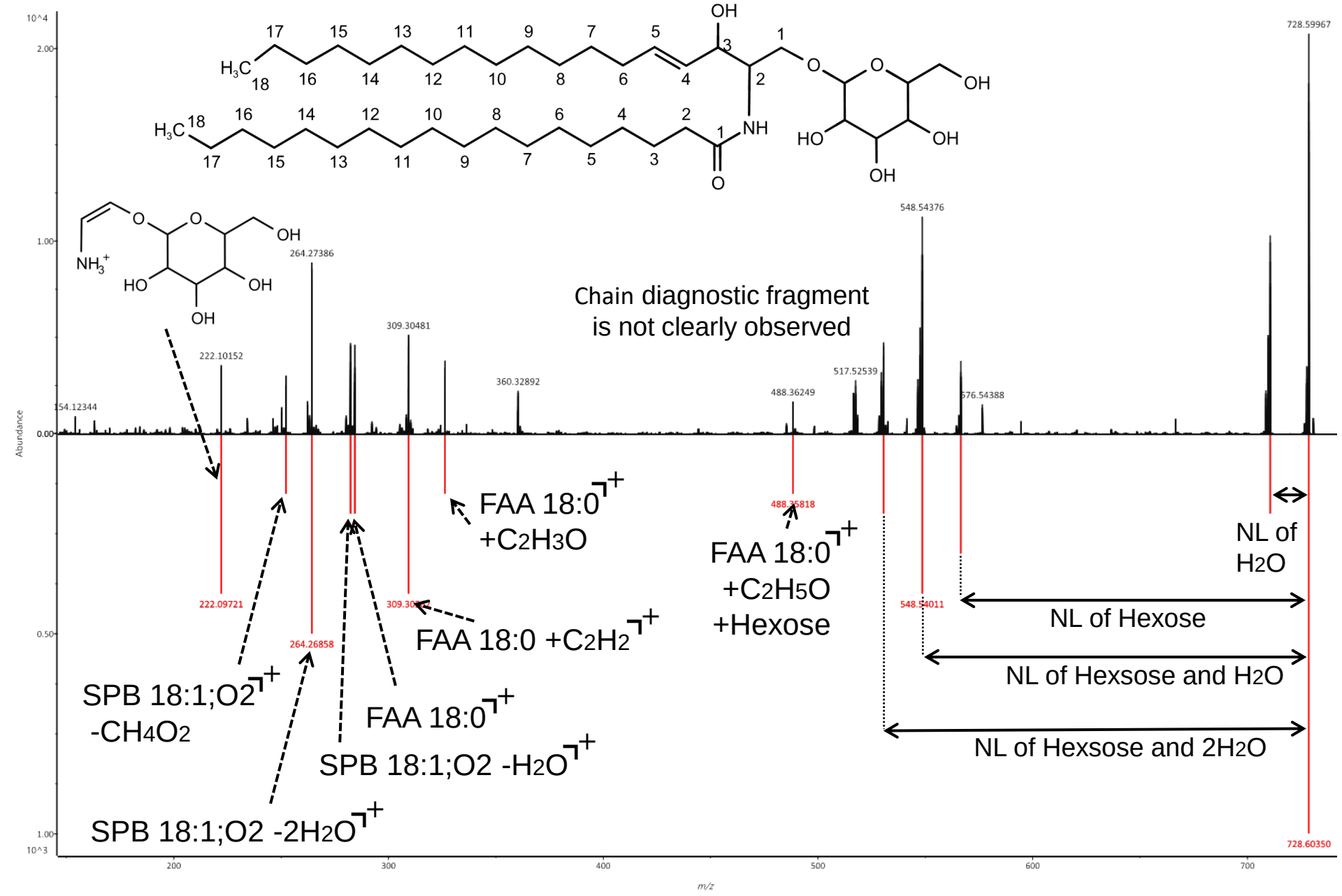
(Cer-AS) Cer 18:1(4)(1OH,3OH)/18:0(2OH) as [M+H]+

(Actual annotation name: Cer 18:1(4)(1OH,3OH)/18:0;O)



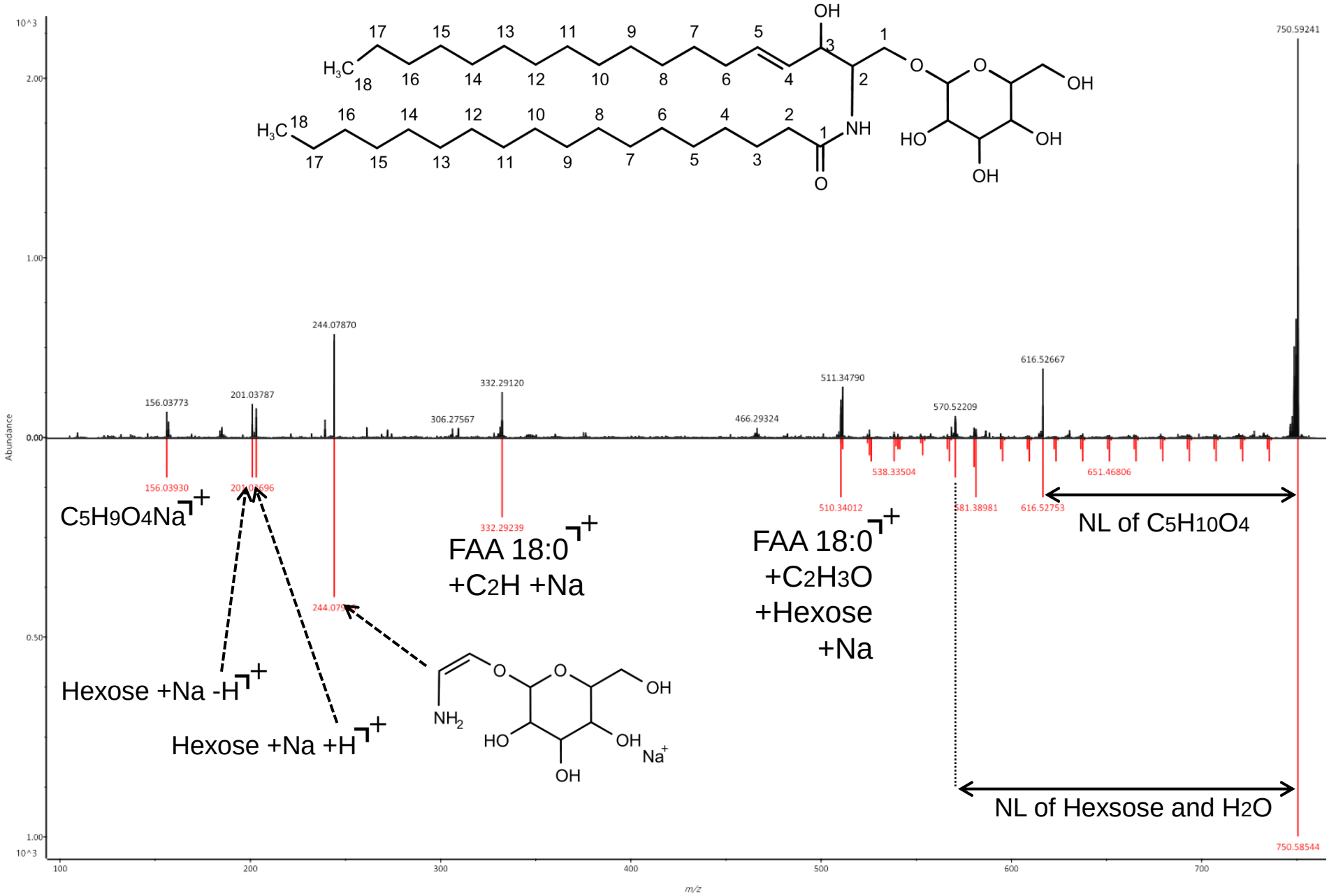
GalCer

(HexCer-NS) HexCer 18:1(4)(1OH,3OH)/18:0 as [M+H]+

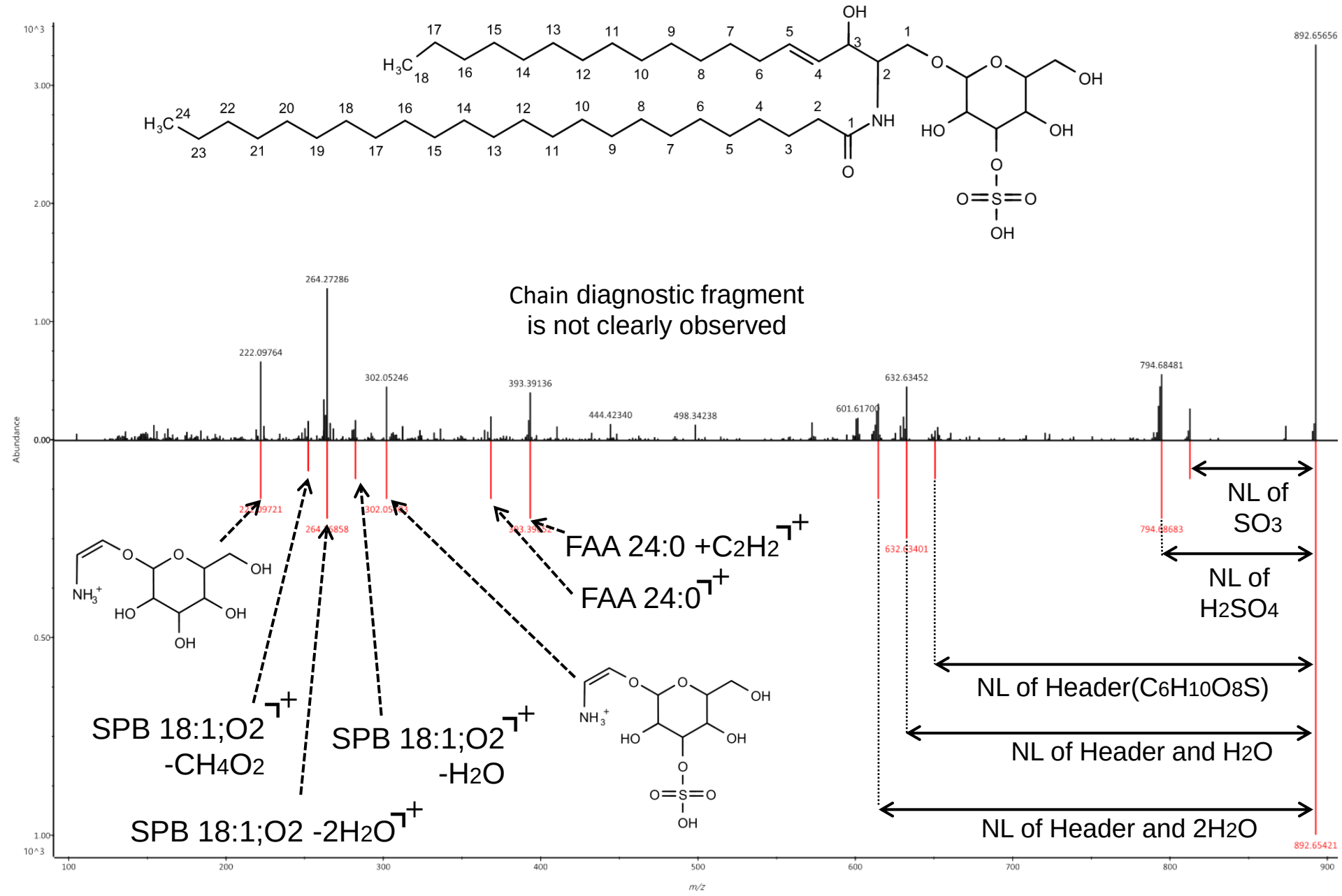


GalCer

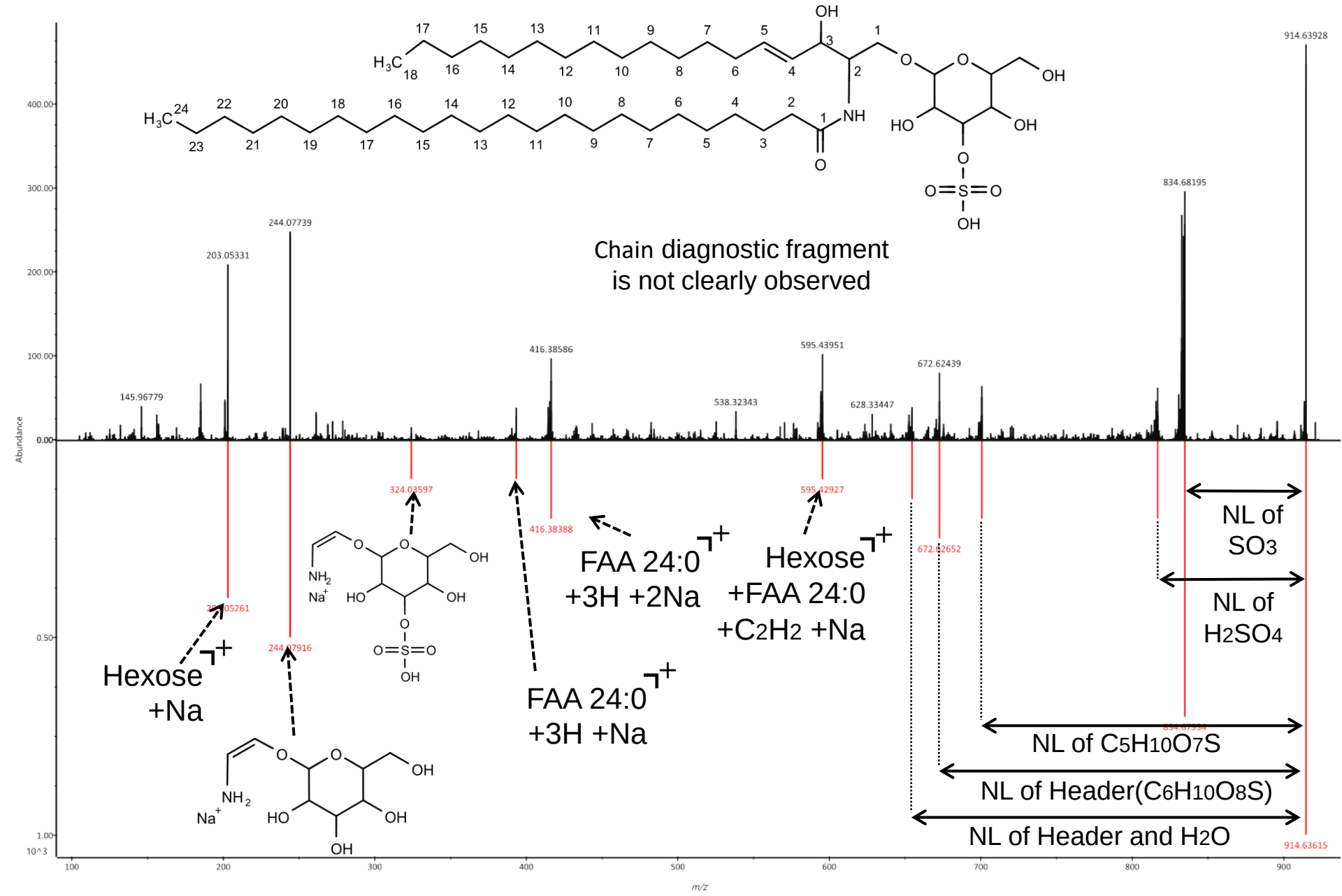
(HexCer-NS) HexCer 18:1(4)(1OH,3OH)/18:0 as [M+Na]+



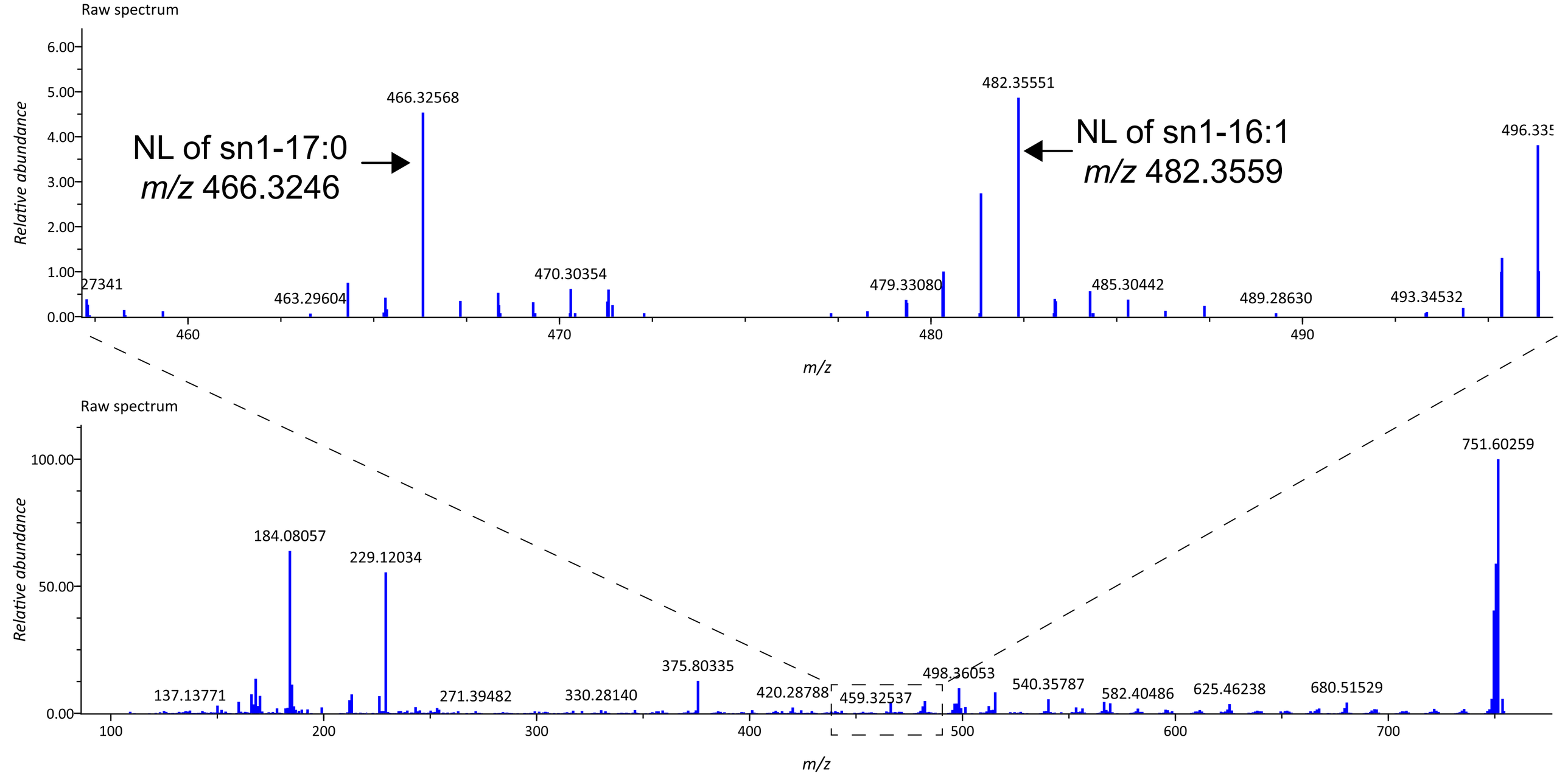
SHexCer 18:1(4)(1OH,3OH)/24:0 as [M+H]+



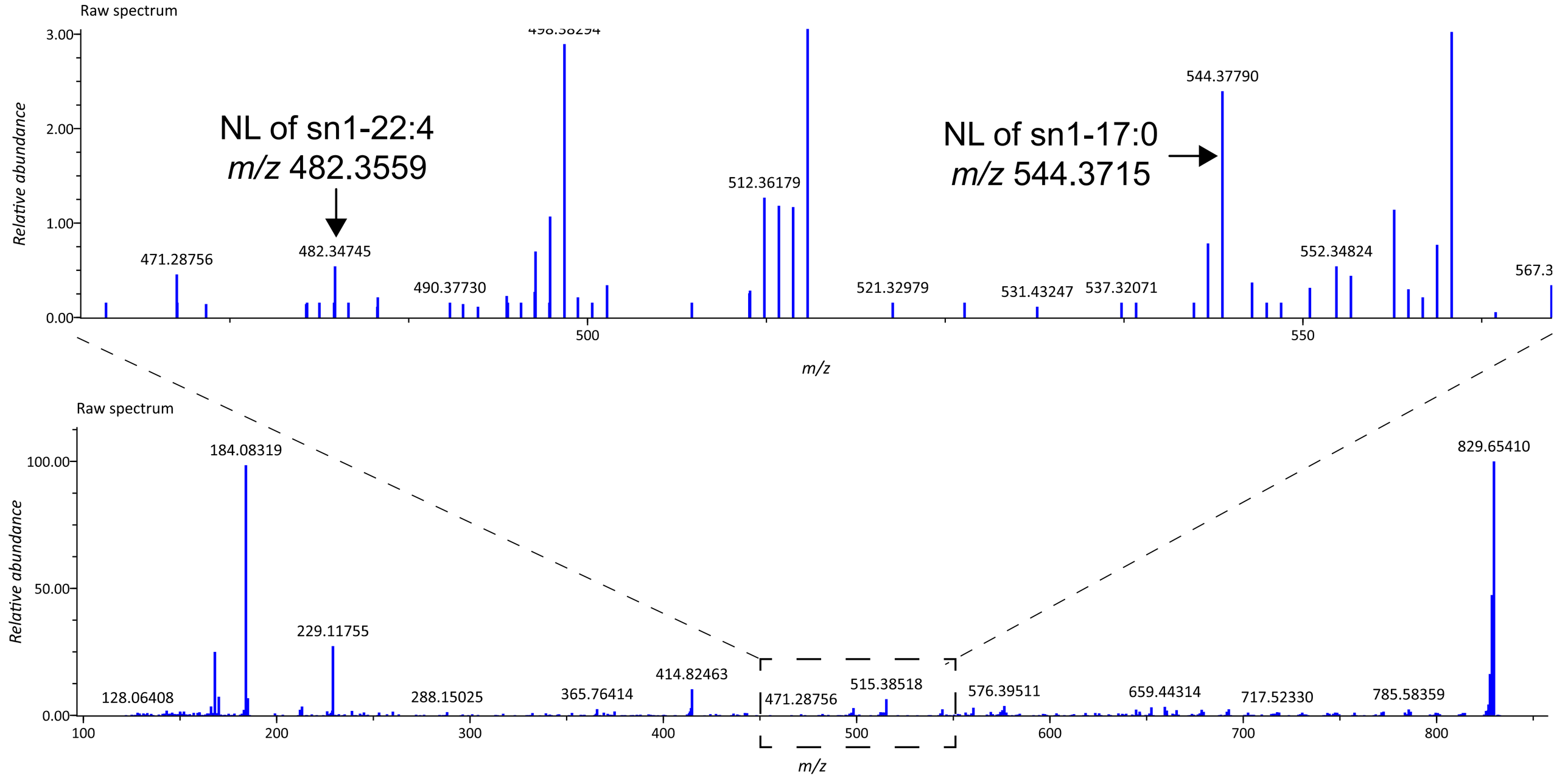
SHexCer 18:1(4)(1OH,3OH)/24:0 as [M+Na]+



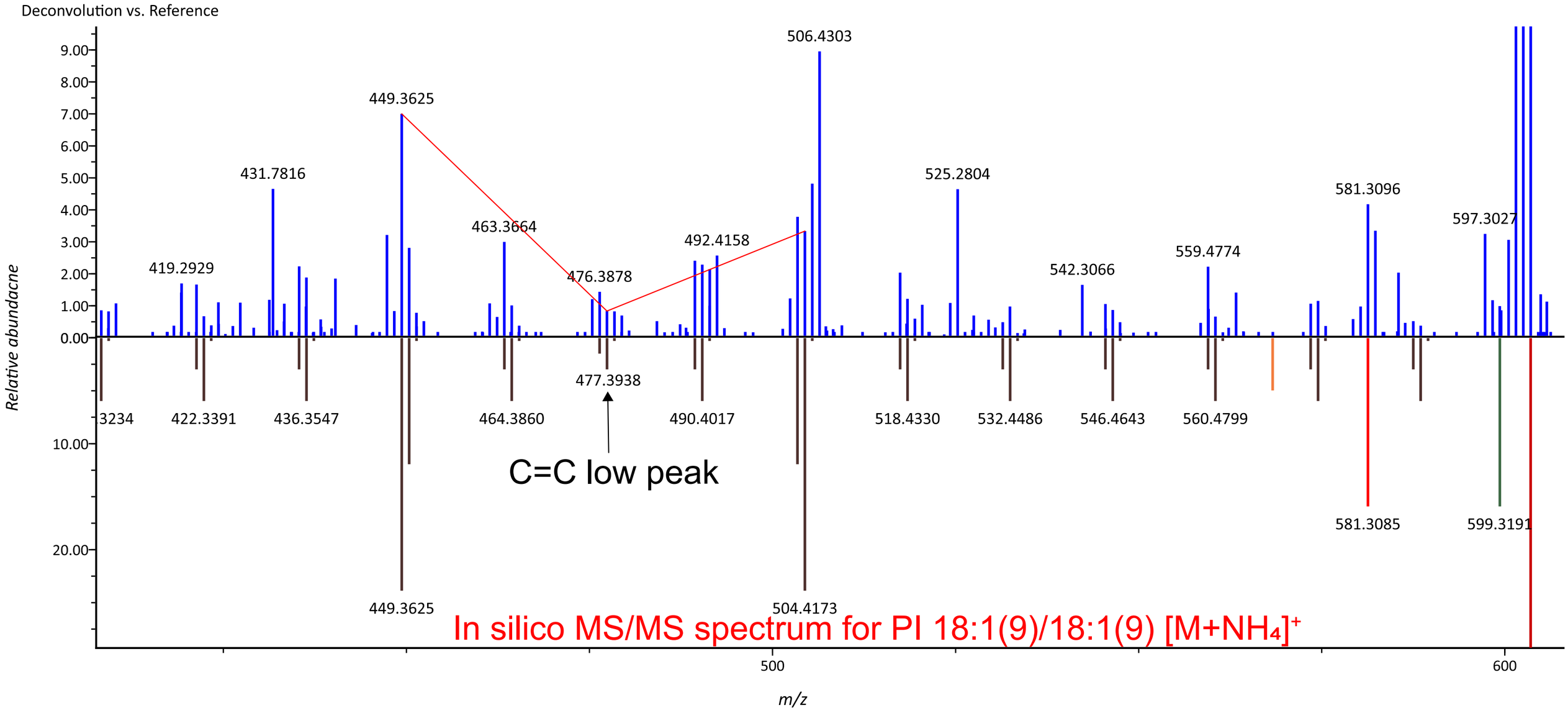
a The original annotation was PC-d5 16:1(9)/17:0, while the correct annotation shold be PC-d5 17:0/16:1(9)



b The annotation of PC-d5 17:0/22:4(7,10,13,16) was correct, but sn1-16:1 exists in the spectrum.

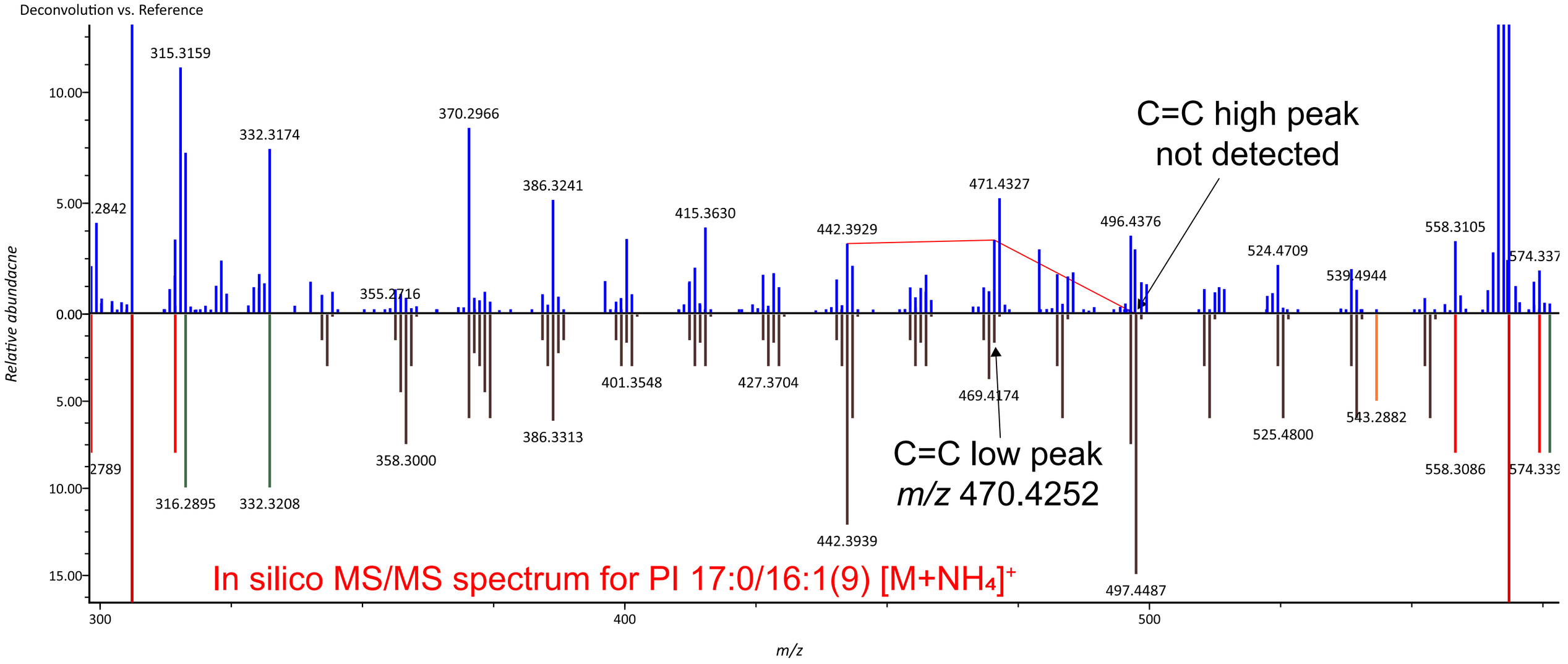


c PI 18:1(9)/18:1(9): the case of correct C=C position annotation in PI.



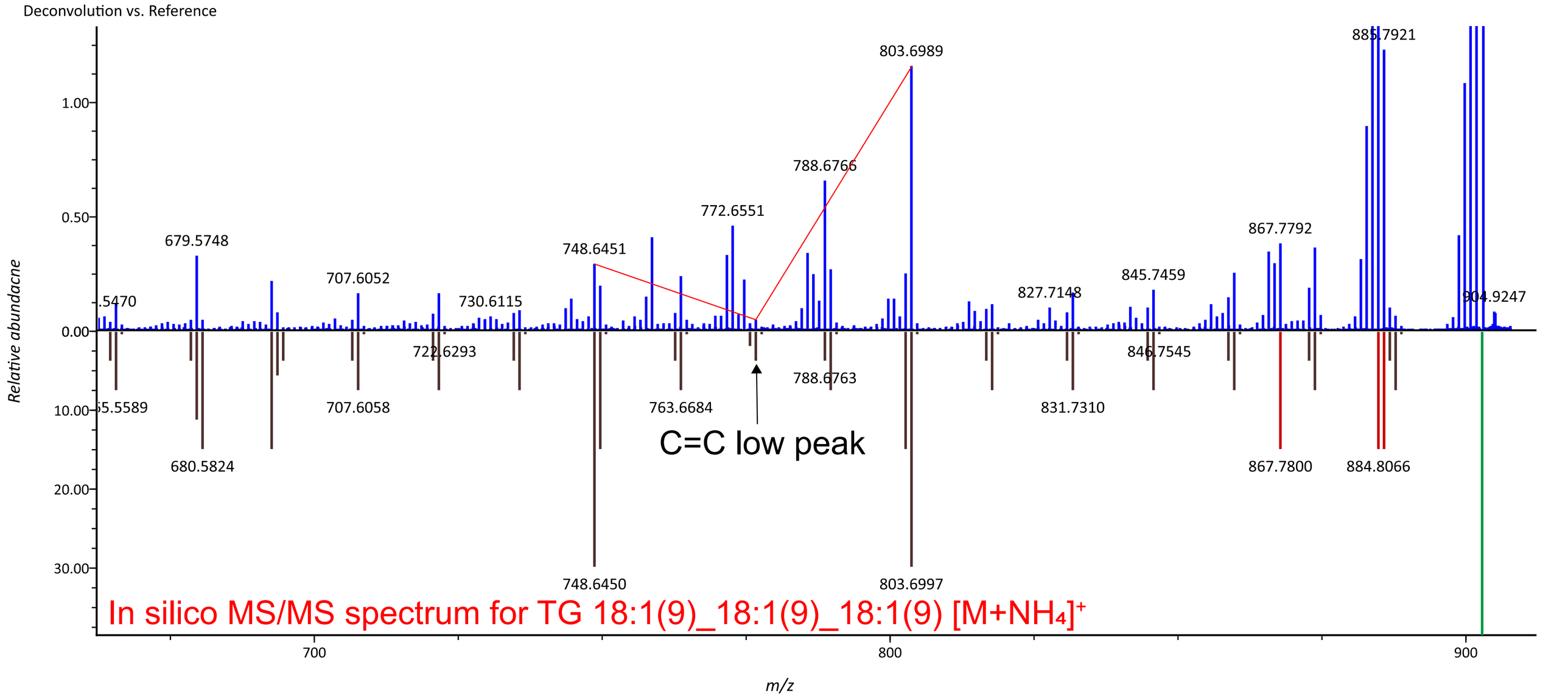
d

The original annotation was PI-d5 17:0/16:1(7), while the correct annotation shold be PI-d5 17:0/16:1(9)

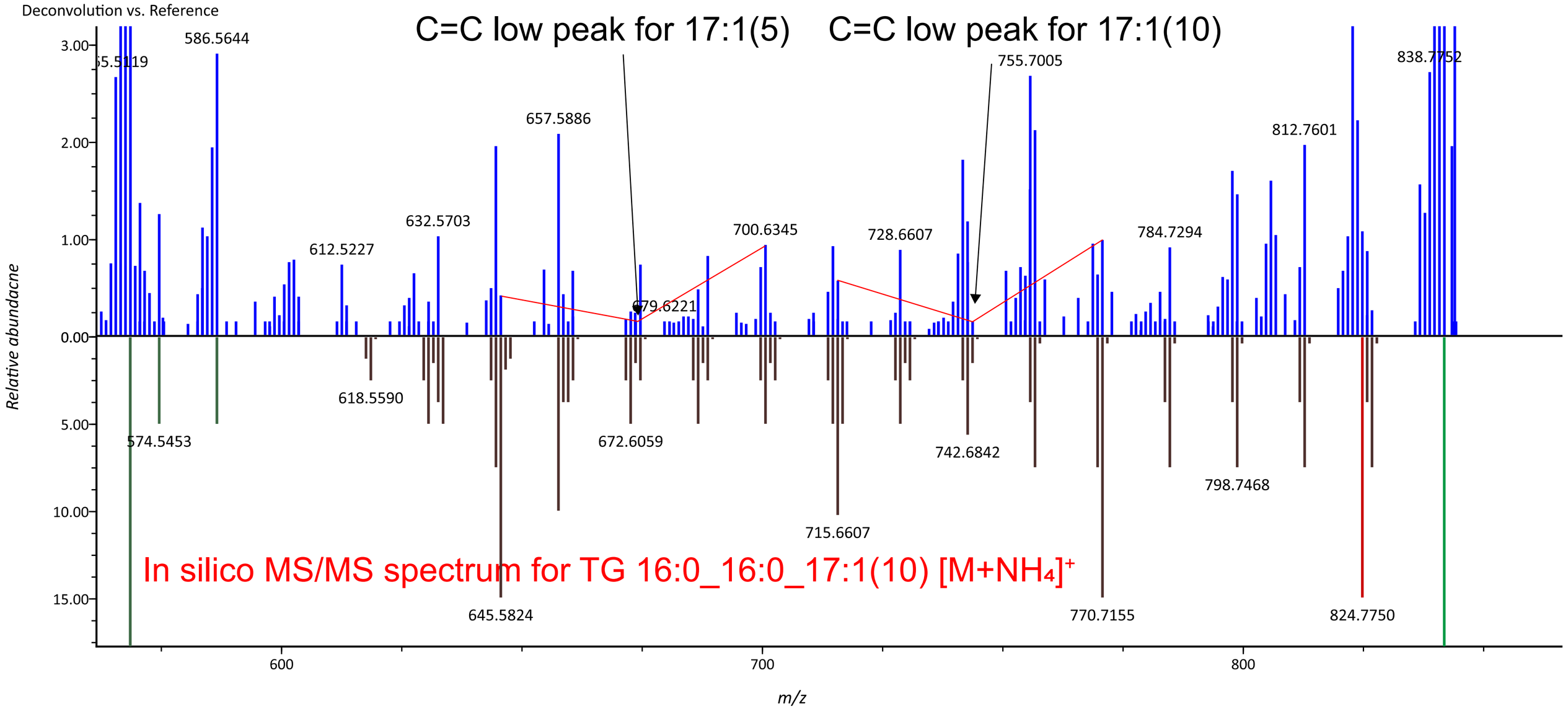


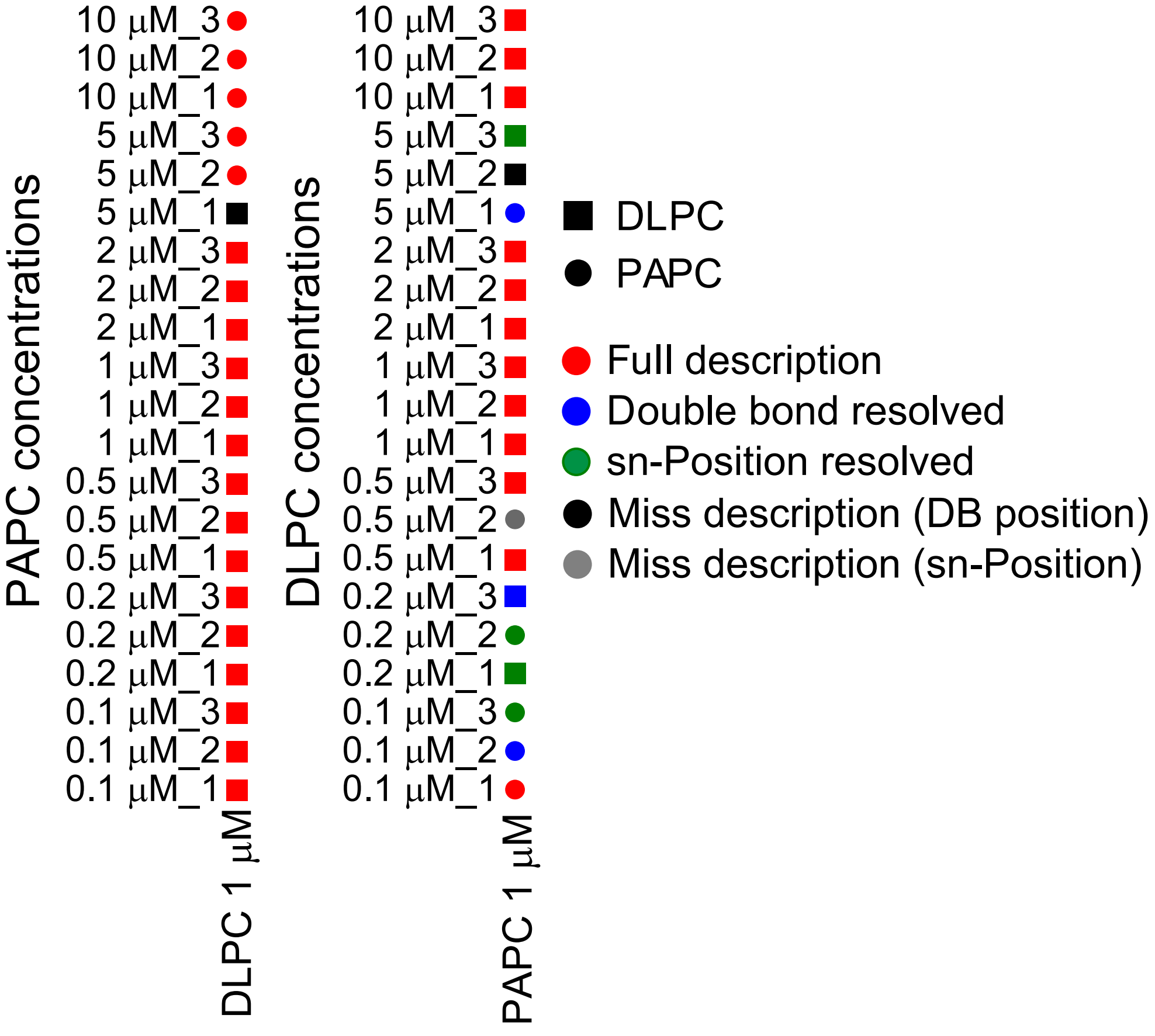
e

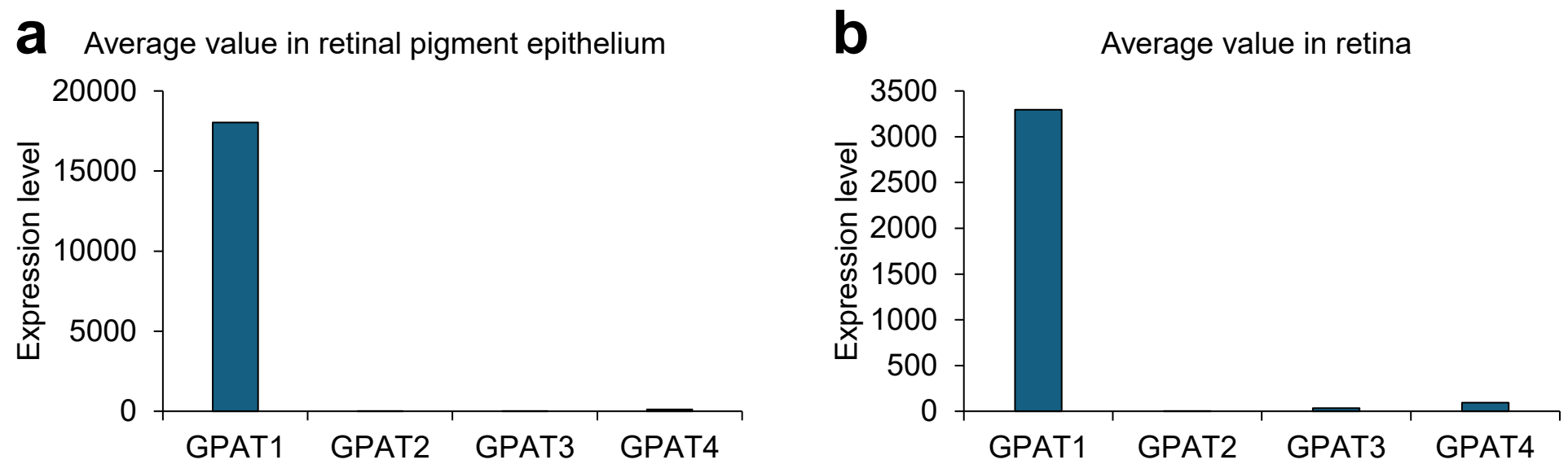
TG 18:1(9)_18:1(9)_18:1(9): the case of correct C=C position annotation in TG.

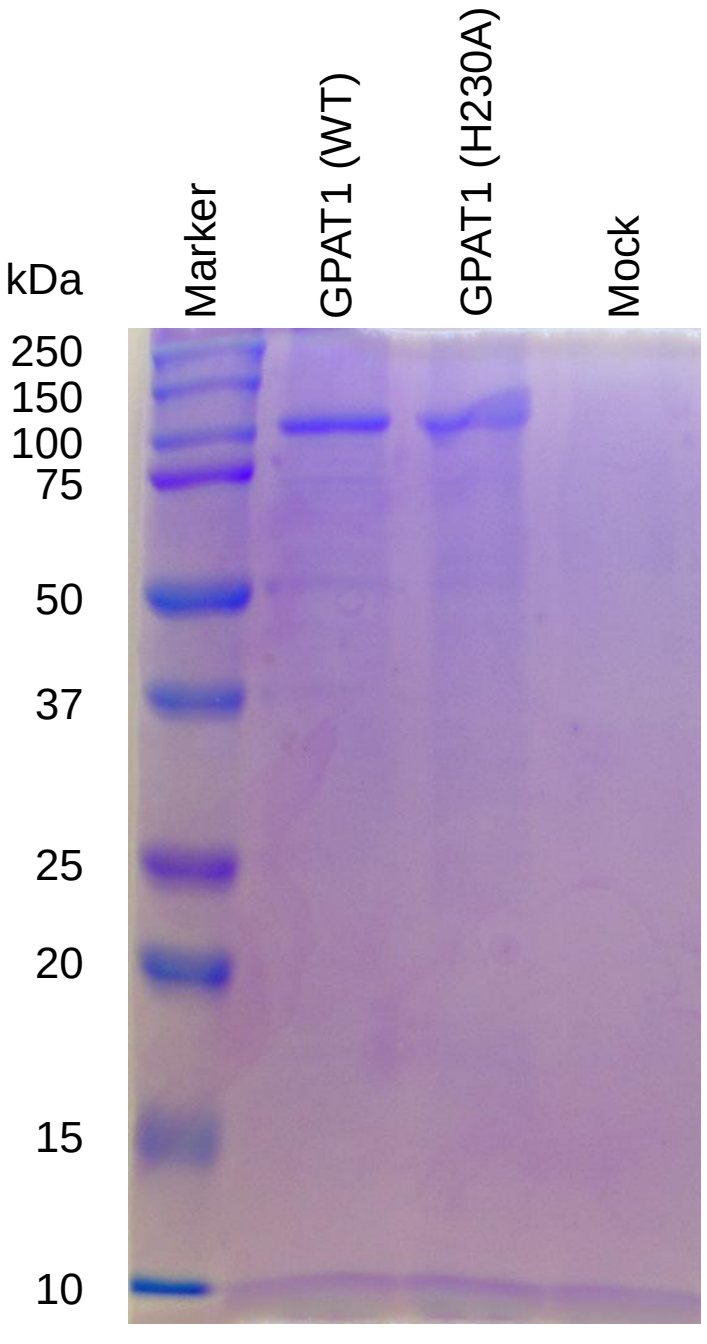


f The original annotation was TG-d5 16:0_16:0_17:1(5),
while the correct annotation shold be TG-d5 16:0_16:0_17:1(10)




PAPC concentrations
10 µM_3
10 µM_2
10 µM_1
5 µM_3
5 µM_2
5 µM_1
2 µM_3
2 µM_2
2 µM_1
1 µM_3
1 µM_2
1 µM_1
0.5 µM_3
0.5 µM_2
0.5 µM_1
0.2 µM_3
0.2 µM_2
0.2 µM_1
0.1 µM_3
0.1 µM_2
0.1 µM_1
DLPC 1 µM
DLPC concentrations
10 µM_3
10 µM_2
10 µM_1
5 µM_3
5 µM_2
5 µM_1
2 µM_3
2 µM_2
2 µM_1
1 µM_3
1 µM_2
1 µM_1
0.5 µM_3
0.5 µM_2
0.5 µM_1
0.2 µM_3
0.2 µM_2
0.2 µM_1
0.1 µM_3
0.1 µM_2
0.1 µM_1
PAPC 1 µM
DLPC
PAPC
Full description
Double bond resolved
sn-Position resolved
Miss description (DB position)
Miss description (sn-Position)


a
Average value in retinal pigment epithelium
Expression level
20000
15000
10000
5000
0
GPAT1
GPAT2
GPAT3
GPAT4
b
Average value in retina
Expression level
3500
3000
2500
2000
1500
1000
500
0
GPAT1
GPAT2
GPAT3
GPAT4


kDa
Marker
GPAT1 (WT)
GPAT1 (H230A)
Mock
250
150
100
75
50
37
25
20
15
10