Published in final edited form as:

Anal Chem. 2018 November 20; 90(22): 13523–13532. doi:10.1021/acs.analchem.8b03436.

Optimization of Electrospray Ionization Source Parameters for Lipidomics To Reduce Misannotation of In-Source Fragments as Precursor Ions

Rose M. Gathungu¹, Pablo Larrea^{1,*}, Matthew J. Sniatynski¹, Vasant R. Marur¹, John A. Bowden^{2,3}, Jeremy P. Koelmel⁴, Pamela Starke-Reed^{5,#}, Van S. Hubbard⁶, and Bruce S. Kristal^{1,&}

¹Department of Medicine, Division of Sleep and Circadian Disorders, Brigham and Women's Hospital and Department of Medicine, Division of Sleep Medicine, Harvard Medical School, Boston, MA 02115

²Center for Environmental and Human Toxicology, Department of Physiological Sciences, College of Veterinary Medicine, University of Florida, Gainesville, FL, 32610

³National Institute of Standards and Technology, Hollings Marine Laboratory, Charleston, SC 29412

⁴Department of Pathology, Immunology, and Laboratory Medicine, College of Medicine, University of Florida, Gainesville, FL 32610

⁵Deputy Director, NIH Division of Nutrition Research Coordination, Bethesda, MD 20892

⁶Director, NIH Division of Nutrition Research Coordination, Bethesda, MD 20892

Abstract

Lipidomics requires the accurate annotation of lipids in complex samples to enable determination of their biological relevance. We demonstrate that unintentional in-source fragmentation (ISF, common in lipidomics) generates ions that have identical masses to other lipids.

Lysophosphatidylcholines (LPC), for example, generate in-source fragments with the same mass

[&]CORRESPONDING AUTHOR: bkristal@bwh.harvard.edu.

^{*}Current Address: AbbVie, Worcester, MA

[#]Current Address: Deputy Administrator, Nutrition, Food Safety and Quality, Agricultural Research Service, USDA, Beltsville, Maryland 20705

Publisher's Disclaimer: Disclaimer:

Publisher's Disclaimer: Certain commercial equipment or instruments are identified in the paper to specify adequately the experimental procedures. Such identification does not imply recommendations or endorsement by the NIST; nor does it imply that the equipment or instruments are the best available for the purpose. Any use of trade, firm, or product names is for descriptive purposes only and does not constitute endorsement by the U.S. Government.

Conflict of Interest:

BSK is the inventor on general metabolomics-related IP that has been licensed to Metabolon via Weill Medical College of Cornell University and for which he receives royalty payments via Weill Medical College of Cornell University. He also consults for and has a small equity interest in the company. Metabolon offers biochemical profiling services and is developing molecular diagnostic assays detecting and monitoring disease. Metabolon has no rights or proprietary access to the research results presented and/or new IP generated under these grants/studies. BSK's interests were reviewed by the Brigham and Women's Hospital and Partners Healthcare in accordance with their institutional policy. Accordingly, upon review, the institution determined that BSK's financial interest in Metabolon does not create a significant financial conflict of interest (FCOI) with this research. The addition of this statement where appropriate was explicitly requested and approved by BWH.

as free fatty acids and lysophosphatidylethanolamines (LPE). The misannotation of in-source fragments as true lipids is particularly insidious in complex matrices since most masses are initially unannotated and comprehensive lipid standards are unavailable. Indeed, we show such LPE/LPC misannotations are incorporated in the data submitted to the NIST interlaboratory comparison exercise. Computer simulations exhaustively identified potential misannotations. The selection of in-source fragments of highly-abundant lipids as features, instead of the correct recognition of trace lipids, can potentially lead to: (i) missing the biologically relevant lipids (i.e., a false negative), and/or; (ii) incorrect assignment of a phenotype to an incorrect lipid (i.e., false positive). When ISF is not eliminated in the negative ion mode, ~40% of the 100 most abundant masses corresponding to unique phospholipids measured in plasma were artifacts from ISF. We show that chromatographic separation and ion intensity considerations assist in distinguishing precursor ions from in-source fragments, suggesting ISF may be especially problematic when complex samples are analyzed via shotgun lipidomics. We also conduct a systematic evaluation of ESI source parameters on a Exactive equipped with a HESI-II source with the objective of obtaining uniformly appropriate source conditions for a wide range of lipids, while, at the same time, reducing in-source fragmentation.

Graphical Abstract

Introduction:

Correct structural annotation of lipids is critical to enable the understanding of their biological role,¹⁻³ but often, it is complicated by the complexity of biological samples and the methodology/instrumentation employed. The multitude of masses in a single mass spectrum of a biological sample, especially in lipid-rich regions, can make it difficult to discriminate a precursor mass from a specific lipid of interest (or, in question) from those masses derived from chemical noise/artifacts.⁴⁻⁶ Mass signals that can complicate the interpretation of a mass spectrum include solvent-based clusters, multiple adducts for the same lipids, isotope peaks, and/or unintended fragment ions⁷. The interference from solvent masses can be resolved by, for example, using background subtraction from a blank sample.^{4,8,9} To avoid observation of multiple adducts for the same lipid, the concentration of the preferred adduct can be increased with the additional of salts, for example, preferentially creating lithiated adducts over protonated and sodiated adducts for some lipids.¹⁰⁻¹² Informatics approaches have also been developed to remove background peaks and identify

specific adducts of interest.^{4,7,13} Isotopic peaks can also be problematic in lipidomics, especially because it can be difficult to resolve the [M+2] ion from lipids that differ by two hydrogens, although this can be resolved when using high resolution instruments.^{14–17} In-source fragmentation can cause similar problems, especially when full scan mass spectra (MS1) are used as a first step in the structural annotation of intact lipids.¹⁸

In-source fragmentation (ISF) is a well-known phenomenon in ESI-based mass spectrometry.^{18–24} ISF occurs within the intermediate pressure region of the mass spectrometer, between the ESI source, which is at atmospheric pressure, and the analyzer, which is under vacuum.^{22–24} The degree of ISF is dependent on the source design and the analyte structure, with some molecules (e.g. sugars) being more prone to ISF than others. ISF occurs when the voltages that accelerate ions through the intermediate pressure region to the analyzer are too high. The high voltages provide additional internal energy to the molecule which leads to collision of the molecule with neutral gas molecules and thus the molecule dissociates.^{22–24} ISF can be problematic, especially if it produces ions that correspond to other endogenous compounds.^{18,20,21} The misannotation of in-source fragments as endogenous compounds has been observed in several fields including in proteomics, metabolomics, and lipidomics.^{18,20,21,25–28} For example, Kim et.al showed that ISF generated artifacts can be misidentified as peptides in proteomics.²¹ When examining the yeast metabolome, Xu et. al found more than 20 cases where in-source fragments mimicked common metabolites.¹⁸ In lipid analysis, Gelb et.al showed that lysophosphatidylserines can also break down in-source to generate masses that correspond to lysophosphatidic acid species.²⁰ These studies underscore the need to identify analytes that are prone to ISF during the method development stages, and, in turn, to reduce or eliminate these analytical artifacts, which can otherwise affect the eventual biological interpretation of the data.

Here we address the impact of unintended ISF on lipid identification and subsequent quantitation in complex biological samples. ISF of lipids has the potential to complicate lipid analysis in three distinct manners: i) since ISF generates ions without precursor pre-selection, if fragment ions have the same mass as endogenous lipids, the fragments may be misannotated as true lipids when they are not, ii) ISF further increases spectral complexity. The presence of both precursor and fragment ions can complicate data analysis (i.e. feature selection and spectral interpretation), and; iii) ISF can lower the intensity of the precursor ion of lipids and alter their quantitation thus producing falsely low concentrations. Moreover, when in-source fragments are misannotated as endogenous lipids, the concentration of the endogenous lipids with the same mass as the fragment ion can be reported as falsely high.²⁷ This work focuses on the misannotation of the in-source fragment ions of lysophospholipids (LPLs) and phosphatidylcholines (PCs). Here we show that LPLs can undergo ISF, producing ions that correspond to their fatty acyl (FA) chains that can in-turn be misannotated as free fatty acids (FFAs). Additionally, lysophosphatidylcholines (LPC) and phosphatidylcholines (PC) fragment in-source to produce ions with the same mass as lysophosphatidylethanolamines (LPE) or phosphatidylethanolamines (PE), respectively.

Additionally, we show that misannotation of in-source fragments remains prevalent in data within the lipidomics community. We specifically focus on the LPEs identified in the recently published National Institute of Standards and Technology (NIST) interlaboratory comparison exercise report.^{29,30} This report collectively analyzed results from a multi-laboratory lipidomics analysis of the NIST Standard Reference Material (SRM 1950). Our analysis of the SRM 1950 indicates that some of the LPEs identified in the interlaboratory report were in-source fragments of LPCs and demonstrates the utility of the chromatographic dimension to distinguish endogenous lipids from in-source fragments. For lipids that cannot be separated due to their co-elution, we show that post-hoc analysis based on retention time (RT), relative quantitation, and structural relationships can be used to identify potential artifacts. We also show the effect of ISF on quantitation of lipids in a plasma sample and on the day-to-day quantitative precision.

Materials and Methods:

Chemicals and Reagents:

LC-MS grade acetonitrile, isopropanol, methanol, and HPLC grade dichloromethane were obtained from ThermoFisher Scientific (Pittsburgh, PA). Ammonium formate was purchased from Sigma-Aldrich (St. Louis, MO). All lipid standards were purchased from Avanti Polar Lipids or Sigma-Aldrich.

Heparinized human plasma from 14 individuals (7 men/7 women) was purchased from Bioreclamation IVT (Waterbury, NY). An equal amount of plasma from each individual was used to create a pool (HHPool), which was used in all studies presented, and NIST SRM 1950 was obtained from NIST (Bethesda MD).

Lipid Extraction and LC-MS Analysis:

Lipid extraction and LC-MS analysis were performed as previously described (detailed in supplemental section).^{31,32}

The following changes were made to the LC-MS method: in addition to full scan spectra collected in the negative mode, full scan spectra of the lipid extract were also acquired by polarity switching between the positive and the negative mode. For polarity switching, the MS method was set to two scan events and the MS alternated between the positive and negative mode with one microscan acquired in each mode for the duration of the chromatographic run.

MS Source Parameters Optimization:

MS analysis was performed on an Exactive Benchtop Orbitrap Mass Spectrometer (Thermo Fisher, San Jose, CA) equipped with a heated electrospray ionization (HESI-II) probe. The HESI-II probe was run in the ESI mode (i.e., the HESI vaporizer temperature was off).

The source parameters assessed (all negative potentials) were the tube lens and the skimmer voltage. The skimmer voltages interrogated were 50, 40, 30, 20, 10 and 5V. The tube lens (TL) was assessed by reducing the TL in steps of 20 starting from 190 down to 90V

Lipid Quantitation:

Relative quantitation of selected lipids in the human plasma pools was done using either TraceFinder v. 4.1 or SIEVE v. 2.1 (Thermo Fisher Scientific, Cambridge, MA). SIEVE analysis was as previously described (details in the supplemental section).

Peak areas of all potential LPEs and LPCs in plasma (based on the results of the interlaboratory report) was obtained using TraceFinder 4.1. In-brief, a compound library based on the precursor ion mass (mass error ± 5 ppm) of all potential LPCs and LPEs was created within the software. TraceFinder data was then exported as a .csv file for further analysis.

Simulation of all Possible In-source Fragments:

An in-house R script (<https://github.com/KristalLab/FragmentationPaperScripts/>) was used to find all precursors with overlapping fragments (± 5 ppm) from other lipid ions. Fragments and precursors were queried from the LipidMatch³³ *in-silico* library. These calculations excluded oxidized lipids contained in LipidMatch *in-silico* library.

Results:

The goal of the work presented was to improve MS-based lipidomics analysis by identifying and addressing unintended ISF issues. Previous work has primarily focused on developing an LC-MS based-platform for broad-based lipidomics profiling^{31,32} and demonstrating the biological utility of this technology.^{34–36} Here, we highlight some challenges observed during structural assignment of lipids using the platforms typically employed by the community. Specifically, we interrogate and aim to resolve the three major issues: (i) unintended in-source fragments misannotated as real lipids; (ii) the effect of ISF on the relative quantitation of lipids, and; (iii) the determination of how specific ESI source parameters can result in ISF and the optimization of these parameters to reduce ISF.

In-source fragments derived from both LPLs and choline-containing lipids can be misannotated as real, biologically-feasible lipids. In the negative ionization mode, choline containing lipids (i.e., LPCs, PCs, and SMs), which ionize as formate or acetate adducts, can undergo ISF to form [M-15] ions that correspond to masses of endogenous lipids.^{37,38} The [M-15]⁻ fragment is obtained from the demethylation of the choline moiety. For LPC and PC, the [M-15]⁻ fragment could also correspond with a [M-H]⁻ of an LPE (for LPC) or PE (for a PC) with two more carbons than the original PC. The in-source fragment of PC(38:5), for example, produces an ion at m/z 792.5549, corresponding to that of PE(40:5). For the most part, when searched in databases (e.g. METLIN^{39,40}), the [M-15]⁻ fragment of SMs did not produce any lipid hits, although a few of the fragment ion masses did correspond to ceramide-phosphoethanolamines (cer-PE), a lipid class currently found only in insects and not humans.⁴¹ The misannotation of in-source fragments can be especially problematic when using single stage mass spectrometers (e.g., stand-alone orbitraps or TOFs), whereby structural annotation relies on the MS1 ion.

The ability of an in-source fragment from a choline-containing lipid to present as the precursor mass of another lipid, is readily illustrated by examining LPC(16:0). The spectrum

in Figure 1A shows the expected LPC(16:0) precursor ion as $[M+FA-H]^-$ at m/z 540.3312, but this spectrum also shows an unexpected base peak at m/z 480.3101. The latter peak at m/z 480.3101 is the in-source fragment from the demethylation of the LPC(16:0) ion. The reaction leading to the formation of the in-source fragment is shown in the schematic on Figure 1B. The mass of the fragment ion (m/z 480), also corresponds to that of the $[M-H]^-$ of LPE(18:0) (Figure 1C top panel). This illustrates the problem caused by ISF of LPC in the analysis of complex mixtures of pre-annotated lipids. LPLs also can undergo ISF to form ions corresponding to their respective FA chains. In addition to the LPE-like ion noted above, for example, LPC(16:0) can also undergo ISF to form an ion that corresponds to its FA chain (m/z 255.2331, FA(16:0)) via the loss of its head group (Figure 1A).

The above examples highlight the potential for misannotation of in-source fragments as endogenous lipids. This is particularly problematic in complex samples due to a lack of commercially available authentic standards and/or chromatographic resolution. Chromatographic separation can provide a potential mechanism to avoid misannotation of in-source fragments as endogenous lipids. For example, because the FFAs often elute later than the LPLs in reverse phase chromatography, they are less likely to be mis-annotated in LC-MS than an infusion-based approach. In lipidomics, LC (both reversed-phase and hydrophilic interaction chromatography (HILIC)) has been shown to separate lipids based on class and/or on FA chain composition.^{31,32,42,43} The lipid profiling method used did not separate PC from PE (or LPC from LPEs); however, we separated LPL from FFAs (Figures 2A and 2B). In the extracted ion chromatograms (XIC) of LPC(18:1), which elutes at 3.79 min, and LPC(20:4), which elutes at 2.57 min, each LPC is aligned with its respective FA fragment ion (Figure 2A and 2B). The XIC for each fatty acyl ion i.e. m/z 281.2486 (FA(18:1)) and m/z 303.2330; (FA(20:4)) showed a second chromatographic peak that eluted later than each LPC (**Bottom panel** Figures 2A and 2B). Separate analysis of authentic standards of FFA(18:1), (Figure 2C) and FFA(20:4) (Figure 2D) indicated that the second peak was a free fatty acyl while the first peak was an ISF fragment of the LPC. This result directly shows that ISF of LPCs yield masses corresponding to FFAs, and that this potential concern can be mitigated through chromatography. Thus, when chromatographic separation is used, and the mass of an LPL and a FFA are observed at the same retention time, the mass of the FFA is likely a fragment ion of the corresponding LPC.

Next, we wanted to establish whether misannotation of in-source fragments as endogenous lipids was a common occurrence in studies within the lipidomics community. We focused on LPEs reported in the recently published interlaboratory comparison exercise for lipidomics using SRM 1950 Metabolites in Frozen Human Plasma.^{29,30} Thirty-one lipidomic laboratories analyzed SRM 1950 using their individual routine lipidomics platforms (targeted or untargeted) with the goal of establishing a consensus of both qualitative and quantitative results. The interlaboratory report had a total of 35 LPEs reported by at least one laboratory. Of these 35 LPEs, 8 were reported by five or more laboratories, 17 were reported by at least 3 laboratories, and the remaining 18 were only reported by one laboratory. We also analyzed SRM 1950 using our lipidomics platform and assessed whether any of the 35 LPEs were potential insource fragments of LPCs. We analyzed the SRM under conditions that previously induced ISF. It should be noted that our platform cannot distinguish O-alkyl

and O-alk-1-enyl LPEs that have the same exact mass (e.g. LPE (O-16:1 and LPE(P-16:0), thus our analysis combined these species, and our list was comprised of 33 target LPEs. The results of our analysis are summarized in Table 1. The table includes the 33 LPEs in the interlaboratory report, the retention time of each LPE using our method (if it was detected), the corresponding LPCs that a misannotated LPE would originate from, and the retention time of the LPC. An LPE was assigned as a fragment ion if it was found at the same retention time as an LPC that had a mass that was 60 mass units (from the loss of the formate and methyl group) above it [LPE+60.0213 +/- 5ppm). With uncontrolled ISF, we observed 28 masses that could potentially be identified as LPEs. After changing to conditions that reduced ISF, we found that 15 of the 33 masses were indeed endogenous LPEs while 11 were potential in-source fragments of an LPC. We note that there is a possibility that some of the LPEs in the interlaboratory report are not detected in our platform because they are found below the limits of detection of our platform or they are too polar for our LC method. Our LC mobile phases have high amounts of organic solvents and the LC gradient starts at a relatively high percentage of organic solvent thus, LPLs with shorter chain lengths e.g. LPC(12:0) are potentially too polar for our method, the shortest chain length LPC we observed was LPC(14:0) at 2.2 minutes. For some LPEs (e.g., LPE (22:2) and LPE(22:3)), we can eliminate the possibility that they were too polar for our chromatographic method due to the general trend of retention order of other LPEs. LPEs with longer chains eluted later than those with shorter chains, and the presence of double bonds lead to an earlier elution. Thus both LPE(22:2) and LPE(22:3), would elute later than LPE(22:4) which has more double bonds and which is observed in our platform.

Unfortunately, not all potential fragmentation events can be conclusively evaluated via chromatographic separation. Arguably, the major problem is the high degree of co-elution between different PLs and SM species (including isomer-dependent chromatographic shifts).⁴⁴ This problem is compounded by the limited availability of lipid reference standards; indeed, only a few are currently commercially available to cover all the possible lipids in biological samples. Moreover, available lipid standards are expensive, and for structural confirmation, require that lipid standards are spiked in biological samples, which can consequently yield a more complex mass spectrum. When combined with ISF, co-elution complicates data analysis. RPLC-MS analysis of a lipid extract from a human plasma pool analyzed in the negative ion mode (Figure 3A) shows that there is little chromatographic overlap between LPLs and FFAs, but LPLs (specifically LPEs and LPCs) elute within the same time range (0-5 minutes). SM and PL species similarly co-elute (between 9-20 minutes, Figure 3A). Plasma samples contain hundreds of SM and PL species differing in signal intensity (and in plasma concentration) by up to six orders of magnitude. The presence of ISF can further complicate the analysis of these co-eluting lipids by increasing spectral complexity. This is illustrated in Figure 3B which shows a spectrum of co-eluting lipids (PCs, PEs and SMs) at 16.8 min, all at different intensities. The spectrum shows ~20 ions of which about half are in-source fragments. Elimination of ISF reduced the number of ions observed at that time point (Supplemental Figure 1)

In addition to complicating mass spectra, ISF of higher abundance lipids, in the context of co-elution of lipids, can limit the detection of trace biologically-relevant lipids that may be present in a sample. In lipidomics (and other 'omics) studies, feature detection and selection

is required prior to statistical analysis. Feature detection involves the selection of an m/z , the intensity of the m/z selected, and if LC is used, the retention time of each selected m/z . Ideally, each feature detected should represent a unique lipid^{45,46}. ISF of highly concentrated lipids can lead to the selection of fragment ions as features instead of endogenous trace lipids. This occurrence can lead to potentially missing biologically relevant low concentration lipids. For example, in two lipidomics studies, where 350 human plasma samples were profiled (data not shown) from the Multiethnic Cohort Study⁴⁷ and 700 human plasma samples were profiled from the CALERIE study^{48–51}, we found that of the top 300 features selected (i.e., the top most intense ions), approximately one-third appeared at first glance to be the precursor ion of a lipid. Of the approximately one hundred that behaved as molecular ions, approximately 30 % were PCs and another 20% were their in-source fragment ions that could be misannotated as PEs (Supplemental Table 1). These results indicate the significance of reducing ISF to ensure that features selected represent the correct endogenous lipids and not fragment ions.

For lipids that ionize in both positive and negative mode, but that only in one of the two modes with masses similar to those of another lipid type, analysis in both modes can potentially be used to eliminate the possibility of misannotation of in-source fragments as endogenous lipids. LPLs, SMs and PLs all ionize in both modes but with different ionization efficiencies. To establish the identity of the lipid, matching is conducted by comparing the m/z from one polarity with the predicted m/z in the opposite polarity at the same retention time. Looking for masses observed in one mode and not the other can potentially distinguish for example, endogenous PEs and PC in-source fragment ions misannotated as PEs. Supplemental Figure 2 shows a negative and positive mode MS spectrum at a select retention time (RT 13.97 min). In the negative ion mode spectrum, the m/z as $[M+FA-H]^-$ of PCs and their in-source fragments that could potentially be misannotated as $[M-H]^-$ of PEs were observed. In the positive mode, only the $[M+H]^+$ ions of the respective PCs were observed, serving as an indication that the other ions observed in the negative mode were in-source fragments. It should be noted that the observation of ions in one mode and not the other (e.g. not observing PEs in positive mode) could be due to their ion suppression (especially with the co-elution of lipids in complex samples) by lipids with better ionization efficiencies. This is especially true in the positive mode, as PCs will exhibit the same exact mass as a PE with 3 more carbons (e.g. PC(32:1) = PE(35:1)).

The examples shown above focus on the misannotation of in-source fragments of choline containing lipids (in the negative ionization mode) as endogenous lipids. Other lipid types can also undergo ISF to form ions that have the same precursor mass of other lipids.^{20,26,27} We therefore queried an *in-silico* library³³ to determine all possible fragment ions which overlapped with endogenous lipid precursor ions. In negative ionization mode, 34.9 % of all precursor ions had at least one fragment ion with overlapping m/z (± 5 ppm). In positive ionization mode (not a focus of this work), 8.1 % of all precursor ions had at least one fragment ion overlap. A non-exhaustive table of precursor ions and fragments from classes that commonly overlapped in mass is shown in the supplemental information (Supplemental Table 2), and a file containing an exhaustive list is also included in supplemental material. Examples include glycosyl-ceramides and ceramide phosphates, losing their glycosyl and phosphate group, respectively, leading to a fragment with the same mass as the respective

ceramide. Others include fragmentation of PLs (i.e. PG, PI, PS, PC) which after the loss of their head groups, form in-source fragments with the same exact mass as phosphatidic acids. Based on the simulated data, we re-analyzed our experimental data, and in accordance with the modelling predictions, we observed masses corresponding to those of phosphatidic acids at the same retention time as that of a PL and LPL with the same number of carbons and double bonds (e.g. PE(34:2) fragments to form PA(34:2)). Unexpected results included overlap of fragment ions $sn1 + sn2 + C3H6PO4$ (or $sn3 + sn4 + C3H6PO4$) from deprotonated cardiolipin with a respective phosphatidic acid species with one less degree of unsaturation. The findings from the *in-silico* exercise signify the importance of further characterizing in-source fragments that can overlap with precursor ions to reduce false positives. Alternatively, the use of MS/MS to aid exact mass identifications is highly advised, although in certain cases when fragments not only overlap in mass with precursor ions but are of similar or the same structure, MS/MS will not deduce these false positives.

One goal of a lipidomics study is to assess quantitative differences between two conditions (e.g., disease vs. control). To facilitate reaching this objective, it is important to ensure that quantitative differences observed are of biological origin and that they are not derived primarily from artifacts in the analytical workflow^{43,52,53}. ISF is potentially difficult to predict and control, and could affect ion reproducibility from injection-to-injection²³. Therefore, we assessed whether ISF was reproducible. This was performed by injecting the same pooled plasma sample multiple times. The injections were performed over several days (four injections/day over 6 days) into the LC-MS system with all instrumental conditions kept constant. This was performed to establish both inter- and intra-day reproducibility of all lipids that exhibited ISF. The coefficient of variation (CV) of the peak areas of select PCs, LPCs, SMs and their respective fragments, were assessed within a day and over the six days. The CVs within each (intra-day) and across the six days (inter-day) are shown in Supplemental Table 3. For each lipid and its fragment, the intraday CV was similar and was below 10%. The inter-day CVs were also at ~10%. This indicated that if all experimental (MS) conditions were maintained, ISF was consistent from one injection to the next. To avoid the challenges discussed in the previous sections (i.e., misannotation of fragments as endogenous lipids and selection of fragments as features), however, ISF still needed to be reduced. For highly abundant lipids, it was impossible to completely eliminate ISF. Our analysis (details in the Supplemental section), indicated that the parameters that dictate the extent of in-source fragmentation were the skimmer voltage and the tube lens voltage.

Discussion:

In summary, we have shown that unintended ISF of LPLs and lipids containing a choline head group commonly occur and produce fragment ions that can be misannotated as endogenous lipids (Supplemental Table 5). As long as all parameters remained unchanged, ISF is reproducible from one sample to the next, as was shown in the repeated analysis of a plasma pool, and thus was not associated with relative (sample-to-sample) quantitation errors. This can be used to confirm assignment of an ISF fragment, as the ratio (putative precursor lipid ion to putative ISF fragment) will have no biological variation across a series of samples. ISF can, however, potentially lead to underestimation of the concentration of lipids that fragment in-source, in -turn affecting biological interpretation. ISF also had

considerable effects on the presence or absence of specific lipids. Specifically, the LPLs fragment in-source producing a fragment ion with a mass similar to a free fatty acid; choline containing lipids i.e., SMs, LPC and PCs, lose a methyl group. For LPCs and PCs, the loss of the methyl group due to ISF leads to formation of an anion that has the same mass as a LPE and a PE, respectively. For SMs, the loss of a methyl group can lead to the formation of an anion that mimics the mass of a cer-PE, which is not found in humans and also, for the most part, this fragmentation event produces an anion that does not correspond to any known lipid. We note that while not the focus of the work presented, when intentionally induced, e.g. by manipulating source parameters, in-source MS/MS fragments can be used to obtain additional structural information of lipids, especially in single stage MS instruments that can otherwise only provide full scan precursor ion information.²²

Due to the abundance of choline-containing lipids in biological samples, these potential misannotations can be widespread. We have illustrated this by analyzing SRM 1950 under conditions yielding ISF and after reducing ISF. The comparison of our results with those in the recently published interlaboratory report, show that some of the LPEs identified in the NIST database/report are likely in-source fragments of LPCs. Querying an *in-silico* library of lipids shows that there are multiple lipid classes that can undergo ISF to produce ions that have the same mass as endogenous lipids. This indicates that the issue of misannotation of fragment ions as endogenous lipids is prevalent within the lipidomics community and should be addressed. These results highlight the importance of considering potential misannotations in developing and utilizing new and existing lipidomics workflows. Additionally, in-source fragments of abundant lipids like PCs can be assigned as features, which can lead to their misassignment as biologically relevant species when they are not endogenous (i.e. false positives). Our analysis of plasma samples from the MEC study suggests that, in the negative ionization mode, approximately 40% of the 100 most abundant masses (primarily phospholipids in negative mode) corresponding to unique lipids in a plasma lipidomics study were artifacts from ISF. The use of chromatography was crucial in establishing that 40% of these ions were fragment ions and not endogenous lipids. Moreover, these in-source fragments can mask the signal of trace lipids, leading to these low abundant lipids not being selected as features when they can potentially be biologically relevant (i.e. false negatives).

Using the data presented, we are now able to reduce potential misannotations using four approaches: (i) judicious choice of LPCs as a tuning lipid minimizes fragmentation while at the same time maintains the ionization efficiency; (ii) recognition of those lipids that are subject to ISF allows those lipids to be examined to determine if they are the molecular or fragment ions, this is especially true of LPEs, PEs, and FFAs; (iii) the use of chromatographic separation enables us to examine co-localization of potentially fragmented lipids and potential fragments to decisively determine the origin of a given lipid in question and (iv) elimination of likely in-source fragments from our data using a developed algorithm that can automatically recognize unintended fragment ions based on the theoretical mass difference between a given mass and the expected in-source fragment (i.e., for each given mass, automatically subtract the exact mass of formate+CH₃ (60.0212 $\pm$ 5 ppm)). This algorithm identifies parent masses and their corresponding fragments based on co-elution (i.e., it has a retention time constraint of $\pm$ 0.05 min) to avoid elimination of endogenous

lipids that have the same mass as in-source fragments (<https://github.com/KristalLab/FragmentationPaperScripts/>).

In the context of an ESI-based lipidomics study, chromatographic co-elution of lipids and likely in-source fragments of those lipids (e.g., LPC(20:0) [the presumptive parent], LPE(22:0) [presumptive fragment 1] and FFA(22:0) [presumptive fragment B]) can be an indication that ISF has occurred. Overall, examples presented in this report raise a concern about untargeted ESI-based (both shotgun and LC-MS based) lipid analysis that relies on MS1 both as a starting point for identification and for feature selection. In the case of LC-MS based lipidomics, we know that the chromatography domain is sufficient to separate FFAs from LPCs and LPEs – this indicates that co-elution (or lack thereof) can be directly used to mitigate potential ISF-derived artifacts in these lipid groups. Similar arguments hold for other transformations that can occur due to ISF, e.g., the conversion of a PC to a PE. Although not consistently chromatographically separable, the co-elution of a PC with a cognate PE must at least raise the possibility that the PE is a fragment of the PC and this must be considered during the data analysis phase. Because chromatographic separation allows one to follow a single peak, it is also possible to distinguish in-source fragments from endogenous lipids by analyzing chromatographically separated samples in both the positive and negative mode, especially if ISF produces ions with masses similar to those of other lipids in one mode and not the other. Overall, these analyses suggest that ISF would especially affect lipidomics analysis when chromatography is not used (i.e., shotgun lipidomics).

Conclusion:

Unintended ISF may prove problematic in complex biological samples, particularly when fragment ions present masses that correspond to endogenous lipids. These fragment ions can then be misannotated as endogenous lipids and also be selected as features over endogenous lipids found at lower concentrations. Fragmentation reduces the ion intensity of the precursor ion of these lipids that fragment in-source. The reduction of the ion intensity of the precursor ion would lead to the reporting of artificially low levels of the precursor ion^{27,28}. While ISF is dependent on the ESI source configuration, our data raise concerns for untargeted ESI-based (both shotgun and LC-MS based) lipids analysis that rely on MS1 both as a starting point for identification and for feature selection; this concern may play out primarily in shotgun lipidomics experiments due to the lack of LC-based separation. It is also important to consider that, as noted in the introduction, the degree of ISF is dependent on the source design and on analyte structure. While we have not tested the range of possible source configurations (e.g. ion-funnel type sources like the S-Lens), these differences at the atmospheric pressure interface potentially alter/increase ISF.^{18,22,23} We therefore emphasize that one key “take home message” of the results reported herein is that a test for ISF should be made a part of lipidomics method development.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

The work was supported by 1-R01-AG0457131 (BSK, PI), 1R01HL132556, 1P01CA168530 (Loic Le Marchand, PI; BSK, Project leader, project 3), P30 DK040561 (W. Allan Walker, MD; BSK Co-Director, metabolomics/lipidomics core), NIH contracts HHSN276201200648P, HHSN276201300457P, HHSN276201400554P (BSK, PI).

References:

- (1). Gross RW; Han X *Chemistry & Biology* 2011, 18, 284–291. [PubMed: 21439472]
- (2). Shevchenko A; Simons K *Nature Reviews. Molecular Cell Biology* 2010, 11, 593–598. [PubMed: 20606693]
- (3). Koelmel JP; Ulmer CZ; Jones CM; Yost RA; Bowden JA *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* 2017, 1862, 766–770. [PubMed: 28263877]
- (4). Kind T; Fiehn O *Bioanalytical Reviews* 2010, 2, 23–60. [PubMed: 21289855]
- (5). Wang SC; Huang CM; Chiang SM *Journal of Chromatography. A* 2007, 1161, 192–197. [PubMed: 17588590]
- (6). Wang S-Y; Kuo C-H; Tseng YJ *Analytical Chemistry* 2015, 87, 3048–3055. [PubMed: 25622715]
- (7). Görlach E; Richmond R *Analytical Chemistry* 1999, 71, 5557–5562.
- (8). Fredriksson M; Petersson P; Jornten-Karlsson M; Axelsson BO; Bylund D *Journal of Chromatography. A* 2007, 1172, 135–150. [PubMed: 17959187]
- (9). Dromey RG; Stefik MJ; Rindfleisch TC; Duffield AM *Analytical Chemistry* 1976, 48, 1368–1375.
- (10). Choi SS; Song MJ *Rapid Communications in Mass Spectrometry : RCM* 2008, 22, 2580–2586. [PubMed: 18646250]
- (11). Hsu FF; Bohrer A; Turk J *Journal of the American Society for Mass Spectrometry* 1998, 9, 516–526. [PubMed: 9879366]
- (12). Hsu FF; Turk J *Journal of the American Society for Mass Spectrometry* 2001, 12, 61–79. [PubMed: 11142362]
- (13). Krishnan S; Verheij EE; Bas RC; Hendriks MW; Hankemeier T; Thissen U; Coulier L *Rapid Communications in Mass Spectrometry : RCM* 2013, 27, 917–923. [PubMed: 23592192]
- (14). Han X; Yang K; Gross RW *Mass Spectrometry Reviews* 2012, 31, 134–178. [PubMed: 21755525]
- (15). Almeida R; Pauling JK; Sokol E; Hannibal-Bach HK; Ejsing CS *Journal of the American Society for Mass Spectrometry* 2015, 26, 133–148. [PubMed: 25391725]
- (16). Schwudke D; Schuhmann K; Herzog R; Bornstein SR; Shevchenko A *Cold Spring Harbor Perspectives in Biology* 2011, 3, a004614. [PubMed: 21610115]
- (17). Wang M; Huang Y; Han X *Rapid Communications in Mass Spectrometry : RCM* 2014, 28, 2201–2210. [PubMed: 25178724]
- (18). Xu YF; Lu W; Rabinowitz JD *Analytical Chemistry* 2015, 87, 2273–2281. [PubMed: 25591916]
- (19). Godzien J; Armitage EG; Angulo S; Martinez-Alcazar MP; Alonso-Herranz V; Otero A; Lopez-Gonzalez A; Barbas C *Electrophoresis* 2015.
- (20). Bollinger JG; Li H; Sadilek M; Gelb MH *Journal of Lipid Research* 2010, 51, 440–447. [PubMed: 19717841]
- (21). Kim JS; Monroe ME; Camp DG, 2nd; Smith RD; Qian WJ *Journal of Proteome Research* 2013, 12, 910–916. [PubMed: 23268687]nd
- (22). Abranko L; Garcia-Reyes JF; Molina-Diaz A *Journal of Mass Spectrometry : JMS* 2011, 46, 478–488. [PubMed: 21500306]
- (23). Gabelica V; De Pauw E *Mass Spectrometry Reviews* 2005, 24, 566–587. [PubMed: 15317019]
- (24). Song F *Journal of Agricultural and Food Chemistry* 2011, 59, 4361–4364. [PubMed: 21417433]
- (25). Farwanah H; Kolter T *Metabolites* 2012, 2, 134. [PubMed: 24957371]
- (26). Xu L; Wang X; Jiao Y; Liu X *Talanta* 2018, 178, 287–293. [PubMed: 29136824]
- (27). Zhao Z; Xu Y *Journal of chromatography. B, Analytical Technologies in the Biomedical and Life Sciences* 2009, 877, 3739–3742. [PubMed: 19734112]

- (28). Onorato JM; Shipkova P; Minnich A; Aubry A-F; Easter J; Tymiak A *Journal of Lipid Research* 2014, 55, 1784–1796. [PubMed: 24872406]
- (29). Bowden JA; Heckert A; Ulmer CZ; Jones CM; Koelmel JP; Abdullah L; Ahonen L; Alnouti Y; Armando AM; Asara JM; Bamba T; Barr JR; Bergquist J; Borchers CH; Brandsma J; Breitskopf SB; Cajka T; Cazenave-Gassiot A; Checa A; Cinel MA, et al. *Journal of Lipid Research* 2017, 58, 2275–2288. [PubMed: 28986437]
- (30). Bowden JA; Heckert A; Ulmer CZ; Jones CM 2017.
- (31). Bird SS; Marur VR; Sniatynski MJ; Greenberg HK; Kristal BS *Analytical Chemistry* 2011, 83, 6648–6657. [PubMed: 21774539]
- (32). Bird SS; Marur VR; Sniatynski MJ; Greenberg HK; Kristal BS *Analytical Chemistry* 2011, 83, 940–949. [PubMed: 21192696]
- (33). Koelmel JP; Kroeger NM; Ulmer CZ; Bowden JA; Patterson RE; Cochran JA; Beecher CWW; Garrett TJ; Yost RA *BMC Bioinformatics* 2017, 18, 331. [PubMed: 28693421]
- (34). Bird SS; Marur VR; Stavrovskaya IG; Kristal BS *Metabolomics : Official Journal of the Metabolomic Society* 2013, 9, 67–83. [PubMed: 23646040]
- (35). Stavrovskaya IG; Bird SS; Marur VR; Baranov SV; Greenberg HK; Porter CL; Kristal BS *Journal of Lipids* 2012, 2012, 797105. [PubMed: 22970378]
- (36). Stavrovskaya IG; Bird SS; Marur VR; Sniatynski MJ; Baranov SV; Greenberg HK; Porter CL; Kristal BS *Journal of Lipid Research* 2013, 54, 2623–2635. [PubMed: 23690505]
- (37). Han X; Gross RW *Journal of the American Society for Mass Spectrometry* 1995, 6, 1202–1210. [PubMed: 24214071]
- (38). Harrison KA; Murphy RC *Journal of Mass Spectrometry* 1995, 30, 1772–1773.
- (39). Guijas C; Montenegro-Burke JR; Domingo-Almenara X; Palermo A; Warth B; Hermann G; Koellensperger G; Huan T; Uritboonthai W; Aisporna AE; Wolan DW; Spilker ME; Benton HP; Siuzdak G *Analytical Chemistry* 2018, 90, 3156–3164. [PubMed: 29381867]
- (40). Smith CA; O'Maille G; Want EJ; Qin C; Trauger SA; Brandon TR; Custodio DE; Abagyan R; Siuzdak G *Therapeutic Drug Monitoring* 2005, 27, 747–751. [PubMed: 16404815]
- (41). Masood MA; Yuan C; Acharya JK; Veenstra TD; Blonder J *Analytical Biochemistry* 2010, 400, 259–269. [PubMed: 20122889]
- (42). Okazaki Y; Kamide Y; Hirai MY; Saito K *Metabolomics : Official Journal of the Metabolomic Society* 2013, 9, 121–131. [PubMed: 23463370]
- (43). Scherer M; Schmitz G; Liebisch G *Analytical Chemistry* 2010, 82, 8794–8799. [PubMed: 20945919]
- (44). Bird SS; Marur VR; Stavrovskaya IG; Kristal BS *Analytical Chemistry* 2012, 84, 5509–5517. [PubMed: 22656324]
- (45). Tautenhahn R; Bottcher C; Neumann S *BMC Bioinformatics* 2008, 9, 504. [PubMed: 19040729]
- (46). Trevino V; Yanez-Garza IL; Rodriguez-Lopez CE; Urrea-Lopez R; Garza-Rodriguez ML; Barrera-Saldana HA; Tamez-Pena JG; Winkler R; Diaz de-la-Garza RI *Journal of Mass Spectrometry : JMS* 2015, 50, 165–174. [PubMed: 25601689]
- (47). Kolonel LN; Henderson BE; Hankin JH; Nomura AM; Wilkens LR; Pike MC; Stram DO; Monroe KR; Earle ME; Nagamine FS *American Journal of Epidemiology* 2000, 151, 346–357. [PubMed: 10695593]
- (48). Fontana L; Villareal DT; Das SK; Smith SR; Meydani SN; Pittas AG; Klein S; Bhapkar M; Rochon J; Ravussin E; Holloszy JO *Aging Cell* 2015.
- (49). Ravussin E; Redman LM; Rochon J; Das SK; Fontana L; Kraus WE; Romashkan S; Williamson DA; Meydani SN; Villareal DT; Smith SR; Stein RI; Scott TM; Stewart TM; Saltzman E; Klein S; Bhapkar M; Martin CK; Gilhooly CH; Holloszy JO, et al. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences* 2015, 70, 1097–1104.
- (50). Redman LM; Kraus WE; Bhapkar M; Das SK; Racette SB; Martin CK; Fontana L; Wong WW; Roberts SB; Ravussin E *The American Journal of Clinical Nutrition* 2014, 99, 71–78. [PubMed: 24257721]

- (51). Stewart TM; Bhapkar M; Das S; Galan K; Martin CK; McAdams L; Pieper C; Redman L; Roberts S; Stein RI; Rochon J; Williamson DA Contemporary Clinical Trials 2013, 34, 10–20. [PubMed: 22981898]
- (52). Cífková E; Holápek M; Lísa M; Ovčáková M; Lyčka A; Lynen F; Sandra P Analytical Chemistry 2012, 84, 10064–10070. [PubMed: 23072569]
- (53). Jung HR; Sylvanne T; Koistinen KM; Tarasov K; Kauhanen D; Ekroos K Biochimica et Biophysica Acta 2011, 1811, 925–934. [PubMed: 21767661]

Fig. 1:

ISF of LPC(16:0) yields biologically plausible fragment artifacts

A: Mass spectrum of LPC(16:0). ISF of the $[M+FA-H]^-$ of LPC(16:0) (m/z 540.3314), lead to the observation of ions at m/z 480.3101 (after demethylation) and m/z 255.2331 from the 16:0 FA chain. **B:** In-source demethylation reaction of LPC(16:0). This schematic shows the fragmentation pathway of the formate adduct of LPC (16:0) which leads to the loss of a CH_3 moiety from the choline. All formate adducts of choline containing lipids undergo this pathway. **C and D:** Structures of lipids that the fragment ions of LPC (16:0) can be misannotated as LPE(18:0). The in-source fragments of LPC(16:0) can be misannotated as LPE(18:0) (C) and FA(16:0) (D).

Fig 2.

Chromatographic separations prevents misannotation of LPL fragments as endogenous FFAs
A and B: The chromatograms of LPC(18:1) (a) and LPC(20:4), in a plasma pool sample and their respective FA in-source fragment ions. The LPCs and the fragment ions of their FA chains align chromatographically which enables the assignment of the FA ions as fragment ions. **C and D:** Chromatograms of the FFA standards of FFA(18:1) and FFA(20:4) which elute later than LPCs.

Fig. 3.

Co-elution of PLs, SMs, and their in-source fragments

a: Negative ion total ion chromatogram showing separation of lipids in a human plasma pool. Regions of the chromatogram where lipids from the different lipid classes are labelled.

b: The corresponding mass spectrum of lipids eluting at ~16.5 min (a mixture of sphingomyelins and phospholipids) showing the complexity observed due to lipid co-elution. Spectra complexity is further increased by ISF leads to the observation of more ions at the same retention time than when ISF is eliminated.

Table 1.

A list of the LPEs found in the NIST interlaboratory report, including the number of labs that observed the LPE^{29,30}. The third column shows the retention time of an LPE if it is observed as an endogenous lipid or whether it is a fragment ion (LPEs observed in our platform as fragments are highlighted in yellow). The fourth column shows the corresponding LPC that an LPE can be derived from if it's a fragment ion and the last column shows the retention time of the LPCs.

LPEs	# of Labs in NIST Report	LPE RT (Minutes)	Corresponding LPC	LPC + ISF RT (Minutes)
LPE(17:0)	1	Fragment at 2.50	LPC(15:0)	2.52
LPE(19:0)	1	Fragment at 3.8	LPC(17:0)	3.8
LPE(19:1)	1	Fragment at 2.73	LPC(17:1)	2.7
LPE(20:0)	1	Fragment at 4.72	LPC(18:0)	4.7
LPE(20:2)	3	Fragment at 2.5	LPC(18:2)	2.5
LPE(22:0)	3	Fragment at 6.6	LPC(20:0)	6.6
LPE(22:1)	5	Fragment at 4.8	LPC(20:1)	4.8
LPE(22:2)	1	Fragment at 3.57	LPC(20:2)	3.6
LPE(22:3)	1	Fragment at 2.69	LPC(20:3)	2.7
LPE(P-20:0) or LPE(O-20:1)	1	Fragments at 4.03 and 5.38	LPC(O-18:1) /LPC(P-18:0)	4.03 and 5.4
LPE(P-20:1)	1	Fragment at 3.9	LPC(O-18:2)/LPC(P-18:1)	3.9
LPE(18:3)	3	2.07	No Corresponding LPC	
LPE(22:6)	12	2.2	LPC(20:6)	Not Observed
LPE(20:4)	14	2.45	No Corresponding LPC	
LPE(18:2)	16	2.5	No Corresponding LPC	
LPE(22:5)	3	2.5	LPC(20:5)	1.9
LPE(20:3)	5	2.8	LPC(18:3)	2.07
LPE(18:1)	14	3.1	LPC(16:1)	2.2
LPE(22:4)	3	3.2	LPC(20:4)	2.3
LPE(16:0)	14	3.6	LPC(14:0)	2.2
LPE(18:0)	15	4.9	LPC(16:0)	3
LPE(20:1)	4	5	LPC(18:1)	3.2
LPE(P-18:0)/LPE(O-18:1)	1	6.17	LPC(O-16:1)/LPC(P-16:0)	4.49
LPE(20:5)	3	2 Ion Count Below 1E3	No Corresponding LPC	
LPE(16:1)	3	2.5 Ion Count Below 1E3	LPC(14:1)	Not Observed
LPE(P-16:0) or LPE(O-16:1)	1	3.8 Ion Count Below 1E3	LPC(O-14:1)/LPC(P-14:0)	Not Observed
LPE(34:1)	1	Observed as fragment of PC(O-32:0)		
LPE(12:0)	1	Not Observed	LPC(10:0)	Not Observed
LPE(14:0)	1	Not Observed	LPC(12:0)	Not Observed
LPE(24:0)	1	Not Observed	LPC(22:0)	8.1
LPE(O-16:0)	1	Not Observed		Not Observed
LPE(P-16:1)	1	Not Observed	No Corresponding LPC	
LPE(P-18:1)	3	Not Observed	LPC(O-16:2)	Not Observed