Analysis of complex lipid samples by mass spectrometry

Perttu Haimi
Institute of Biomedicine/Biochemistry,
University of Helsinki, Finland

April 18, 2005

Abstract

We developed a software tool for the analysis of mass spectra derived from complex lipid samples. The software corrects for isotopic overlap of peaks in the mass spectrum using calculated theoretical isotopic distributions.

1 Introduction

Mass spectrometry (MS) is a technique for analyzing chemical compounds. A mass spectrometer consists of ion source, mass analyzer and detector. The ion source produces free (gas phase) ions from the sample. Mass analyzer contains a magnetic field which creates a force on the passing ions, proportional to their mass to charge ratio (m/z). The field in the mass analyzer can be tuned so that only ions with particular mass to charge ratio are allowed to pass to the detector. Mass spectrum is obtainded when the instrument scans a selected range of m/z values and records the detected signal intensity at each point.

In the field of lipidomics, mass spectrometry, coupled with an electrospray interface (ESI) as the ion source, offers an alternative for the traditional quantitative methods, which require several analytical steps and relatively high amounts of sample. Mass spectrometry has been used in the study of intracellular lipid trafficking [6], the composition of plasma membrane domains [4], and disorders in lipid-mediated signaling [7]. Together with heavy isotope-labeled lipid precursors, it has been used in the elucidation of lipid metabolic pathways [3], [5].

Many approaches have been adopted for the analysis of lipidomes by ESI-MS [10]. In one of these, a crude lipid extract is infused to the MS instrument and either direct MS scans or specific precursor ion or neutral-loss scans are used to identify the different lipid species. Another commonly used method (LC-MS) employs liquid chromatography with on-line mass spectrometric detection, in which case several scans, at different time points, are collected. The software presented here is applicable for both methods. Current commercial MS software are ill-suited for lipidome analyses because of the great number of partially overlapping lipid species in most samples.

2 Materials and methods

The software is written in ANSI C++ and compiles with GCC 3.3.5. The gnu scientific library is used for linear least squares fitting. Mass spectra were acquired with Quattro Micro triple quadrupole mass spectrometer (Micromass, Manchester, U.K.).

2.1 Data preprocessing

The first step of the analysis is the conversion of the raw data to netCDF format. This is done using the Micromass DataBridge program, which converts the raw data after the acquisition of the spectrum. Next, the spectral data is linearly interpolated to obtain an evenly spaced grid with spacing of 0.1 m/z units. After this, a Gaussian filter with width of 0.3 m/z (full width at half maximum) is applied. This width was chosen because it was wide enough to smooth out the detector noise but did not cause loss of true peaks in our datasets. Next, all local maxima are collected to obtain a list of peak positions. The integrals of the peaks are approximated by calculating the sum of data points within of 0.3 m/z units from the peak center. The output of the preprocessing is a list of peak positions and their corresponding integrals. Alternatively, a peak list generated by any other software can be used.

2.2 Calculation of isotopic distributions

Since most elements occur in several isotopic forms, mass spectrometric analysis of a compound gives rise to a pattern of multiple peaks. The peak with the lowest m/z value is called the mono-isotopic peak, and it contains only atoms of the lightest isotope of each element. Subsequent peaks have one or more extra neutrons in their constituent atoms.

Lipids occur in most biological samples as a series of molecules, differing from one another only by one double bond (which corresponds to two hydrogen atoms). Thus, the isotopic patterns of different species frequently overlap, making the identification and especially the quantitation problematic. The program constructed here allows for the correction of isotopic overlap by deconvolution of the peaks according to calculated theoretical isotopic peak abundances.

As input, the program needs a list of expected analytes and their molecular formulas. It must be noted that if specific MS/MS-modes are used, the isotopic pattern differs from the one calculated for the mother ion. Hence, the formula of the specific fragment is also required. The isotope patterns are calculated from molecular formulas and the relative isotope abundances [9] of the elements by a method utilizing fast Fourier transform [2]. The isotopic patterns for the spectra acquired in specific MS/MS-modes are calculated by adding the mass of the selected fragment to the isotopic pattern of the complementary fragment [1]. Variations in the natural abundancies of stable isotopes exist, so it is possible that the theoretical peak patterns do not correspond exactly to the measured ones in a specific sample. However, the differences are expected to be fairly small in most cases.

2.3 Peak assignment and deconvolution

The assignment begins by comparing the m/z values of the determined peak positions with those in the theoretical peak list. If the m/z of the theoretical peak is within 0.3 m/z distance from the found, the peak is assigned to that species. Next, all the assigned peaks are divided into connected groups. This is done by constructing a graph from the peaks such that the vertexes correspond to peaks and an edge exists between vertexes if the peaks are assigned to the same compound. Connected components can then be found by traversing the graph. The peaks in one connected group are deconvoluted by linear least squares fitting of the theoretical isotopic distributions to the peak intensities of the acquired mass spectrum as described by [8] and the error estimates for the deconvoluted intensities are determined. In problematic cases, e.g. when a large unassigned peak is found close to the peaks to be deconvoluted, the equation system is underdetermined (i.e. has more compounds than there are measured peaks), or if the residuals of the fit are large, an error condition is reported.

3 Results and discussion

We compared the calculated isotope patterns to measured patterns from known compounds. Figure 1 shows the results of fitting the theoretical pattern to 48 repeated measurements of one lipid, PC-28:0 in two scan modes. The agreement is relatively good, and the relative error increases for smaller peaks.

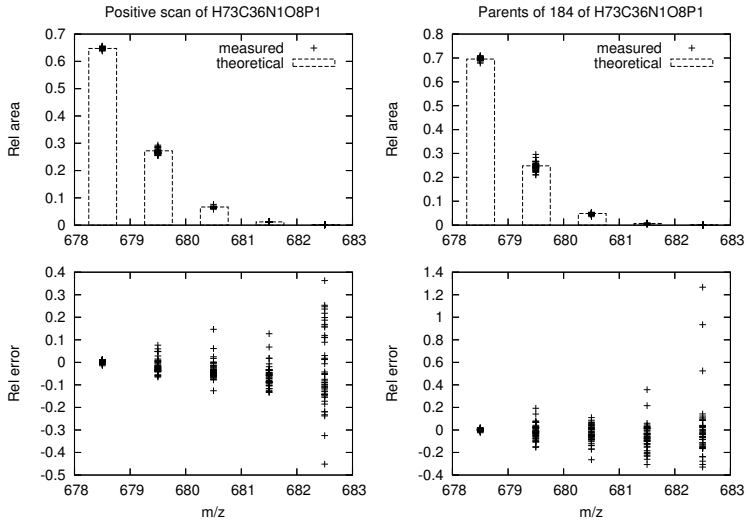

Figure 1: Comparison of theoretical isotope distribution to measured distributions. Relative error is calculated as $(A_t - A_m)/A_t$, where A_t and A_m are the theoretical and measured peak areas, respecively.

The performance of the deconvolution algorithm was tested with a biological sample from HDL (High density lipoprotein) extracts. These samples contain two different classes of lipids (phosphatidyl choline and sphingomyelin) with heavy overlap (see figure 2). These classes have a very similar fragmentation behavior and are not separated during the mass spectrometric analysis. Separation of sphingomyelin can be done chemically by destroying the phosphatidyl choline species by hydrolysis. The performance of the algorithm in resolving minor species of sphingomyelin in the presence of more abundant phosphatidylcholine was tested by comparing the results from samples hydrolyzed in the presence of NaOH to the original samples.

Table 1 shows the results from deconvolution algorithm from eight replicate measurements. The peak areas are divided by the peak area of internal standard (SM-25:0). For comparison, the relative areas of the monoisotopic peaks are shown without any attempt to correct for overlap. Deconvolution

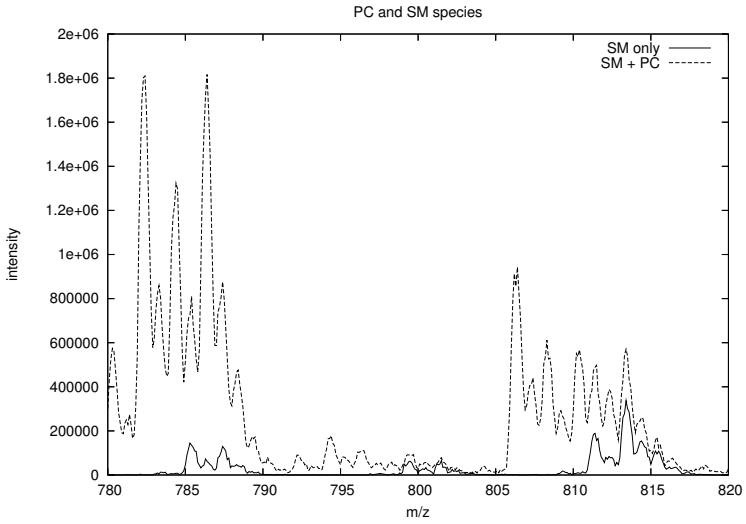

Figure 2: Overlapping peaks of phosphatidyl choline and sphingomyelin

Lipid name	Monoisotopic mass	Hydrolyzed	Deconvoluted	Monoisotopic peak
SM-22:1	785.7	0.67 (0.03)	0.81 (0.13)	3.3 (0.2)
SM-22:0	787.7	0.51 (0.04)	0.63 (0.13)	3.7 (0.3)
SM-24:2	811.7	0.97 (0.04)	0.99 (0.10)	2.1 (0.2)
SM-24:1	813.7	1.58 (0.06)	1.58 (0.06)	2.17 (0.06)

Table 1: Average relative areas of four sphingomyelin species. Standard deviation in parenthesis.

step seems to have some benefit (i.e. the relative areas are closer to the areas in the hydrolyzed sample).

3.1 Improvements

The spectrum preprocessing could be improved with more sophisticated signal processing methods. Currently, the spectrum preprocessing step does not handle high-resolution spectra and thus will not work well with ions having multiple charges. The decovolution step could be improved by constraining the problem to positive solutions. Also, combining results from different scan modes could make it possible to better resolve lipids with the same molecular formula [11].

References

[1] Rockwood AL, Kushnir MM, and Gordon NJ. Dissociation of individual isotopic peaks: Predicting isotopic distributions of product ions in msn. *J Am Soc Mass Spectrom*, 14:311–322, 2003.

[2] Rockwood AL and Van Orden SL. Ultrahigh-speed calculation of isotope distributions. *Anal Chem*, 68:2027–2030, 1996.

[3] Hunt AN, Clark, GT, Attard GS, and Postle AD. Highly saturated endonuclear phosphatidylcholine is synthesized in situ and colocated with CDP-choline pathway enzymes. *J Biol Chem*, 276:8492–9, 2001.

[4] Blom TS et al. Mass spectrometric analysis reveals an increase in plasma membrane polyunsaturated phospholipid species upon cellular cholesterol loading. *Biochemistry*, 40:14635–44, 2001.

[5] Boumann HA et al. The two biosynthetic routes leading to phosphatidylcholine in yeast produce different sets of molecular species. evidence for lipid remodeling. *Biochemistry*, 42:3054–9, 2003.

[6] Schneiter R et al. Electrospray ionization tandem mass spectrometry (esi-ms/ms) analysis of the lipid molecular species composition of yeast subcellular membranes reveals acyl chain-based sorting/remodeling of distinct molecular species en route to the plasma membrane. *J Cell Biol*, 146:741–54, 1999.

[7] Wenk MR et al. Phosphoinositide profiling in complex lipid mixtures using electrospray ionization mass spectrometry. *Nat Biotechnol*, 21:813–7, 2003.

[8] Meija J and Caruso JA. Deconvolution of isobaric interferences in mass spectra. *J Am Soc Mass Spectrom*, 15:654–658, 2004.

[9] Rosman KJR and Taylor PDP. Isotopic compositions of the elements 1997. *Pure Appl Chem*, 70:217–236, 1998.

[10] Pulfer M and Murphy RC. Electrospray mass spectrometry of phospholipids. *Mass Spectrom Rev*, 22:332–64, 2003.

[11] Han X and Gross RW. Shotgun lipidomics; electrospray ionization mass spectrometric analysis and quantitation of cellular lipidomes directly from crude extracts of biological samples. *Mass Spectrom Reviews*, 23:1–53, 2004.