Automatic Lipid Characterization Based on Charge Remote and Fatty Acid Fragmentation in MALDI MSMS

Ningombam Sanjib Meitei[1], Arun Apte[2], Jonggyu Park[5], Dietmar Waidelich[3], Fadi Abdi[4], Matthias Glueckmann[3]
[1]PREMIER Biosoft, Indore, India, [2]PREMIER Biosoft, Palo Alto, U.S.A., [3]AB SCIEX, Darmstadt, Germany, [4]AB SCIEX, Framingham, U.S.A, [5]NetWell Corporation, Tokyo, Japan

PREMIER Biosoft

Introduction

MALDI mass spectrometry has been used for detecting phosphatidylcholines by direct tissue analysis (1-4). Recently, triacylglycerols and phosphatidylcholine in complex mixtures have been identified using MALDI MS/MS (5,6) wherein charge remote fragmentation (CRF) as well as other A-, B-, C-, G and J- ions are the determinants as described recently (7,8,9). The CRF ions facilitate localization of double bonds and branching in fatty acid chains. However, the major challenge in MALDI MS/MS data analysis is the huge amount of data generated in the process. Annotation of the signals mentioned above requires time consuming manual spectral interpretation. SimLipid® software is a comprehensive informatics tool for characterizing lipids by MS/MS data which streamlines this type of data analysis.

Methods

Commercial standards and lipids from fats such as olive oil were analyzed using the AB SCIEX TOF/TOF™ 5800 system in MS and MS/MS mode prepared with DHB matrix in positive ion mode. MS/MS data was acquired using 1 keV collision energy and air as collision gas (5 x10^{-6} Torr). The automatic interpretation of MS/MS data as well as the structural identification of lipids are performed using SimLipid®, a software product of PREMIER Biosoft.

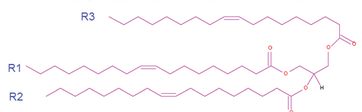

Figure 1. Structure of the TG(18:1(9Z)/18:1(9Z)/18:1(9Z))

SimLipid® is a comprehensive informatics tool for characterizing lipids using MS/MS data. For each candidate, in silico fragment ions are matched against the experimental MS/MS data. A proprietary scoring mechanism is developed in order to differentiate isobaric candidates.

SimLipid® database has been created containing 36299 lipids and 1305386 structure-specific in silico charge remote fragmentation (CRF) ions as well as other characteristic ions (7,8,9). Table 1 presents a list of prototype in silico fragments generated by SimLipid® for the lipid TG(18:1(9Z)/18:1(9Z)/18:1(9Z)) (Figure 1) and corresponding nomenclature used by the program to annotate MS/MS spectra.

Ion Type	Fragment Structure Displayed by SimLipid®	Description	SimLipid® Nomenclature
B - ion		Loss of one neutral fatty acid	M - <C:DB> - H where <C:DB> is the number of carbons and double bonds in the fatty acid. E.g., M – 18:1-H
C - ion		Loss of one sodium carboxylate residue	M - <C:DB > - Na-H E.g., M – 18:1-Na-H
G - ion		Loss of 2 fatty acids; one from position 2 and the other from either position 1 or position 3. It contains all 3 carbons from glycerol backbone. SimLipid® generates 3 G fragments by considering loss of any of the fatty acids.	Gk where k represent the fatty acid retained in the fragment structure. E.g., G3
J - ion		Loss of 2 fatty acids on position 1 and 3. It contains 2 carbons out of the 3 in glycerol backbone. In order to generalize the fragmentation pattern, SimLipid® generates 3 J fragments by considering loss of any of the 2 fatty acids unconstrained to position 1 and 3.	Jk where k represent the fatty acid retained in the fragment structure. E.g., J3
A - ion		Loss of a part of 1 fatty acid	F(Ri); i=1,2 & 3 and j can be any number between 2 to n where n = # carbons in the fatty acid chain – 2 - # double bonds in the fatty chain. E.g., F2(R1)
F(Ri)		Charge remote fragment ions	F(Ri); i=1,2 & 3 and j can be any number between 2 to n where n = # carbons in the fatty acid chain – 2 - # double bonds in the fatty chain.
Acylium ion		Protonated acylium ion (9,10)	RC=O$^+$ where R represents the alkyl group that is attached to the CO group with a single bond. Simply, it is the number of carbon atoms -1 present in the fatty acid. E.g., 17:1C=O$^+$

Table 1: Prototype in silico fragments generated by SimLipid® for the lipid TG(18:1(9Z)/18:1(9Z)/18:1(9Z)).

Results

MS and MS/MS data was acquired in positive ion mode. Figure 2 shows an example of a lipid standard showing different Glycerolipids [GL] of the lipid class Triradylglycerols [GL03] and sub Class Triacylglycerols [GL0301]. Interestingly, the triacylglycerols are all detected as sodiated molecular species. As typical for MALDI analysis, all molecular ions are detected as singly charged ions. For the Lipids analyzed here, the DHB matrix gave the best performance at the laser wavelength of λ = 354 nm as used here.

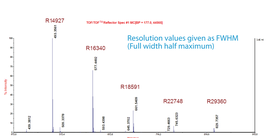

Figure 2. Example of MS data of a Triacylglycerol lipid standard acquired in positive ion mode using the AB SCIEX TOF/TOF™ 5800 system.

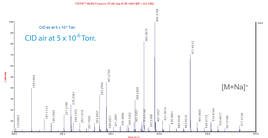

Figure 3. Example of MS/MS data of a Triacylglycerol lipid standard, precursor mass m/z 577.449, acquired in positive ion mode using the AB SCIEX TOF/TOF™ 5800 system. The MS/MS spectrum was acquired using 1kV lab frame collision energy.

Each of the triacylglycerols (TAGs) was subjected to MS/MS analysis as well (using 1 keV collision energy with air as collision gas). Figure 3 shows the MS/MS spectrum obtained for the TAG at m/z 577.5, which corresponds to the 30:0 fatty acid chain distribution with 3 fatty acids of the same length.

As shown in figure 4, the dominant fragments observed in the MALDI MS/MS spectrum are the result of the B- and the C-ions.

In Figure 5, another example is shown for the TAG 48:0, exhibiting the same fragmentation behavior. Note the CRF ion signals of lower intensity as well, which are separated by a mass difference of 14 Da, indicating the loss of a CH_2 group, which is a result of the cleavages of C-C bonds within the fatty acid chains. Odd fragment masses indicate saturated fatty acids, while the double bonds can be identified using the losses of 12 Da, indicating the cleavage of a C=C bond.

The MS/MS data of TAGs were subjected to batch analysis in the SimLipid® software. All the structures are identified correctly by the program.

Distinguishing Isobaric Candidate Structures using MS/MS Data

SimLipid® analysis of the MS/MS spectrum obtained for the TAG at m/z 577.4492 (known 30:0 fatty acid chain distribution with 3 fatty acids of the same length), results in 5 different candidate structures. The correct structure bearing the common name '1,2,3-tricaprinoyl-glycerol' is the highest scored lipid (Figure 6) followed by four different diacylglycerols that correspond to the 31:0 fatty acid chain distribution with 2 fatty acids of different lengths. Figure 8 shows the MS/MS spectrum (same as the spectrum depicted in the Figure 3) annotated by SimLipid® with fragments of (A) 1st ranked lipid 1,2,3-tricaprinoyl-glycerol and (B) 2nd ranked lipid DG(15:0\16:0\0:0)[iso2]. CRF ions region in (A) shows matched ions that are separated by a mass difference of 14 Da indicating the loss of a CH_2 group, which is a result of the cleavages of C-C bonds within the fatty acid chains. This pattern of ions indicates the absence of any double bond in the fatty acids of the identified structure. Manual spectrum interpretation in Figure 4 annotates the base peak that corresponds to the B – ion and C ions. As shown in Figure 8 (A), SimLipid® identifies and annotates these peaks automatically (with nomenclature M-10:0-H and M-10:0-Na-H). In addition, A, J, G and acylium ions are also identified by the program.

CRF ions region in (B) completely explains the two fatty acids of the 2nd ranked lipid DG(15:0\16:0\0:0)[iso2]. Note that the CRF ions F1(R1) to F13(R1) identifies the 15:0 fatty acid chain while F1(R2) to F14(R2) identifies the 16:0 fatty acid of the lipid structure. As the ions of Fj(R1) = Fj+1(R2) for j = 2,3,...13 for this structure, we have annotated only the CRF ions corresponding to R1 chain. Here, F7(R1) and F8(R2) are the A – ions. It is observed that the CRF ions region cannot distinguish the two structures. Hence, SimLipid® delves into other fatty acid chain specific characteristic ions. The program penalized the DG(15:0\16:0\0:0)[iso2] for its failure to generate fragments that match the high intensity peaks observed in the MS/MS spectrum as well as the typical characteristic ions that are determinants of the fatty acids. In addition, penalty is imposed on the structure that has typical characteristic ions but matched to low intensity peaks. For the structure DG(15:0\16:0\0:0)[iso2], B- ions that are characteristics of both the fatty acids 15:0 and 16:0 are observed and also one of the acylium ions that is characteristic of the fatty acid 15:0 are matched to peaks with very low intensity. However, typical characteristic peaks corresponding to C, J and G ions are not detected.

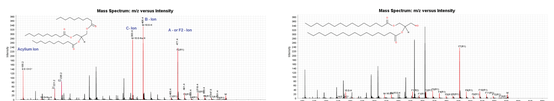

Figure 8. SimLipid® Software Interface: (A) MS/MS spectrum annotation of TAG 30:0 and (B) MS/MS spectrum annotation of 2nd ranked lipid DG(15:0\16:0\0:0)[iso2]. Fonts in blue color and the lipid structures on the MS/MS spectrum are manually written.

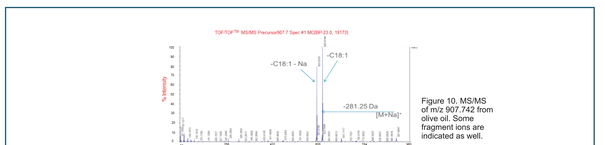

Figure 10. MS/MS of m/z 907.742 from olive oil. Some fragment ions are indicated as well.

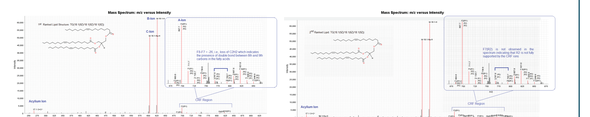

Figure 11. SimLipid® Software Interface: (A) MS/MS spectrum annotation of TG(18:1(9Z)/18:1(9Z)/18:1(9Z) (B) MS/MS spectrum annotation of 2nd ranked lipid TG(18:1(9Z)/18:1(9Z)/18:1(9Z). Font in blue color on the MS/MS spectra is manually added, while signals in red are automatically annotated by the software.

Acylium Ions in the MALDI-MS/MS Spectra

In ESI-mass spectrometry platforms, determination of the fatty acyl components that make up TAG and DAG molecular species requires MS^3 analysis of each of the major ESI - MS/MS product ions (10). Acylium ions, typical to MS^3 spectra, are the characteristic ions of fatty acids (10,11). A list of common fatty acyl groups and nominal masses (m/z) of corresponding ESI-MS^2 fragment ions (including acylium ions) has been provided in (10,11).

Our analysis shows that the MALDI-MS/MS with high-energy collision induced dissociation (HE-CID) capability is the superior ion dissociation technique for challenging structural analysis. Acylium ions are readily generated in MS/MS spectra itself. The observation of acylium ions in addition to CRF, G- and J-ions provides unparalleled information that facilitates accurate structural identification of complex lipid structures.

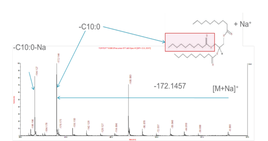

Figure 4. Example of MS/MS data interpretation of a Triacylglycerol TAG 30:0 lipid standard, precursor mass m/z 577.449, acquired in positive ion mode using the AB SCIEX TOF/TOF™ 5800 system. The spectrum is the same as figure 3, additionally indicating here the dominant loss of either the C10:0 fatty acid or the corresponding salt [C10:0+Na] as neutral loss.

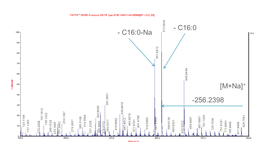

Figure 5. Example of MS/MS data obtained for a TAG 48:0 with three fatty acids of 16:0 composition. The dominant fragments corresponding to the neutral losses and the fatty acid 16:0 itself or the salt -[C16:0+Na] are indicated as well.

Figure 6 shows the typical identification result in the SimLipid® Software. The navigation window on the left displays the lipids identified in each MS/MS spectrum as the spectrum name is appended to the abbreviation/common name of the lipid identified. Search results are displayed in two panes: The search results display pane and the annotation pane. Here the results for the MS/MS of the precursor m/z 577.449 (TAG 30:0) are shown. SimLipid® ranks them on the basis of the peaks observed in the MS/MS data that correspond to diagnostic ions.

In Figure 7, the MS/MS Annotation tab of the matched m/z values for TAG 30:0 are shown in blue color. Additionally, the fragment structures can be displayed and the whole annotated table including each fragment structure can be exported as well.

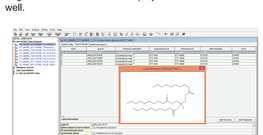

Figure 6: SimLipid® Software Interface: Search results for TAG 30:0. Four additional diacylglycerols that correspond to the 31:0 fatty acid chain distribution with 2 fatty acids of different lengths are also reported as lower ranked probable structures.

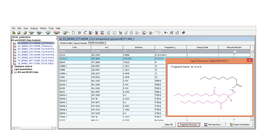

Figure 7. SimLipid® Software Interface: Search results for TAG 30:0, the MS/MS Annotation tab displays matched m/z values in blue and fragment structures can be displayed for a selected ion.

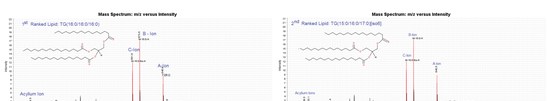

Figure 9. SimLipid® Software Interface: (A) MS/MS spectrum annotation of TG(16:0/16:0/16:0) (B) MS/MS spectrum annotation of 2nd ranked lipid TG(15:0/16:0/17:0)[iso 6]. Fonts in blue color on the MS/MS spectra are manually written.

Similarly, Figure 9 shows the annotated MS/MS spectrum of TAG 48:0 with three fatty acids of 16:0 composition. SimLipid® identifies the structure with abbreviation TG(16:0/16:0/ 16:0) correctly. From Figure 9 (A), it is observed that CRF ions identify the fatty acids correctly. B- and C- ions corresponding to the 16:0 fatty acid are observed as the highest intensity peaks in the spectrum. G- and J- ions are observed. Finally, acylium ion corresponding to 16:0 fatty acid is also observed. MS/MS spectrum annotation of 2nd ranked lipid TG(15:0/16:0/17:0)[iso 6] in Figure 9 (B) shows that CRF ions completely explain the three fatty acids. However, B- and C- ions corresponding to only 16:0 fatty acid are observed. Even though the B-ion corresponding to 17:0 is observed, the peak intensity is very low while the corresponding C-ion is missing. The B- and C-ions corresponding to 15:0 and J2 ion are not observed. Acylium ions corresponding to 16:0 and 17:0 are observed.

Identifying Unknown Lipids

In order to investigate the features of SimLipid® Software, we applied the technology on the analysis of lipids from different origins such as like olive oil brands, margarine and butter. The loadings plot constructed after Principal Component Analysis for the comparison of lipid fractions from different origins identifies the relevant m/z values which were responsible for variations in oil samples. In this case, the lipid with precursor m/z 907.742 was found to have elevated ion intensities for margarine and olive oils. Figure 10 shows a plot of the intensity profile for this precursor. The MS/MS spectrum was subjected to a search in SimLipid® Software which resulted in the identification of a TAG 54:3. The annotated MS/MS spectrum and the lipid structure are shown in Figure 11. In this case, CRF ions are the determinants for identifying the lipid structures. Note that acylium ion is also observed. Additional lipids could be identified within this workflow as well, which include e.g. m/z 689.75 from butter which were identified as Diradylglycerols and Diacylglycerophosphoglycerols, and from Margarine, ID of m/z 901.65, which lead to an unsaturated TAG as well (data not shown).

Conclusion

Structural analysis of a TAG mixture, including differentiation of regiosomers, without chromatographic separation, using MS/MS data generated by AB SCIEX TOF/TOF™ 5800 system (data acquired using 1kV lab frame collision energy) has been demonstrated. AB SCIEX TOF/TOF™ 5800 system supplies highly detailed fragmentation information for the structural elucidation of lipids that includes CRF-, B- , C-, J-, G- and acylium ions. It may be noted that acylium ions are important characteristic ions for identification of fatty acids in ESI mass spectrometry technique, although these ions are typical to MS^3 spectra.

SimLipid® software is a powerful informatics tool that facilitates automatic identification of structures of complex lipids mixture in batch mode. SimLipid® utilizes a database of lipids and corresponding structures based on in silico fragment ions (that includes A-, B-, C-, G-, J-, CRF- and acylium ions) for identification. For every precursor m/z, all the possible lipid structures are ranked based on SimLipid®'s proprietary search and scoring algorithm. Lipids randomly matching the precursor m/z are removed from the list of candidates by filtering out those lipids that do not have any of its characteristic fragments matching the product ions observed on the MS/MS spectra. In combination with statistical analysis routines such as PCA that allow for unbiased differentiation of lipid preparations, SimLipid® Software provides a rapid means of identifying relevant lipid species at high confidence. This was demonstrated identifying unknown lipids from fats, e.g. m/z 907.742, which was identified as TAG 54:3 within olive oil.

References

[1] Harvey, Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry of Phospholipids. J. Mass Spectrom. 1995, 30, 1333–1346.
[2] Marto, White, Seldomridge, Marshall, Structural Characterization of Phospholipids by Matrix-Assisted Laser Desorption/Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, Anal. Chem. 1995, 67,3979–3984.
[3] Al-Saad, Siems, Hill, Zabrouskov, Knowles, Structural Analysis of Phosphatidylcholines by Post-Source Decay Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. J. Am. Soc. Mass Spectrom. 2003, 14, 373–382. 8.
[4] Schiller, Sub, Arnhold, Fuchs, LeBig, Muller, Petkovic, Spalteholz, Zschornig, Arnold, Matrix- Assisted Laser Desorption and Ionization Time-of-Flight (MALDI-TOF) Mass Spectrometry in Lipid and Phospholipid Research. Prog. Lipid Res. 2004, 43, 449–488.
[5]. Ayumi Kubo, Takaya Satoh, Yoshiyuki Itoh, Masahiro Hashimoto, Jun Tamura, Robert B. Cody. Structural Analysis of Triacylglycerols by Using a MALDI TOF/TOF System with Monoisotopic Precursor Selection, J. Am. Soc. Mass Spectrom. (201) 24:6847689.
[6] Takaya Satoh, Ayumi Kubo, Shuichi Shimma, and Michisato Toyoda, Mass Spectrometry Imaging and Structural Analysis of Lipids Directly on Tissue Specimens by Using a Spiral Orbit Type Tandem Time-of-Flight Mass Spectrometer, SpiralTOF-TOF. Mass SPectrometry, 2012. DOI: 10.5702/massspectrometry.A0013
[7] 1. M. L. Gross. Charge-remote fragmentations: Method, mechanism and applications. Int. J. Mass Spectrom. Ion Process. 1992, 118/119, 137–165.
[8]. S. Trimpin, D. E. Clemmer, C. N. McEwen. Charge-remote fragmentation of lithiated fatty acids on a TOF-TOF instrument using matrix-ionization. J. Am. Soc. Mass Spectrom. 2007, 18, 1967–1972.
[9] Cheng. C., Gross, M. L., Pittenauer, E. Complete structural elucidation of triacylglycerols by tandem sector mass spectrometry . Anal. Chem. 1998, 70, 4417-4426.
[10] McAnoy A.M., Wu C.C., and Murphy R.C. Direct Qualitative Analysis of Triacylglycerols by Electrospray Mass Spectrometry Using a Linear Ion Trap. J Am Soc Mass Spectrom 2005, 16, 1498–1509
[11] Murphy R.C., James P.F., McAnoy A.M., Krank J., Duchoslav E., Barkley R.M. Detection of the abundance of diacylglycerol and triacylglycerol molecular species in cells using neutral loss mass spectrometry. Analytical Biochemistry 2007, 366, 59–70.

Trademarks/Licensing: For Research Use Only. Not for use in diagnostic procedures. SimLipid® is a registered trademark of PREMIER Biosoft. © 2014 PREMIER Biosoft.