Characterization of *Silybum marianum* triacylglycerol regioisomers using accurate mass quadrupole time of flight mass spectrometry

Thao L. Nguyen¹, Duma Hlangothi¹ and Mahmoud A. Saleh¹*

**Abstract:** Milk thistle (*Silybum marianum*) is a medicinal plant commonly used as a food supplement for liver protection against hepatitis C virus. The fruits are very rich in lipids (>30%); however, their chemical structures are not well characterized. In this investigation, dry fruits were extracted using CO₂ supercritical fluid extraction to obtain the crude lipid components. Triacylglycerols (TAGs) were isolated from the crude lipid extract using high-performance thin layer chromatography. Fatty acid composition of the purified TAGs was identified as their fatty acids methyl esters (FAMES) using gas chromatography mass spectrometry. Regioisomers of TAGs and their relative abundances were determined using accurate mass ultra-high-performance liquid chromatography quadrupole time of flight mass spectrometry both in the full scan and in MSMS electrospray ionization modes. Major identified fatty acids were 18:2, 18:1 and 16:0. Twenty-six different TAGs were identified with LLL, OLL, PLL, LOO and PLO as the major components.

**Subjects:** Analytical Chemistry; Natural Products; Chemical Spectroscopy

**Keywords:** milk thistle; QTOF; lipids; fatty acids; medicinal plants; liver protection; complementary alternative medicine (CAM)

1. Introduction

*Silybum marianum* (L.) Gaertn (Asteraceae), known as milk thistle, is an annual to biennial plant and grows widely in North African regions (Boulos, 2000). The fruit of milk thistle is one of the most important herbal liver medicines. The plant has been used for more than 2000 years to treat a range of liver and gallbladder disorders, including hepatitis, cirrhosis and jaundice (Cheilari, Sturm, Intelmann, Seger, & Stuppner, 2016). It was also shown to protect the liver against poisoning from chemical and environmental toxins, including snake bites, insect stings and mushroom poisoning (Pan, Lai, Dushenkov, & Ho, 2009). The medicinal part of the plant is the fruits (often referred to as “seeds”). The active constituents of milk thistle fruits are a mixture of flavonolignans, namely, silibinin (silybin A & B), isosilibinin (isosilybin A & B), silychristin and silidianin, collectively known as silymarin (Anthony & Saleh, 2012, 2013; Anthony, Subramanya, Uprichard, Hammouda, & Saleh, 2013).

**PUBLIC INTEREST STATEMENT**

*Silybum marianum* is also known as Milk thistle, Mary thistle, Saint Mary’s thistle and Mediterranean milk thistle. The plant has red to purple flowers and shiny pale green leaves with white veins. Milk thistle extract consists of about 65-80% silymarin (a flavonolignan complex) and 20-35% lipids. It has been used for treatment of liver disease, prevention and treatment of cancer and appeared to stimulate prolactin due to possible estrogenic activity. Most of the available scientific data on milk thistle is focused on silymarin components and very limited information about lipidd composition. This article describes in detail the chemical composition of lipids in the seeds of milk thistle and their potential nutritional values.
Lipids are essential metabolites in cells and they fulfill a variety of functions which include structural components of cellular membranes, energy storage, cell signaling and membrane trafficking (Akoh & Min, 2008; Nguyen, Aparicio, & Saleh, 2017; Welti et al., 2007). To understand the function of lipids in biological systems, it is important to profile the diverse lipid species present in any given organism and quantify their levels. Naturally occurring triacylglycerols (TAGs) represent a considerable challenge to the analytical chemist (Miroslav, Velínská, & Holčapek, 2009; Rustam & Reid, 2018; William, 2017). A typical animal fat or vegetable oil may contain up to 40 different TAG molecular species, while more complex mixtures such as milk fat or fish oil may comprise 150 or more individual species. The range of fatty acyl moieties present may include a wide range of carbon numbers and double bonds. In addition to determining which fatty acyl moieties are present in a TAG, it is also important to be able to determine their positions on the glycerol backbone of the molecule, since this has considerable significance from biochemical (Aoyama et al., 1996; Pufal, Quinlan, & Salter, 1995), nutritional (Martin, Bougnoux, Antoine, Lanson, & Couet, 1993; Zampelas, Williams, Morgan, Wright, & Quinlan, 1994) and biotechnological points of view (Rustam & Reid, 2018; Valenzuela & Nieto, 1994). TAGs can occur in nature as geometrical or stereoisomers (R&S) depending on the positions of the fatty acids on the glycerol backbone (normally denoted as sn-1, sn-2 or sn-3). TAGs containing two different acyl chains (XY) may exist as six (excluding enantiomers) regioisomers ($2^n - 2$) as XXY (chiral), YXX (chiral), YYX (chiral), XYY (achiral) and XYX(achiral) depending on which chain occupies the sn-2 position. Clinical studies have demonstrated that the location of acyl chain on the glycerol backbone affects lipid metabolism and showed that fatty acid at the sn-2 position is more bioavailable to rats and humans (Wakil, Mir, Mellor, Mellor, & Atkin, 2010). Therefore, structural analysis of TAGs is essential for development of dietary food supplements, natural health products and functional foods. Three methods that are currently used for identifying TAG regioisomers are as follows: method 1 is based on enzymatic hydrolysis of the TAGs to selectively remove acyl side-chains in specific positions followed by separation and determination of the various reaction products (Brockerhoff, 1965); method 2 uses silver-ion chromatography to resolve TAGs according to the carbon number, geometrical configuration and position of double bonds (Holčapek et al., 2010; Miroslav & Holčapek, 2013). Such complete separations require high efficiency and thus long columns, often with two or more columns connected in tandem, and long separation times, and method 3 uses chromatography, but does not require chromatographic separation of the regioisomers and relies on mass spectrometry to provide information on the individual isomers based on the selective abundance of certain fragment or product ions (Herrera, Ramaley, Potvin, & Melanson, 2013; Hsu & Turk, 2010). However, all the abovementioned methods used unit mass resolution instruments creating some uncertainty in identifying molecular ions and fragments.

In this investigation, we use accurate mass ultra-high-performance liquid chromatography quadrupole time of flight mass spectrometry to identify and quantitate TAG regioisomers of the fruits of *Silybum marianum*.

## 2. Materials and methods

### 2.1. Materials

*Silybum marianum* fruits were purchased from officially certified seed company (Frontier Natural Products CO-OP, 2012 Frontier, Norway, IA 52318, USA; Certified Organic Non-irradiated, USDA, patch # 9523128). All widespread use solvents and chemical reagents were purchased from VWR (Sugar Land, TX, USA). Liquid chromatography mass spectrometry (LCMS) water and solvents were purchased from J.T. Baker (Sugar Land, TX, USA). Isooctane and 1% boron trifluoride in methanol were
obtained from Sigma-Aldrich (Milwaukee, WI, USA). Analytical TAG standards: glyceryl trilinoleate (Beilstein Registry Number 1718698, EC Number 208-666-6, MDL number MFCD00042909, PubChem Substance ID 24900425), 1,2-dioleoyl-3-linoleoyl-rac-glycerol (CAS # 2190-20-7, MDL number MFCD00214281, PubChem Substance ID 24894336), 1,3-dipalmitoyl-2-oleoylglycerol (CAS # 2190-25-2, MDL number MFCD00673528, PubChem Substance ID 329764658), 1,2-dilinoleoyl-3-oleoyl-rac-glycerol (CAS # 2190-21-8, MDL number MFCD00083310, PubChem Substance ID 329751654), 1,3-dipalmitoyl-2-linoleoylglycerol (CAS # 2442-56-0, MDL number MFCD00214266 PubChem Substance ID 329767138), 1,2-dilinoleoyl-3-palmitoyl-rac-glycerol (CAS # 2190-15-0, MDL number MFCD00214267, PubChem Substance ID 24893225) and 1-palmitoyl-2-oleoyl-3-lino-
leoyl-rac-glycerol (CAS # 1587-93-5, MDL number MFCD00214386, PubChem Substance ID 24898953) were obtained from Sigma-Aldrich (USA).

2.2. Supercritical fluid extraction of lipids using CO$_2$
About 100 g of dry seed powder were extracted using CO$_2$ supercritical fluid performed with Model SFT-100XW (Proras, Rome, Italy) without modifier. The pressure was set at 8700 psi, oven temperature at 40°C, restrictor temperature at 80°C, CO$_2$ flow rate at 10 mL/min, and 9 static and dynamic steps at 10 min cycles yielding 20 g of a light-yellow oil.

2.3. High-performance thin layer chromatography (HPTLC)
TAG fraction of the crude lipid extract was obtained using Camag automated HPTLC instrument (CAMAG USA, Denver, CO) on normal phase plates (Merk KGaA Silica gel 60 F254s plates, Lot#: HX67519396) developed in hexane, diethyl ether, acetic acid (70:30:1 v/v).

2.4. Preparation of fatty acids methyl esters (FAMES)
The transesterification was carried as we described earlier (Nguyen, Aparicio, & Saleh, 2015) by adding 10 mg of the purified TAGs to 2 mL of 2% BF$_3$ in methanol, mixed by vortex and placed on the heating block at 75°C for 1 h. After cooling down to room temperature, 1 mL of saturated sodium chloride solution was added to stop the reaction and FAMES were extracted in 2 mL of isoctane, passed through anhydrous sodium sulfate and transferred to gas chromatography (GC) vials for gas chromatography mass spectrometry (GCMS) analysis. All experiments were conducted in three replicates.

2.5. Analysis of FAMES by GCMS
GCMS analysis of the FAMES was carried out using Agilent Technologies 7890-B gas chromatograph equipped with 5977A mass spectrometer. GC separation was done using a BPX90 SGE Analytical Science column 100 m × 0.25 mm × 0.25 μm column (SGE Analytical Science, Austin, TX, USA). Average linear velocity of 21 cm/s with a holdup time of 7.934 min. Temperature program started at 150°C held for 10 min and heated up to 230°C at a rate of 2°C/min; total runtime was 50 min; solvent delay was 9 min; and equilibration time was 0.25 min. Mass spectral acquisitions were performed in electron ionization (EI) mode. The post-run analyses were performed by Mass Hunter software Qualitative Analysis B.07.00 SP1. FAMES were identified based on their retention times and spectra. The mass spectral data were analyzed using Wiley10/NIST mass spectral database.

2.6. LC-QTOF
Agilent 6530 accurate mass liquid chromatography quadrupole time-of-flight mass spectrometry (LC-QTOF) system equipped with 1290 binary pump, Phenomenex Kinetex C18 100 mm × 3.0 mm 2.6 μm column. Solvent A1 was made of 90% acetonitrile, 10% water, 5.5 mM ammonium formate; B1 was 90% isopropanol, 10% water, 5.5 mM ammonium formate, run in a gradient protocol starting at 50% A1, 50% B1 with a flow 0.15 mL/min reaching 100% solvent B1 at 25.0 min from the start, and held at 100% B1 for additional 5 min. Injection volume was 1 μL, with a total runtime of 33 min. The LC-QTOF instrument was operated under the following conditions: The Dual Agilent Jet Stream ESI source was operated in positive mode, and instrument parameters were set as follows: sheath gas temperature, 350°C; sheath gas flow, 11 L/min; nebulizer, 50 psi; dry gas temperature, 300°C; dry gas flow, 5 L/min; and
capillary entrance voltage, 3500 V. Fragmentor and Skimmer1 were operated at 175 and 65 V, respectively. The MS scan data were collected at a rate of 1.0 spectra/s in the range 100–1500 m/z. All the MS data were collected with MassHunter Data Acquisition B.07.00 (Agilent Technologies) and MassHunter Qualitative Analysis B.07.00.

2.7. LC-QTOF MS/MS experiment
Accurate mass molecular ions [M+NH₄]⁺ for each individual predicted TAG were subjected to an MS/MS experiments at collision energies of 5, 10, 15, 20, 25, 30, 35 and 40 eV using the following conditions: dual AJS ESI positive-ion source, source gas temperature at 300°C; drying gas 5.0 L/min; nebulizer 50 psig; sheath gas temperature 350°C; sheath gas flow 11.0 L/min; V cap 3500 V; capillary 5.941 μA; nozzle voltage 1000 V; chamber 6.27 μA; MS/MS mass range 100–1500 m/z; acquisition rate 1 spectra/s; acquisition time 1000 MS/spectrum; 9783 transitions/spectrum; auto recalibration reference mass window, detection window 100 ppm; minimum height 1000 counts; polarity type: positive; offset: 5; MSMS absolute threshold 5; reference threshold 0.01%.

3. Results

3.1. Fatty acid composition
The fatty acid composition of milk thistle TAGs was analyzed in their methyl ester forms using GCMS technique as described in the experimental section (2.5). Baseline separation of all the detected fatty acids was achieved as shown in Figure 1. Identification and relative percentage quantitative analysis of the fatty acids are shown in Table 1. Unsaturated fatty acids constituted more than 60% of the total fatty acids. Among the unsaturated fatty acids, linoleic acid is the major constituent and forms 45.31% of the total composition followed by oleic acid (14.29%). Saturated fatty acids, namely, palmitic acid, 22.34%, and stearic acid, 8.07%, were the main saturated constituents.

3.2. Triacylglycerol composition
TAG composition in the purified lipid extract of S. marianum fruits was invigilated using the accurate mass LC-QTOF system. Eight major HPLC peaks were observed and were further quantitatively analyzed by accurate mass target deconvolution module of the Masshunter software to reveal that those peaks were clusters of one or more different TAGs. All clusters were found to be positive ammonium adducts [M+NH₄]⁺ ions, and their molecular formula was calculated based on their accurate masses. Fatty acid composition of each detected TAG peaks was identified by searching their accurate mass using Lipid Maps Database and Human Metabolome Database. Although the ion [M+NH₄]⁺ with the accurate mass equivalent to 896.7707 represents 66 possible TAG isomers, this number can be reduced to 6, by eliminating TAGs that have any
fatty acid that was not detected in the oil (Table 2). This concept was applied to all the detected peaks in eliminating TAGs that have any fatty acid that was not detected in the oil (Figure 2 and Table 2). Twenty-six possible TAG structures were predicted based on their accurate molecular weights and fatty acid composition; however, the complete conformation of the actual structure cannot be confirmed until MSMS experiments are performed.

### 3.3. Determination of regioisomers by MSMS

To establish a protocol for using accurate mass MSMS experiments for elucidation of the regioisomers of TAGs, we selected five commercially available authentic TAGs with confirmed configuration to test our prediction accuracy. These standards were selected to present different types of configuration with examples of XXX, XXY, YYX, XYX and XYZ, where X, Y and Z represent different fatty acids. In the full-scan MS mode, all of the individual standard showed only the molecular ion as \([\text{M}+\text{NH}_4]^+\) with no fragments. For each standard, the selected molecular ion \([\text{M}+\text{NH}_4]^+\) was selected by the QTOF system and subjected to collision ionization.

### Table 1. Fatty acid composition of the triacylglycerols (TAGs) of *Silybum marianum*

| RT (min) | %* | Shorthand | Common name                  | IUPAC name                                         |
|---------|-----|-----------|--------------------------------|-----------------------------------------------------|
| 15.58   | 0.14| 14:0      | Myristic acid                  | Tetradecanoic acid                                  |
| 18.49   | 0.10| 15:0      | NA                            | Pentadecanoic acid                                  |
| 20.05   | 22.34| 16:0      | Palmitic acid                  | Hexadecanoic acid                                  |
| 21.04   | 0.51| 16:1 n-7 | Palmitoleic acid (Z)-hexadec-9-enoic acid |
| 25.45   | 8.07| 18:0      | Stearic acid                   | Octadecanoic acid                                  |
| 27.13   | 14.29| 18:1 n-9 | Oleic acid                     | (Z)-octadec-9-enoic acid                            |
| 29.84   | 45.31| 18:2 n-6 | Linoleic acid                  | (9Z,12Z)-octadec-9,12-dienoic acid                 |
| 31.15   | 3.32| 20:0      | Arachidic acid                 | Icosanoic acid                                     |
| 32.70   | 0.30| 18:3 n-3 | α-Linolenic acid               | (9Z,12Z,15Z)-octadec-9,12,15-trienoic acid          |
| 36.77   | 1.70| 20:1 n-11| Gadoleic acid                  | (Z)-icos-9-enoic acid                              |
| 39.42   | 0.23| 22:0      | Behenic acid                   | Docosanoic acid                                    |
| 42.32   | 0.02| 24:0      | Lignoceric acid                | Tetracosanoic acid                                 |

*Results are shown as the average of three replicates.
MSMS experiment and detected the exact masses of the product ions. The results for each individual standard were as follows.

| ID # | RT (min) | %* | Accurate mass | Δ mass PPM | Molecular formula | No. of possible TAG isomers | No. of TAGs according to fatty acids found |
|------|----------|----|---------------|------------|-------------------|-----------------------------|------------------------------------------|
| 1    | 6.98     | 30.26 | 896.7707     | 0.6        | C_{57}H_{102}NO_{6} | 66                          | 6                                        |
| 2    | 7.95     | 0.17  | 846.7555     | 1.1        | C_{53}H_{100}NO_{6} | 48                          | 7                                        |
| 3    | 8.11     | 8.78  | 872.7701     | 0.1        | C_{53}H_{102}NO_{6} | 63                          | 8                                        |
| 4    | 8.15     | 17.73 | 898.7858     | 0.0        | C_{53}H_{100}NO_{6} | 72                          | 7                                        |
| 5    | 9.35     | 1.09  | 848.7709     | 0.8        | C_{53}H_{102}NO_{6} | 46                          | 6                                        |
| 6    | 9.35     | 6.02  | 874.7709     | 24.5       | C_{53}H_{102}NO_{6} | 59                          | 8                                        |
| 7    | 9.47     | 11.35 | 900.8017     | 0.2        | C_{53}H_{104}NO_{6} | 70                          | 9                                        |
| 8    | 9.52     | 1.09  | 926.8173     | 0.2        | C_{53}H_{104}NO_{6} | 73                          | 5                                        |
| 9    | 10.62    | 0.45  | 850.7863     | 0.6        | C_{53}H_{102}NO_{6} | 38                          | 6                                        |
| 10   | 10.69    | 1.6   | 876.8014     | 0.1        | C_{53}H_{102}NO_{6} | 52                          | 7                                        |
| 11   | 10.79    | 4.45  | 902.8167     | 0.4        | C_{53}H_{104}NO_{6} | 67                          | 14                                       |
| 12   | 11.06    | 2.85  | 928.8179     | 16.0       | C_{53}H_{104}NO_{6} | 69                          | 8                                        |
| 13   | 12.09    | 0.41  | 878.8171     | 0.0        | C_{53}H_{102}NO_{6} | 39                          | 7                                        |
| 14   | 12.2     | 2.02  | 904.8327     | 0.0        | C_{53}H_{102}NO_{6} | 50                          | 9                                        |
| 15   | 12.33    | 2.92  | 930.8481     | 0.3        | C_{53}H_{104}NO_{6} | 57                          | 9                                        |
| 16   | 12.56    | 3.04  | 956.8633     | 0.8        | C_{53}H_{104}NO_{6} | 56                          | 6                                        |
| 17   | 13.57    | 0.33  | 906.8484     | 0.0        | C_{53}H_{112}NO_{6} | 36                          | 8                                        |
| 18   | 13.68    | 0.75  | 932.8638     | 0.3        | C_{53}H_{114}NO_{6} | 45                          | 7                                        |
| 19   | 13.8     | 1.20  | 958.8793     | 0.4        | C_{53}H_{114}NO_{6} | 47                          | 7                                        |
| 20   | 14.04    | 0.70  | 984.8957     | 0.3        | C_{53}H_{118}NO_{6} | 41                          | 1                                        |
| 21   | 14.97    | 0.21  | 934.8801     | 0.4        | C_{53}H_{116}NO_{6} | 31                          | 8                                        |
| 22   | 15.07    | 0.93  | 960.8952     | 0.2        | C_{53}H_{118}NO_{6} | 35                          | 6                                        |
| 23   | 15.24    | 0.45  | 986.9102     | 0.8        | C_{53}H_{120}NO_{6} | 34                          | 5                                        |
| 24   | 16.34    | 0.27  | 1017.8747    | 1.0        | C_{53}H_{126}NO_{6} | 31                          | 3                                        |
| 25   | 16.44    | 0.08  | 988.9270     | 0.3        | C_{53}H_{122}NO_{6} | 25                          | 4                                        |
| 26   | 16.48    | 0.51  | 993.9217     | 0.6        | C_{53}H_{122}NO_{6} | 34                          | 4                                        |

*Results are shown as the average of three replicates.

Table 2. TAG composition and their normalized % of the neutral lipid fraction of *S. marianum*
3.3.1. Glyceryl trilinoleate (182182182)
This standard represents the type of TAG which is made of only one fatty acid (XXX). The molecular ion [M+NH₄]+ has an accurate mass of 896.7708 and its MSMS product ions are shown in Figure 3. The [M+NH₄]+ molecular ion lost a neutral NH₃ to give protonated molecular ion [M+H]+ with an accurate mass of 879.7167 m/z. Since the TAG is made of three identical fatty acid, only one diacyl product ion was detected at 599.4848 m/z corresponding to the loss of a neutral 18:2 RCOOH from protonated molecular ion or the loss of 297.2583 corresponding to the loss of neutral 18:2 RCOONH₄ from [M+NH₄]+. Other product ions with less abundance but highly important in stereoisomers assignments were ions corresponding to 18:2 [RCO-H₂O]+, [RCO]+, [RCO + 74]+ and [RCO + 74-H₂O]+ with accurate masses of 245.2190, 263.2358, 337.2637 and 319.2550 m/z, respectively. The product ion at 261.2208 m/z represents the typical loss of the middle fatty acid at the sn-2 position as the α,β-unsaturated acid [18:2 RCH₂CH = CH-CO]+ which is very useful in identifying fatty acids at the 2-position.

3.3.2. 1,2-Dioleoyl-3-linoleoyl-rac-glycerol (181181182)
This standard represents the type of TAGs which is made of two different fatty acids (XXY) which are chiral isomers; however, the one used in this work is the racemic mixture. The molecular ion [M+NH₄]+ has an accurate mass of 900.8015 m/z and its MSMS product ions are shown in Figure 5. The TAG [M+NH₄]+ parent ion undergoes the neutral loss of ammonia or a
fatty acyl group and ammonia, yielding a corresponding protonated molecular ion \([M + H]^+\) at 883.1142 m/z and DAG fragment ions, respectively. The product ion of 603.5335 m/z corresponded to the loss of 18:2 fatty acid, whereas the product ion of 601.5189 m/z corresponded to the loss of 18:1 fatty acid. This is consistent with our previous work that 18:1 and 18:2 acids are located at sn1 and sn3 positions, respectively (Nguyen et al., 2015). If the fatty acid 18:2 was in the sn2 position, an ion corresponding to \(\alpha,\beta\)-unsaturated acid 18:2 \([18:2\ RCH=CH\ \text{CO}]^+\) will be shown at 261.2205 m/z which is not found in Figure 4. Other product ions with less abundance but highly important in stereoisomers assignments were ions corresponding to 18:2 \([\text{RCO-H}_2\text{O}]^+,\ \text{18:1 [RCO-H}_2\text{O}]^+,\ \text{18:2 [RCO]}^+,\ \text{18:2 [RCO+74]}^+,\ \text{18:1[RCO+74]}^+,\ \text{18:2[RCO+74-H}_2\text{O}]^+\) with accurate masses of 245.2190, 247.2392, 263.2355, 265.2521, 337.2637, 339.2888, 319.2550 and 321.2750 m/z, respectively. The product ion at 339.2883 m/z corresponding to 18:1 [RCO+74] is much more abundant than the ion at 337.2637 corresponding to 18:2 [RCO+74] showing that the 18:1 fatty acid not 18:2 is located at the sn2 position.

### 3.3.3. 1,2-Dilinoleoyl-3-oleoyl-rac-glycerol (182182181)

This standard represents the type of TAGs which are made of two different fatty acids (YYX) which are chiral isomers; however, the one used in this work is the racemic mixture. The molecular ion \([M+\text{NH}_4]^+\) has an accurate mass of 898.7547 m/z and its MSMS product ions are shown in Figure 5. The TAG \([M+\text{NH}_4]^+\) parent ion undergoes the neutral loss of ammonia or a fatty acyl group and ammonia, yielding a corresponding protonated molecular ion \([M + H]^+\) at 881.7579 m/z and DAG fragment ions, respectively. The product ion of 601.5180 m/z corresponded to the loss of 18:2 fatty acid, whereas the product ion of 599.5028 m/z corresponded to the loss of 18:1 fatty acid. This is consistent with our previous work that 18:2 and 18:1 acids are located at sn1 and sn3 positions, respectively (Nguyen et al., 2015). It can also be seen that loss of fatty acids from sn1 position gave a higher abundant ion than the one produced from sn3 position. The fatty acid 18:2 is in the sn2 position, an ion corresponding to \(\alpha,\beta\)-unsaturated acid 18:2 \([18:2\ RCH=CH\ \text{CO}]^+\) is shown at 261.2205 m/z which is not found in Figure 4. Other product ions with less abundance, but highly important in stereoisomers assignments were ions corresponding to 18:2 \([\text{RCO-H}_2\text{O}]^+,\ \text{18:1 [RCO-H}_2\text{O}]^+,\ \text{18:2 [RCO]}^+,\ \text{18:2 [RCO+74]}^+,\ \text{18:1[RCO+74]}^+,\ \text{18:2[RCO+74-H}_2\text{O}]^+\) with accurate masses of 245.2190, 247.2392, 263.2355, 265.2521, 337.2637, 339.2888, 319.2550 and 317.2456 m/z, respectively. The product ion at 339.2883 m/z corresponds to 18:1 [RCO+74] and the ion at 337.2729 corresponds to 18:2 [RCO+74].
3.3.4. 1,2-Dilinoleoyl-3-palmitoyl-rac-glycerol (182182160)

This standard represents TAG made of two different fatty acids (YYZ) which are chiral isomers; however, the one used in this work is the racemic mixture. The molecular ion [M+NH₄]⁺ has an accurate mass of 872.7700 m/z and its MSMS product ions are shown in Figure 6. The TAG [M+NH₄]⁺ parent ion undergoes the neutral loss of ammonia or a fatty acyl group and ammonia, yielding a corresponding protonated molecular ion [M + H]⁺ at 855.7432 m/z and DAG fragment ions, respectively. The product ion of 599.5033 m/z corresponded to the loss of 16:0 fatty acid, whereas the product ion of 575.5033 m/z corresponded to the loss of 18:2 fatty acid. This is consistent with our previous work that 18:2 and 16:0 acids are located at sn1 and sn3 positions, respectively (Nguyen et al., 2015). The fatty acid 18:2 is in the sn2 position, an ion corresponding to α,β-unsaturated acid 18:2 [18:2 RCH₂CH=CH-CO]⁺ is shown at 261.2210 m/z. Other product ions with less abundance but highly important in stereoisomer assignments were ions corresponding to 18:2 [RCO-H₂O]⁺, 18:2 [RCO]⁺, 16:0 [RCO]⁺, 18:2 [RCO+74], 16:0[RCO+74], 18:2[RCO+74-H₂O]⁺ and 16:0[RCO+74-H₂O]⁺ with accurate masses of 239.2368, 245.2190, 247.2392, 263.2355, 265.2550, 337.2637, 339.2888, 319.2550 and 317.2456 m/z, respectively. The product ion at 313.2736 m/z corresponds to 16:0[RCO+74] and the ion at 337.2730 corresponds to 18:2 [RCO+74].

3.3.5. 1-Palmitoyl-2-oleoyl-3-linoleoyl-rac-glycerol (160181182)

This standard represents TAG made of three different fatty acids (XYZ) which are chiral isomers; however, the one used in this work is the racemic mixture. The molecular ion [M+NH₄]⁺ has an
accurate mass of 874.7854 m/z and its MSMS product ions are shown in Figure 7. The TAG 
[M+NH4]+parent ion undergoes the neutral loss of ammonia or a fatty acyl group and ammonia,
yielding a corresponding protonated molecular ion [M + H]+ at 857.7576 m/z and DAG fragment
ions, respectively. Loss of the acyl group from sn1 and sn2 positions as expected produced highly
abundant ions (601.5191 m/z from the loss of 16:0 fatty acid at the sn1 position and 577.5183 m/z
from the loss of 18:2 at the sn3 position). The loss of the 18:1 fatty acid from the sn2 position gave
a lower abundant ion of 575.5035 m/z as expected. The formation of the
\(\alpha,\beta\)-unsaturated acid 18:1 from the 18:1 fatty acid at sn2 position cannot be detected because it
is isobaric with the 18:2 fatty acid at sn3 position. Other product ions with less abundance but
highly important in stereoisomers assignments were ions corresponding to 16:0 \([RCO-H_2O]+\)
(239.2362 m/z), 16:0 \([RCO+74]+\) (313.2734 m/z), 18:2 \([RCO]+\) (263.2364 m/z), 18:1 \([RCO]+\)
(265.2512 m/z), 18:2 \([RCO+74]\) (337.2715), 18:1 \([RCO+74]\) (339.2883), 18:2 \([RCO+74-H_2O]+\)
(319.2631).

4. Discussion
Based on what was found from accurate mass MSMS experiments using five different authentic
standards, TAG regioisomers can be predicted as follows:

| ID # | Shorthand name | IUPAC chemical name |
|------|----------------|---------------------|
| 1    | 18:2/18:2/18:2 | 1,2,3-Tri-(9Z,12Z-octadecadienoyl)-glycerol |
| 2    | 14:0/18:1/18:2 | 1-Tetradecanoyl-2-(9Z-octadecenoyl)-3-(9Z,12Z-octadecadienoyl)-sn-glycerol |
| 3    | 16:0/18:2/18:2 | 1-Hexadecanoyl-2,3-di-(9Z,12Z-octadecadienoyl)-sn-glycerol |
| 4    | 18:1/18:2/18:2 | 1-(9Z-Octadecenoyl)-2,3-di-(9Z,12Z-octadecadienoyl)-sn-glycerol |
| 5    | 16:0/16:0/18:2 | 1,2-Dihexadecanoyl-3-(9Z,12Z-octadecadienoyl)-sn-glycerol |
| 6    | 16:0/18:2/18:1 | 1-Hexadecanoyl-2-(9Z,12Z-octadecadienoyl)-3-(9Z-octadecenoyl)-3-sn-glycerol |
| 7    | 18:2/18:1/18:1 | 1,(9Z,12Z-Octadecadienoyl) 2,3-di-(9Z-octadecenoyl)-3-sn-glycerol |
| 8    | 20:1/18:1/18:2 | 1-(11Z-Eicosenoyl),2-3-di-(9Z,12Z-octadecadienoyl)-3-sn-glycerol |
| 9    | 16:0/16:0/18:1 | 1,2-Dihexadecanoyl-3-(9Z-octadecenoyl)-sn-glycerol |
| 10   | 16:0/18:1/18:2 | 1-Hexadecanoyl-2-(9Z,12Z-octadecadienoyl)-3-octadecanoyl-sn-glycerol |
| 11   | 18:0/18:1/18:2 | 1-Octadecenoyl-2-(9Z-octadecenoyl)-3-(9Z,12Z-octadecadienoyl)-sn-glycerol |
| 12   | 18:2/18:2/20:0 | 1,2-Di-(9Z,12Z-octadecadienoyl)-3-eicosanoyl-sn-glycerol |
| 13   | 16:0/18:0/18:1 | 1-Hexadecanoyl-2-octadecanoyl-3-(9Z-octadecenoyl)-sn-glycerol |
| 14   | 16:0/18:2/20:0 | 1-Hexadecanoyl-2-(9Z,12Z-octadecadienoyl)-3-eicosanoyl-sn-glycerol |
| 15   | 18:2/18:1/20:0 | 1-(9Z,12Z-Octadecadienoyl) 2,2-(9Z-octadecenoyl)-3-eicosanoyl-sn-glycerol |
| 16   | 18:2/18:2/22:0 | 1,2-Di-(9Z,12Z-octadecadienoyl)-3-docosanoyl-sn-glycerol |
| 17   | 16:0/18:1/20:0 | 1-Hexadecanoyl-2-(9Z-octadecenoyl)-3-eicosanoyl-sn-glycerol |
| 18   | 18:1/18:1/20:0 | 1,2-Di-(9Z-octadecenoyl) 3-eicosanoyl-sn-glycerol |
| 19   | 18:1/18:2/22:0 | 1-(9Z-Octadecenoyl)-2-(9Z,12Z-octadecadienoyl)-3-docosanoyl-sn-glycerol |
| 20   | 18:3/20:1/22:0 | 1-(9Z,12Z,15Z-Octadecatrienoyl) 2,2-(11Z-eicosanoyl)-3-docosanoyl-sn-glycerol |
| 21   | 16:0/18:1/22:0 | 1-Hexadecanoyl-2-(9Z-octadecenoyl)-3-docosanoyl-sn-glycerol |
| 22   | 18:1/18:1/22:0 | 1,2-Di-(9Z-octadecenoyl)-3-docosanoyl-sn-glycerol |
| 23   | 20:1/20:1/20:1 | 1,2,3-Tri-(11Z-eicosanoyl)-glycerol |
| 24   | 20:1/20:1/22:0 | 1,2-Di-(11Z-eicosanoyl)-3-docosanoyl-sn-glycerol |
| 25   | 20:0/20:1/20:1 | 1-Eicosanoyl-2,3-di-(11Z-eicosanoyl)-sn-glycerol |
| 26   | 20:0/20:0/20:0 | 1,2,3-Trieicosanoyl-sn-glycerol |
(1) Fatty acids at sn1 and sn3 are preferably lost from the molecular ion than the fatty acids in the sn2 position producing DAG ions with higher abundances from losses of fatty acids than the DAG produced from the loss of fatty acid at the sn2 position. Steric and electronic energy favor the loss of fatty acids from the sn1 and sn3 where both will give an intermediate of six-member ring; however, the loss of fatty acid from sn2 position gives an intermediate of 5-membered ring, as shown below.

(2) If the fatty acids in the sn1 and sn3 positions are the same, the TAG is achiral, and the two positions are identical and cannot be distinguished from each other. However, if the two fatty acids are different, we found that based on this work and previously reported results (Nguyen, Aparicio, & Saleh, 2015, 2017) that loss of fatty acid at the sn1 position gave relatively higher abundant DAG ion than the one produced from the loss of fatty acid at the sn3 position.

(3) Fatty acids at the sn2 position tend to be lost as α,β-unsaturated acyl and therefore introducing additional double bonds (i.e. 18:0, 18:1 and 18:2 are lost as 18:1, 18:2 and 18:3, respectively).

Figure 8. MSMS product ions and fragmentation pattern of 1-Palmitoyl-2-oleoyl-3-linoleoyl-rac-glycerol.
(4) Each fatty acid in the TAG produces an RCO+ ion, if two or three of the chains are different, two or three different RCO+ are found, but more for the sn1 and sn3 positions. Another ion of relatively high intensity is that corresponding to the acyl ion plus 74 amu equivalent to C3H6O2 corresponds to the glyceryl moiety [RCO+74] + ion as shown in Figure 8.

Applying the abovementioned rules for fragmentation patterns with accurate mass of product ions, all the individual TAGs of S. marianum were identified as listed in Table 3. Linoleic, oleic, and palmitic acids were the most abundant fatty acids found in the identified TAGs, which is consistent in the relative percentages of total fatty acids revealed from the GCMS analytical results of the total FAMES analysis. The single 1,2,3-tri-(9Z,12Z-octadecadienoyl)-glycerol (18:2/18:2/18:2) TAG constituted 30.26% of the total identified TAGs. The majority of the fatty acids located in the sn2 position were the unsaturated acids with linoleic acid being the predominant acid in the sn2 position suggesting that the S. marianum oil possess a good nutritional property since it is known that fatty acids in the sn2 position the TAG are the most bioavailable acids compared to the one in the sn1 or sn3 positions.

Accurate mass MSMS analysis using LCQTOF provides an accurate and straightforward procedure for designated regioisomers of TAG. Accurate mass provided a more precise way of eliminating many isobaric ions (with unit mass resolution) which require further MSMS experiments (MSMS2 and MSMS3).

**Funding**

This publication was made possible in part by research infrastructure support from grant number 2G12MD007605 from the NIMHD/NIH.

**Author statement**

Our research group under the leadership of Professor Mahmoud A. Saleh are conducting laboratory investigations on the effect of chronic exposures to environmental pollutants (chemical, physical and biological) which were found to increase reactive oxygen and nitrogen species and ultimately initiate oxidative damage. Advance analytical techniques are used to search for natural product chemicals such as the ones that we found in *Silybum marianum*, which are capable of counteracting oxidative damage.

**Competing interests**

The author declares no competing interests.

**Author details**

Thao L. Nguyen
E-mail: nguyenthao190@yahoo.com
Duma Hlangothi
E-mail: dhlangothi@gmail.com
Mahmoud A. Saleh
E-mail: saleh ma@tsu.edu

1 Department of Chemistry, Texas Southern University, Houston, TX, USA.

**Citation information**

Cite this article as: Characterization of *Silybum marianum* triacylglycerol regioisomers using accurate mass quadrupole time of flight mass spectrometry, Thao L. Nguyen, Duma Hlangothi & Mahmoud A. Saleh, Cogent Chemistry (2018), v: 4: 1477246.

**References**

Akoh, C., & Min, D. B. (2008). Food lipids. In Chemistry, nutrition, and biotechnology (3th ed.). New York: Marcel Dekker Inc.

Anthony, K., & Saleh, M. A. (2013). Chemical profiling and antioxidant activity of commercial milk thistle food supplements. Journal Chemical Pharmaceutical Researcher, 4, 4440–4450.

Anthony, K., & Saleh, M. A. (2013). Free radical scavenging and antioxidant activities of silymarin components. Antioxidants, 2, 398–407. doi:10.3390/antioxi2040398

Anthony, K., Subramanya, G., Uprichard, S., Hammouda, F., & Saleh, M. A. (2013). Antioxidant and anti-hepatitis C viral activities of commercial milk thistle food supplements. Antioxidants, 2, 23–36. doi:10.3390/antioxi2010023

Aoyama, T., Fukui, K., Taniguchi, K., Nagoaka, S., Yamamoto, T., & Hashimoto, Y. (1996). Absorption and metabolism of lipids in rats depends on fatty-acid isomeric position. The Journal of Nutrition, 126, 225–231. doi:10.1093/jn/126.1.225

Boulos, L. (2000). Flora of Egypt (1st ed.). Cairo: Al Hadara Publishing Inc.

Brockerhoff, H. A. (1965). Stereospecific analysis of triglycerides. Journal of Lipid Research, 6, 10–15.

Cheilari, A., Sturm, S., Intelmann, D., Seger, C., & Stuppner, H. (2016). Head-to-head comparison of ultra-high-performance liquid chromatography with diode array detection versus quantitative nuclear magnetic resonance for the quantitative analysis of the Silymarin complex in *Silybum marianum* fruit extracts. Journal of Agricultural and Food Chemistry, 64, 1618–1626. doi:10.1021/acs.jafc.5b05494

Csupora, D., Csorba, A., & Hohmann, J. (2016). Recent advances in the analysis of flavonolignans of *Silybum marianum*. Journal Pharmaceutical Biomedical Analysis, 130, 301–317. doi:10.1016/j.jpba.2016.05.034

Evans, W. C. (2002). *Trease and Evans pharmacognosy* (15th ed.). New Delhi: Reed Elsevier India Pvt. Ltd.

Herrera, L. C., Ramaley, L., Patvin, M. A., & Melanson, J. E. (2013). A method for determining regioisomer abundances of polyunsaturated triacylglycerols in omega-3 enriched fish oils using reversed-phase liquid chromatography and triple-stage mass spectrometry. Food Chemistry, 139, 655–662. doi:10.1016/j.foodchem.2012.12.059

Hlangothi, D., Abdel-Rahman, F., Nguyen, T., Anthony, K., & Saleh, M. A. (2016). Distribution of Silymarin in the Fruit of *Silybum marianum*. Pharmaceutical Analysis Acta, 7, 11–19.
Holc'apek, M., Dvort'oková, H., Lisa, M., Girón, A. J., Sandra, P., & Cvc'ak'a, J. (2010). Regioisomeric analysis of triacylglycerols using silver-ion liquid chromatography-atmospheric pressure chemical ionization mass spectrometry: Comparison of five different mass analyzers. Journal of Chromatography A, 1217, 8186–8194. doi:10.1016/j.chroma.2010.10.064

Hsu, F. F., & Turk, J. (2010). Electrospray ionization multiple-stage linear ion-trap mass spectrometry for structural elucidation of triacylglycerols: Assignment of fatty acid groups on the glycerol backbone and location of double bonds. Journal of the American Society for Mass Spectrometry, 21, 657–669. doi:10.1016/j.jasms.2010.01.007

Martin, J. C., Bougnoux, P., Antoine, J. M., Lanson, M., & Couet, C. (1993). Triacylglycerol structure of human colostrum and Mature Milk. Lipids, 28, 637–643. doi:10.1007/BF02536059

Miroslav, L., & Holcapek, M. (2013). Characterization of triacylglycerol enantiomers using chiral HPLC/APCI-MS and synthesis of enantiomeric triacylglycerols. Analytical Chemistry, 85, 1852–1859. doi:10.1021/ac303237a

Miroslav, L., Velinska, H., & Holcapek, M. (2009). Regioisomeric characterization of triacylglycerols using silver-ion HPLC/MS and randomization synthesis of standards. Analytical Chemistry, 81, 3903–3910. doi:10.1021/ac900150j

Nguyen, T., Aparicio, M., & Saleh, M. A. (2015). Accurate mass GC/MS–quadrupole time of flight mass spectrometry analysis of fatty acids and triacylglycerols of spiced oils from the apiaceae family. Molecules, 10, 33–90.

Nguyen, T., Aparicio, M., & Saleh, M. A. (2017). Lipid profiling of the carob fruit (Ceratonia Siliqua L.) using GC/MS/QTOF accurate mass spectrometry. SOJ Chromatograph Science, 1, 6–18.

Pan, M. H., Lai, C. H., Dushenkov, S., & Ho, C. T. (2009). Modulation of inflammatory genes by natural dietary bioactive compounds. Journal of Agricultural and Food Chemistry, 57, 4467–4477. doi:10.1021/jf900612n

Pufal, D. A., Quinlan, P. T., & Salter, A. M. (1995). Effect of dietary triacylglycerol structure on lipoprotein metabolism—A comparison of the effects of dioleoylpalmitoyl glycerol in which palmitoyl is esterified to the 2- or the 1 (3)-position of the glycerol. Biochim Biophysica Acta, 1258, 41–48. doi:10.10160005-2760(95)00095-T

Rustam, Y., & Reid, G. (2018). Analytical challenges and recent advances in mass spectrometry based lipidomics. Analytical Chemistry, 90, 374–397. doi:10.1021/acs.analchem.7b04836

Upton, R., Graff, A., Jolliffe, G., Langer, R., & Williamson, E. (2010). American herbal pharmacopoeia: Botanical pharmacognosy microscopic characterization of botanical medicines (pp. 616–618). Boca Raton, FL: CRC Press.

Valenzuela, A., & Nieto, S. (1994). Biotechnology of lipids—the use of lipases for the structural modification of fats and oils. Grasar Acetas, 45, 337–343. doi:10.3989/gya.1994.v45.i5.1024

Wakil, A., Mir, M., Mellor, D. D., Mellor, S. F., & Atkin, S. L. (2010). The bioavailability of eicosapentanoic acid from reconstituted triglyceride fish oil is higher than that obtained from the triglyceride and monoglyceride forms. Asia Pacific Journal of Clinical Nutrition, 19, 499–505.

Wallace, W., Vaughn, K., Bradford, W., Stewart, B. W., Viswanathan, T., Clausen, E., . . . Carrier, D. J. (2008). Milk thistle extracts inhibit the oxidation of Low-Density Lipoprotein (LDL) and subsequent scavenger receptor-dependent monocyte adhesion. Journal of Agricultural and Food Chemistry, 56, 3966–3970. doi:10.1021/jf703684u

Welte, R., Shah, J., Li, W., Li, M., Chen, J., & Burke, J. J. (2007). Plant lipidomics: Discerning biological function by profiling plant complex lipids using mass spectrometry. Frontiers in Bioscience, 12, 2494–2506. doi:10.2741/2250

William, C. (2017). Comprehensive dual liquid chromatography with quadruple mass spectrometry. Analytical Chemistry, 89, 10537–10546. doi:10.1021/acs.analchem.7b02753

Zampelas, A., Williams, C. M., Morgan, L. M., Wright, J., & Quinlan, T. (1994). The effect of triacylglycerol fatty-acid positional distribution on postprandial plasma metabolite and hormone responses in normal adult men. British Journal of Nutrition, 71, 401–410. doi:10.1079/BJN19940147
