Characterizing Phytochemicals in Camelina Seed Meal by LC-MS^N

Mark A. Berhow and Micheal J. Bowman, USDA, ARS, NCAUR, Peoria, IL

Ming Gu and Yongdong Wang, Cerno Bioscience, Norwalk, CT

Overview

- LC-MS analyses were performed on an extract of defatted camelina seed meal to determine identification of the chemical composition.
- Through peak shape calibration and spectral accuracy calculation, the potential chemical formulas were prioritized.
- Identification of several chemicals in the extract was achieved by a combination of MS-MS-MS fragmentation analysis and spectral accuracy calculation.

Introduction

An optimized single run evaluation that would accurately determine the elemental composition of as many compounds present in an extract would greatly aid in the evaluation of plant tissues. For phytochemicals, we have used accurate mass analysis to quickly characterize the potential chemical formulas for both known and unknown compounds in the seed extracts. We were able to identify several phenolics in the methanol extracts of defatted camelina seeds including rutin and two derivatives of rutin. The identification of the later two compounds was accomplished by elemental composition determination through high spectral accuracy and confirmed with CID/HCD spectra. This result allowed for the focus of isolation and characterization of other unknown compounds in the camelina seed extract.

Methods

Extracts were prepared from defatted camelina seed meal (1). Samples were run on an Thermo Electron LTQ Orbitrap Discovery Mass Spectrometer -- a linear ion trap (LTQ XL) MS, coupled to a high precision electrostatic ion trap (Orbitrap) MS with a higher energy C-trap dissociation (HCD) cell attached -- with an Ion Max electrospray ionization (ESI) source; a Thermo Scientific ACCELA series HPLC system all running under Thermo Scientific Xcalibur 2.1.0.1140 LC-MS software.

HPLC conditions: Column 3 mm x 150 mm Inertsil reverse phase C-18, ODS 3, 3 µ column. Initial solvent conditions: 40% methanol 0.1% formic vs water 0.1% formic; flow rate 0.25 mL per min. Injection 1 µl or less. Development: hold at initial conditions 2 minutes; develop linear gradient to 100% methanol, 0.1% formic acid over 55 additional min. Effluent monitored at 280nm by PDA detector before entering the MS.

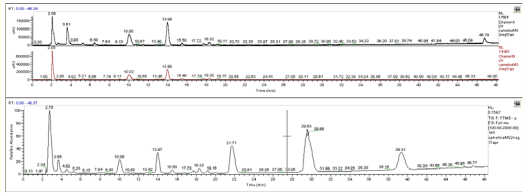
The MS conditions: ESI probe negative mode; source inlet 300 ˚C; sheath gas rate 50 arbitrary units; auxiliary gas rate 5 arbitrary units; sweep gas rate 2 arbitrary units. Maximal mass resolution set at 30,000; spray voltage 3.0 kV; tube lens -100 V. The MS is calibrated with a standard calibration mixture recommended by Thermo Scientific. The software package was set to collect mass data between 100-2000 AMUs. Generally the most significant sample ions generated under these conditions were $[M-1]^-$ and $[M+HCOO]^-$.

For the evaluation of Xcalibur accurate mass data by the Cerno BioScience LLC MassWorks 4.0.0.0 software the FTMS was set to collect spectra at a resolution of 7500 and a range of m/z of 100 to 2000 and then were evaluated by sCLIPS (self Calibrating Line-shape Isotope Profile Search) that enhances formula ID accuracy without the need to run calibration standards.

Results and Discussion

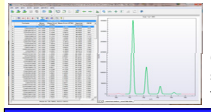
Six mass spec "events" were programed to run in sequence in the MS as a "general" approach.

1. LTQ(IT)-MS full scan m/z 150 to 2000.
2. LTQ(IT)-MS trap most abundant ion perform CID at 35% energy.
3. FT-MS full scan m/z 150 to 2000.
4. IT-MS/MS from Event 1 to HCD at 25% energy.
5. MS3 most abundant fragment ion from Event 2, HCD at 25% energy.
6. MS3 most abundant fragmentation ion from Event 2, CID at 35% energy.

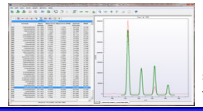


LC-MS separation and analysis of a methanol extract of defatted camelina seeds. Top trace UV trace at 280 nm, bottom trace total ion chromatogram.. Evaluation of peak {M-H}- ions by sCLIPS, peak 21.71, m/z = 506.11895

Rank	Formula	Mono Isotope	Mass Error (PPM)	Spectral Accuracy
1	C14H24O4N11S3	506.1169	3.2862	99.3346
2	C17H32O10NS3	506.1183	0.6229	98.3726
3	C18H28O6N5S3	506.1196	-2.0196	97.2241
4	C19H24O2N9S3	506.121	-4.6621	95.2479
5	C11H28O4N11S4	506.1203	-3.374	95.2439



Good sCLIPS fit

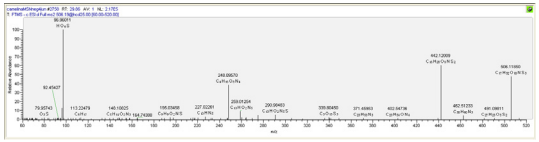


Poor sCLIPS fit

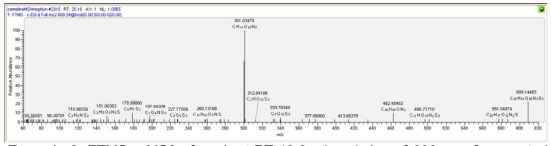
Predicted Formulas for Detectable [M-H]- ions

RT	Exact mass	Mass Error (PPM)	Best Formula	Rank by Spectral Accuracy	Spectral Accuracy (%)	Formula Hits
3.7	451.1220	3.5	C21H23O11	1	96.00	41
10.0	741.1849	4.9	C32H29O20	1	99.60	15
14.0	609.1428	3.6	C27H29O16	1	99.20	114
15.5	947.2408	4.1	C43H47O24	2	98.34	6
17.7	593.1484	3.5	C27H29O15	1	99.29	131
18.3	623.1583	1.5	C28H31O16	2	99.63	83
19.2	623.1616	1.5	C28H31O16	2	99.63	83
21.7	506.1189	0.6	C17H32O10NS3	2	98.37	1
22.5	417.2111	2.4	C20H33O9	1	96.00	5
22.9	417.2111	2.4	C20H33O9	1	96.00	5
24.4	431.2269	1.8	C21H35O9	1	96.00	10
26.2	413.2269	1.8	C21H35O9	1	96.00	10
26.8	329.2316	1.3	C18H33O5	1	96.00	15
27.2	329.2316	1.3	C18H33O5	1	96.00	15
27.8	293.1749	4.9	C17H25O4	1	96.00	61
28.4	329.2316	2.9	C17H29O6	1	96.00	3
29.5	520.1373	1.2	C18H34O10NS3	2	99.11	1
31.2	407.2062	1.3	C22H31O7	1	96.00	36
39.3	534.1511	1.2	C19H36O10NS3	2	98.35	1

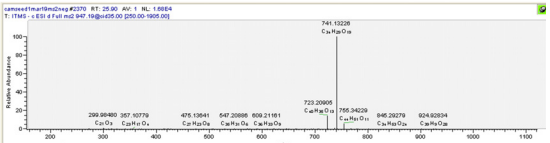
To determine which chemical formula is present MS-MS-MS fragmentation experiments were performed to select the best candidate chemical identification.



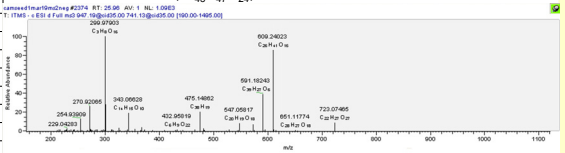
Example 1: FTMS – MS2 of peak at RT 21.7 min m/z ion of 506 was fragmented by the HCD to provide ions of m/z of 442.1, 248.1, and 96.96. The 96.96 is the SO4- ion which is only generated by glucosinolate class compounds. Only one compound was found in a plant chemical database search to match this formula: glucoarabin (9-methyl-sulfinyl-nonyl-glucosinolate). This compound has been reported in the literature to be found in camelina.



Example 2: FTMS – MS2 of peak at RT 13.9 min m/z ion of 609 was fragmented by the HCD to provide ions of m/z of 462 and 301. The chemical formula selected by the sCLIPS analysis corresponds to that of the flavonol rutin (quercetin-3-O-(6"-O-rhamnosyl)-glucoside) which has been reported in the literature to be found in camelina. The 462 ion is the loss of a terminal rhamnose sugar unit, the 301 ion is the flavonol aglycone quercetin. This can be confirmed by both retention time and MS-MS fragmentation of standard rutin.



Example 3: Analysis of unknown UV peak Rt = 26.0 MS event 1, Linear ion-trap (IT-MS) major negative mass ion m/z 947.29. MS event 2, major mass ion m/z 947.29) fragmented by CID 35% yielded major daughter ion m/z 741.13 and smaller fragments of m/z 755.34 and m/z 723.209. MS event 3, FT-MS major mass ion m/z 947.24768, $C_{43}H_{47}O_{24}$, error 2.5%.



MS Event 5, IT-MS3 of daughter ion m/z 741 produced by MS event 2 was fragmented with 35% CID energy, producing granddaughter ions m/z 723.07 (loss of water); m/z 609.24 (loss of apiose); m/z 591.18 (loss of apiose and water), m/z 475.15 (loss of rhamnose and 2/3 glucose), m/z 343.06 (loss of apiose, rhamnose and 2/3 glucose); and m/z 299.97 (loss of apiose, rhamnose, glucose and H).

A search was made in a database of phytochemical compounds for the chemical formula of $C_{43}H_{47}O_{24}$. Six compounds were listed, all isomeric variations of similar compounds: quercetin (or kaempferol)-triglycosides conjugated to a substituted cinnamic acid. Based on the LC-MS analysis of the other flavonoids found in camelina seeds, it is probable that this quercetin glycoside has an apiose sugar group instead of a second glucose. In order to make up the difference in MW, the phenolic acid group would have to be sinapic acid instead of ferulic acid.

The tentative identification of this minor constituent of camelina seeds is quercetin-O-glucosyl-?-O-sinopyl-2"-O-apiosyl-6"-O-rhamnoside.

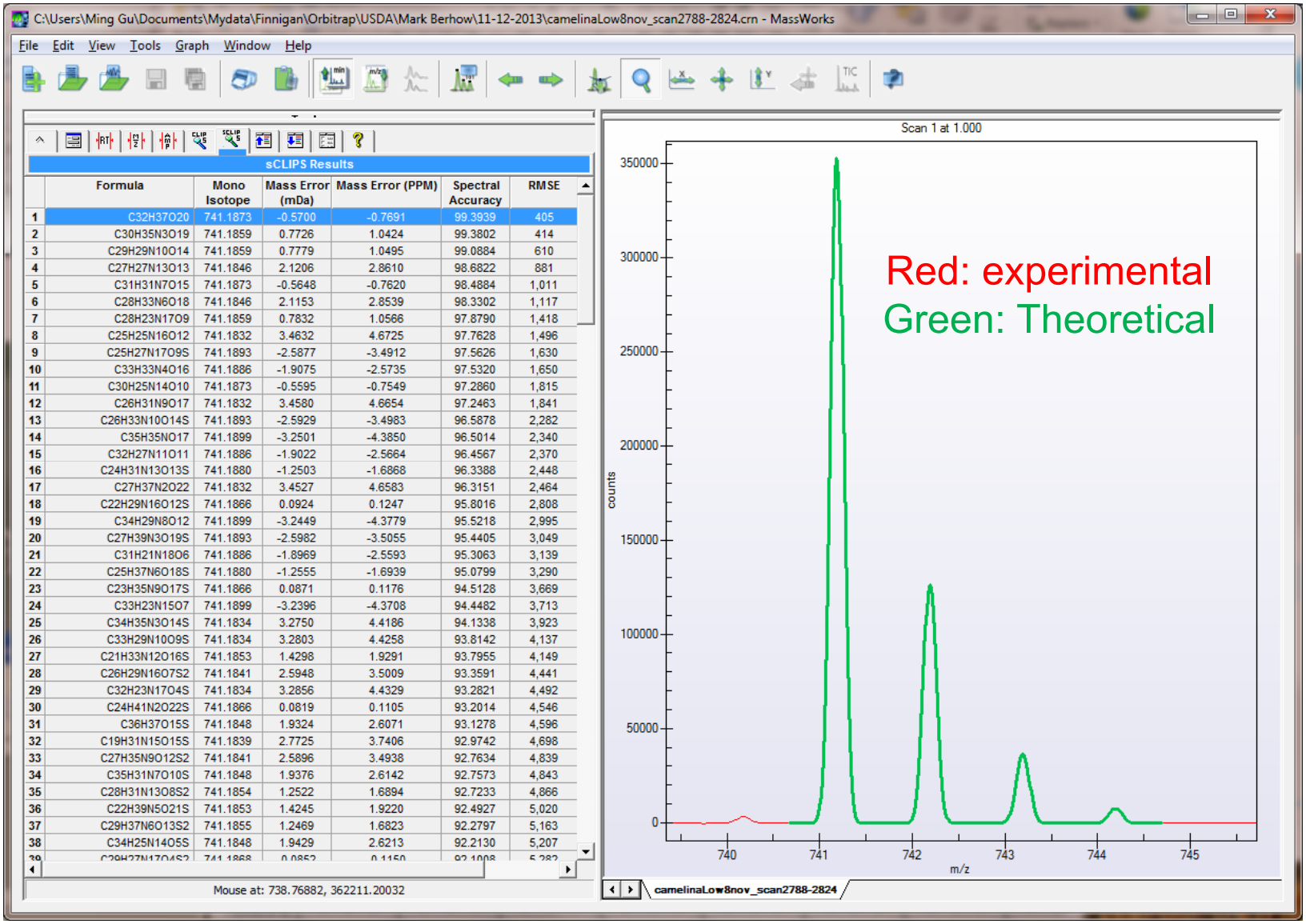
Identification of the Major Peaks

RT	accurate mass	Best Formula	Chemical ID	confirmation	Literature
3.65	451.12198	C21H23O11	6-OH-flavane-rhamnoside		
10	741.18494	C32H29O20	Quercetin-3-O-gluc-rham-apiose	NMR	
14	609.14283	C27H29O16	Quercetin-3-O-gluc-rham (rutin)	standard	yes
15.5	947.24075	C43H47O24	Quercetin 3-O-gluc-rham-apio-sinopic acid	MS-MS frag	
17.7	593.14837	C27H29O15	Kaempferol-7-O-gluc-rham	MS-MS frag	aglycone
18.3	623.15831	C28H31O16	Rhamnetin-O-gluc-rham	MS-MS frag	aglycone
19.2	623.16162	C28H31O16	Rhamnetin-O-gluc-rham	MS-MS frag	aglycone
21.7	506.11709	C17H32O10NS3	9-methylsulfinyl-nonyl-glucosinolate (glucoarabin)	standard	yes
22.5	417.21107	C20H33O9	11-abeo-11taxene-9 hydroxyl		
22.9	417.21113	C20H33O9	11-abeo-11taxene-9 hydroxyl		
24.4	431.2269	C21H35O9	Eudesmene-4 hydroxyl-acetate		
26.2	413.22691	C21H35O9	Eudesmene-4 hydroxyl-acetate		
26.8	329.23156	C18H33O5	no likely candidates		
27.2	329.23155	C18H33O5	no likely candidates		
27.8	293.17487	C17H25O4	2 OH-eudesmen-2-one-acetate		
28.4	329.23161	C17H29O6	bisabolene-5 hydroxyl-acetate		
29.5	520.13732	C18H34O10NS3	10-methylsulfinyl-decyl-glucosinolate (glucocamelinin)	standard	yes
31.2	407.2062	C22H31O7	3 OH-eudesmen-6-one-epoxy-methylbutanoyl-tetra-acetate		
39.3	534.15105	C19H36O10NS3	11-methylsulfinyl-undecyl-glucosinolate	standard	yes

Conclusions

Evaluation of mass accurate data generated by an Orbitrap MS by MassWorks self Calibrating Line-shape Isotope Profile Search (sCLIPS) provided a series of 17 chemical formula identifications in a extract from camelina seeds. Only four compounds were previously described in the literature for camelina. Five other flavonoids were tentatively identified by the interpretation of MS-MS-MS experiments, and nine more were identified by potential formula matches alone. This method shows great promise for accurate unknown formula predictions that can be further refined to a chemical structure determination by MS-MS fragmentation.

Good Match


C:\Users\Ming Gu\Documents\Mydata\Finnigan\Orbitrap\USDA\Mark Berhow\11-12-2013\camelinaLow8nov_scan2788-2824.crn - MassWorks
File Edit View Tools Graph Window Help
sCLIPS Results
Formula | Mono Isotope | Mass Error (mDa) | Mass Error (PPM) | Spectral Accuracy | RMSE
1 C32H37O20 741.1873 -0.5700 -0.7691 99.3939 405
2 C30H35N3O19 741.1859 0.7726 1.0424 99.3802 414
3 C29H29N10O14 741.1859 0.7779 1.0495 99.0884 610
4 C27H27N13O13 741.1846 2.1206 2.8610 98.6822 881
5 C31H31N7O15 741.1873 -0.5648 -0.7620 98.4884 1,011
6 C28H33N6O18 741.1846 2.1153 2.8539 98.3302 1,117
7 C28H23N17O9 741.1859 0.7832 1.0566 97.8790 1,418
8 C25H25N16O12 741.1832 3.4632 4.6725 97.7628 1,496
9 C25H27N17O9S 741.1893 -2.5877 -3.4912 97.5626 1,630
10 C33H33N4O16 741.1886 -1.9075 -2.5735 97.5320 1,650
11 C30H25N14O10 741.1873 -0.5595 -0.7549 97.2860 1,815
12 C26H31N9O17 741.1832 3.4580 4.6654 97.2463 1,841
13 C26H33N10O14S 741.1893 -2.5929 -3.4983 96.5878 2,282
14 C35H35NO17 741.1899 -3.2501 -4.3850 96.5014 2,340
15 C32H27N11O11 741.1886 -1.9022 -2.5664 96.4567 2,370
16 C24H31N13O13S 741.1880 -1.2503 -1.6868 96.3388 2,448
17 C27H37N2O22 741.1832 3.4527 4.6583 96.3151 2,464
18 C22H29N16O12S 741.1866 0.0924 0.1247 95.8016 2,808
19 C34H29N8O12 741.1899 -3.2449 -4.3779 95.5218 2,995
20 C27H39N3O19S 741.1893 -2.5982 -3.5055 95.4405 3,049
21 C31H21N18O6 741.1886 -1.8969 -2.5593 95.3063 3,139
22 C25H37N6O18S 741.1880 -1.2555 -1.6939 95.0799 3,290
23 C23H35N9O17S 741.1866 0.0871 0.1176 94.5128 3,669
24 C33H23N15O7 741.1899 -3.2396 -4.3708 94.4482 3,713
25 C34H35N3O14S 741.1834 3.2750 4.4186 94.1338 3,923
26 C33H29N10O9S 741.1834 3.2803 4.4258 93.8142 4,137
27 C21H33N12O16S 741.1853 1.4298 1.9291 93.7955 4,149
28 C26H29N16O7S2 741.1841 2.5948 3.5009 93.3591 4,441
29 C32H23N17O4S 741.1834 3.2856 4.4329 93.2821 4,492
30 C24H41N2O22S 741.1866 0.0819 0.1105 93.2014 4,546
31 C36H37O15S 741.1848 1.9324 2.6071 93.1278 4,596
32 C19H31N15O15S 741.1839 2.7725 3.7406 92.9742 4,698
33 C27H35N9O12S2 741.1841 2.5896 3.4938 92.7634 4,839
34 C35H31N7O10S 741.1848 1.9376 2.6142 92.7573 4,843
35 C28H31N13O8S2 741.1854 1.2522 1.6894 92.7233 4,866
36 C22H39N5O21S 741.1853 1.4245 1.9220 92.4927 5,020
37 C29H37N6O13S2 741.1855 1.2469 1.6823 92.2797 5,163
38 C34H25N14O5S 741.1848 1.9429 2.6213 92.2130 5,207
Mouse at: 738.76882, 362211.20032
Scan 1 at 1.000
Red: experimental
Green: Theoretical
counts
m/z
camelinaLow8nov_scan2788-2824