Postprint of Food Chemistry Vol. 187, 15 November 2015, Pages 166–173
doi:10.1016/j.foodchem.2015.04.022

TITLE: **Isolation and identification of minor secoiridoids and phenolic components from thermally treated olive oil byproducts.**

FÁTIMA RUBIO-SENENT[1], SERGIO MARTOS[2], ANTONIO LAMA-MUÑOZ[1], JOSÉ G. FERNÁNDEZ-BOLAÑOS[2], GUILLERMO RODRÍGUEZ-GUTIÉRREZ[1], AND JUAN FERNÁNDEZ-BOLAÑOS[1]*.

(1) Instituto de la Grasa, Consejo Superior de Investigaciones Científicas (CSIC). Ctra Utrera km 1, Campus Universitario Pablo de Olavide, Edificio 46, Seville-41013, Spain.

(2) Universidad de Sevilla, Departamento de Química Orgánica, Facultad de Química, Profesor García González 1, Seville-41012, Spain

* Corresponding author: Tel: 34-954691054; Fax 34-954691262; E-mail: jfbg@cica.es

ABSTRACT

The application of a novel industrial process based on the hydrothermal treatment of 160 ºC/60 min of alperujo, a by-product of olive oil extraction, allows the formation of a liquid phase containing a high concentration of phenolic and secoiridoid compounds. Ethyl acetate was used to extract these phenolic compounds from the aqueous matrix. The phenols present in high concentrations in the ethyl acetate extract have been studied in previous work. In this study, we confirmed the presence of several phenolic compounds existing in minor concentrations in the steam-treated alperujo extract, but which contribute to the characteristics of this organic extract. The polar compounds that remain in the aqueous fraction after extraction with ethyl acetate were also studied. We report the presence of known compounds and we also detected an unknown molecule with a molecular weight of 408 whose structure was characterized. This new secoiridoid glucoside was identified as β-D-glucopyranosyl acyclodihydroelenolic acid.

KEYWORDS: pomace olive waste (Alperujo); steam-treatment; ethyl acetate extract; secoiridoid derivatives;.acteosides; flavonoids; phenolic compounds.

1. Introduction.

A modern two-phase centrifugation system to obtain extra virgin olive oil is now widely used in Spain (accounting for 90% of the total production). Besides olive oil, the process releases a semisolid by-product called two-phase pomace or "alperujo", with over 4 million of tonnes generated annually in Spain. Olive fruits and derived products represent a recognized valuable source of phenolic compounds with important beneficial effects for human health including antioxidant, anti-inflammatory, antimicrobial, and anticarcinogen activities (Estruch et al., 3013; Kalogerakis, Polito, Foteinis, Chatzisymeon & Mantzavinos, 2013; Ruiz-Canela, Estruch, Corella, Salas-Salvador, & Martínez-González, 2014). The possible recovery of bioactive compounds from this by-product, containing some 98% of the total phenolic content of the olive fruit in comparison to only 2% in the oil, is an attractive way of valorising it. In addition, to improve the productivity of olive oil processing, using the waste material is being promoted. One interesting approach toward the valorisation of alperujo is the use of a hydrothermal treatment at high temperature (150-170 ºC) and pressure (0.6-1.2 MPa), which allows an easy separation of the solid and liquid phases and the recovery of the bioactive compounds from each phase (Fernández-Bolaños, Rodríguez-Gutiérrez, Lama-Muñoz, & Sánchez, 2010), and whose methodology is already implemented on an industrial scale. Recent studies have shown that a substantial part of the phenolic compounds, the pectins and hemicelluloses present in alperujo could also be recovered from the water soluble fraction after the steam treatment (Rubio-Senent, Lama-Muñoz, Rodríguez-Gutiérrez, & Fernández-Bolaños, 2013; Rubio-Senent, Rodríguez-Gutiérrez, Lama-Muñoz, & Fernández-Bolaños, 2012; Rubio-Senent, Rodríguez-Gutiérrez, Lama-Muñoz, García, & Fernández-Bolaños 2015). Many of these phenols were isolated and identified from the ethyl acetate extract obtained from the aqueous phase of steam-

65 treated alperujo. Moreover, many interesting antioxidant activities have been attributed
66 to this class of compounds (Rubio-Senent et al., 2012; Rubio-Senent et al., 2013).
67 However, the complete characterization of the ethyl acetate extract has yet to be
68 achieved. In this work we considered the minor phenols present in the organic fraction,
69 which could contribute to the antioxidant activity of the total extract. In fact, Obied,
70 Prenzler, Konczak, Rehman, & Robards (2009) showed that a mixture of olive
71 biophenols in ethyl acetate extacts of olive mill waste were more effective in protecting
72 DNA from oxidative damage and inhibiting the growth of cancer cells than individual
73 biophenols.

74 The more polar components that remain in the aqueous fraction after the ethyl
75 acetate extraction were also studied in previous work. After ultra-filtration through a
76 3000 Da molecular weight cut-off membrane, a fraction over 3000 Da was found to be
77 enriched in pectin. The fraction below 3000 Da was mainly composed of neutral
78 oligosaccharides with a degree of polymerisation (DP) between 4−10, together with
79 several secoiridoid glycosides esterified to phenolic compounds, their aglycones, and
80 mono- and disaccharides linked to phenolic compounds (Fernández-Bolaños, Rubio-
81 Senent, Lama-Muñoz, García, & Rodríguez-Gutiérrez, 2014). The occurrence of these
82 interesting compounds prompted us to continue our investigation on the same aqueous
83 fraction. In this work, we identified another series of known phenol and secoiridoid
84 derivatives by the analysis of their spectral mass data and comparison with literature
85 values. In addition, an unknown compound with a molecular weight of 408 was
86 identified for the first time based on nuclear magnetic resonance (NMR) spectra of both
87 the unknown compound and its acetylated derivative.

88

2. Materials and methods.

2.1. Materials.

The sample of alperujo (a semisolid residue composed of olive peels, pulp, seeds, and ground stones) was obtained in March 2009, from Picual olives processed at a Spanish oil mill (Almazara Experimental, Instituto de la Grasa, Seville). The alperujo was processed in the pilot reactor without removal of the stones.

2.2. Thermal Treatment.

The hydrothermal treatment used has been patented (Fernández-Bolaños, Rodríguez-Gutiérrez, Lama-Muñoz, & Sánchez, 2010) and was performed using a prototype designed by our research group at the Instituto de la Grasa (Seville, Spain). The reactor has a stainless steel reservoir (100 L capacity) that can operate at temperatures between 50 and 190 °C and at a maximum pressure of 1.2 MPa.

20 kg of fresh alperujo was loaded into the reactor and treated for 60 min at 160 °C. Then the wet material was centrifuged at 4,700 g (Comteifa, S.L., Barcelona, Spain) to separate the solids and liquids. After centrifugation, 10 L of the liquid phase was concentrated to 1 L by rotary evaporation at 30 °C under reduce pressure.

2.3. Phenol Extraction.

The aqueous portion obtained after thermal treatment was washed with hexane to remove the lipid fraction: 1 L of liquid was mixed with 500 mL of hexane; the mixture

was shaken vigorously, and then the phases were separated by decantation and washed twice. The liquid−liquid extraction of phenolic compounds was carried out with ethyl acetate (500 mL per 200 mL of sample; refluxed at 77 °C) in a continuous extractor of a heavier liquid (water) by a lighter liquid (ethyl acetate) for 8 h.

2.4. Concentration of minor phenols in phenolic extract.

Approximately 3 g of extract was dissolved in 50 mL of H_2O/MeOH (80:20). First, the extracts were passed through a polyamide column 3.5 cm in diameter and 30 cm in height. The elution was performed stepwise with 500 mL of H_2O, 25% MeOH (v/v), 50% MeOH (v/v), 75% MeOH (v/v) and 100% MeOH (v/v), and each fraction was collected. The fractions were analyzed by HPLC, and those that contained a greater number of unknown phenols were passed through a second column of Amberlite XAD16, 3.5 cm in diameter and 15 cm in height. The elution was performed with 250 mL of H_2O, 25% MeOH (v/v), 50% MeOH (v/v), 75% MeOH (v/v), and 100% MeOH (v/v), and each collected fraction was monitored by HPLC. The phenols were purified and concentrated using a cartridge of Sep-Pak C_{18}, with elution by H_2O, MeOH 20 % (v/v), MeOH 40 % (v/v),, MeOH 80 % (v/v) and MeOH 100 % (v/v). This process allowed the collection of three main fractions: Polyamide water/XAD 50/SPE 20, Polyamide water/XAD75/SPE 20 and Polyamide water/XAD 75/SPE 40, referred to as PXS50/20, PXS75/20 and PXS75/40 respectively. These three fractions were enriched in minor compounds, which were later identified by HPLC-MS.

2.5. Isolation of phenolic compounds in aqueous fraction after extraction with ethyl acetate.

In order to increase the proportion of low molecular weight oligosaccharides in the initial liquid, a mild chemical hydrolysis with 0.5 N HCl was carried out (Lama-Muñoz, Rodríguez-Gutiérrez, Rubio-Senent, & Fernández-Bolaños, 2012). A protocol for the isolation of these oligosaccharides, including separation by ultrafiltration, adsorption chromatography and size exclusion, has been developed according to a previously published procedure (Rubio-Senent et al., 2013). Briefly, hydrolyzed fractions were ultra-filtered at room temperature using a Prep/Scale®-TFF Cartridge of 3000 Da regenerated cellulose (Millipore Corp., Bedford, MA, USA). Fractions with a size smaller than 3000 Da were fractionated and purified by adsorption on Amberlite XAD-16 resin and eluted with water, methanol/water 20% (v/v) and 50% (v/v), successively. After the separation of the monosaccharides and fractionation of the oligosaccharides by adsorption chromatography, the fractions were fractionated by size exclusion chromatography on two Superdex Peptide HR 10/30 (30 x 1 cm) columns (Pharmacia Biotech, Uppsala, Sweden) connected in line to a Jasco LC-Net II/ADC (Easton, MD, USA) HPLC system.

Samples (100 μL) were eluted with a flow rate of 0.5 mL/min. The peaks were monitored by a Jasco MD-1550 Diode Array detector (DAD) and a Jasco RI-1530 Refraction Index detector. Column calibration was performed with a variety of standards, using compounds of a series of cellobiose (RT= 66.5 min), cellotriose (RT= 63.5 min), cellotetraose (RT= 61.1 min), cellopentaose (RT= 59.0 min) and cellohexaose (RT= 56.2 min) (purchased from Megazyme).

2.6. HPLC-DAD.

The different phenols were analyzed using a Hewlett-Packard 1100 liquid chromatography system with a C-18 column (Teknokroma Tracer Extrasil ODS-2, 250 mm × 4.6 mm i.d., 5 μm). The system was equipped with a diode array detector and Rheodyne injection valves (20 μL loop). The mobile phases were 0.01% trichloroacetic acid in water (A) and acetonitrile (B), utilizing the following gradient over a total run time of 55 min: 95% A initially, 75% A at 30 min, 50% A at 45 min, 0% A at 47 min, 75% A at 50 min, and 95% A at 52 min, until the run was completed.

2.7. HPLC-MS.

The phenolic compounds present in the different fractions were identified by electron impact mass spectra collected on a quadrupole mass analyzer (ZMD4, Micromass, Waters Inc., Manchester, UK). Electron spray ionization (ESI) mass spectra were obtained at ionization energies of 50 and 100 eV in negative mode and of 50 eV in positive mode; the capillary voltage was 3 kV, the desolvation temperature was 200 °C, the source temperature was 100 °C, and the extractor voltage was 12 V. The flow was maintained at 1 mL min^{-1} in split mode (UV detector MS) for each analysis. A Tracer Extrasil ODS-2 column (250 mm × 4.6 mm i.d., 5 μm) (Teknokroma, Barcelona, Spain) was used. The mobile phases were 0.01% trichloroacetic acid in water and acetonitrile, utilizing the gradient used in HPLC-DAD.

2.8. Structural analysis.

2.8.1. ^{1}H and ^{13}C-NMR, COSY, HSQC and NOESY spectra.

^{1}H (500.1 MHz) and ^{13}C (125.8 MHz) NMR spectra were recorded on Bruker Avance-500 spectrometer using D_2O or CD_3OD as solvent. Chemical shifts are reported in δ units (ppm) relative to the solvent peak (D_2O set at 4.79 for ^{1}H-NMR, CD_3OD set at 33.31 and 49.0 ppm for ^{1}H- and ^{13}C-NMR, respectively). The assignments of ^{1}H and ^{13}C signals were confirmed by homonuclear COSY, and heteronuclear HSQC spectra, respectively. The configuration of the double bond was confirmed by a 2D-NOESY experiment.

2.8.2. Mass spectra.

High Resolution Electrospray Ionization (HR-ESI) was performed using a Qexactive Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Scientific) with a resolution of 70,000. The spray needle voltage was +3.5 kV in positive ionization, or -2.5 kV in negative ionization, as indicated for each compound. Samples were solved in MeOH:H_2O (50%) with 0.1% formic acid as additive and automatically introduced using a UHPLC Ultimate 3000 (Dionex).

2.8.3. IR spectra.

IR spectra were recorded on Jasco FT/IR 4100 equipped with an ATR accessory.

4-[(β-D-Glucopyranosyloxy)methyl]-3-[3-hydroxy-2-(methoxycarbonyl)propyl)] hex-4-enoate 1 (Trivial name: **1-glucosyl acyclodihydroelenolic acid**)

IR ν_{max} 3650-2500 (br), 1717, 1438, 1366, 1268, 1233, 1163, 1068 (sh), 1019 (str) cm^{-1}. ^{1}H-NMR (500.1 MHz, D_2O): δ 5.97 (q, 1H, $J_{8,10}$ = 7.0 Hz, H-8), 4.53 (d, 1H, $J_{1',2'}$ = 8.0 Hz, H-1'), 4.43 (d, 1H, $J_{1a,1b}$ = 11.7 Hz, H-1a), 4.19 (d, 1H, $J_{1b,1a}$ = 11.7 Hz, H-1b), 3.99 (dd, 1H, $J_{6a',5'}$ = 2.1 Hz, $J_{6a',6b'}$ = 12.3 Hz, H-6a'), 3.83 (s, 3H, OCH_3), 3.81-3.76 (m, 3H, H-6b', H-3a, H-3b), 3.55 (t, 1H, $J_{3',2'}$ = $J_{3',4'}$ = 9.1 Hz, H-3'), 3.53-3.42 (m, 3H, H-4', H-5', H-5), 3.34 (dd, 1H, $J_{2',1'}$ = 8.0 Hz, $J_{2',3'}$ = 9.1 Hz, H-2'), 3.04 (ddd, 1H, $J_{4,3a}$ = 4.8 Hz, $J_{4,5}$ = 8.9 Hz, $J_{4,3b}$ = 11.3 Hz, H-4), 2.67 (dd, 1H, $J_{6a,5}$ = 9.6 Hz, $J_{6a,6b}$ = 15.4 Hz, H-6a), 2.56 (dd, 1H, $J_{6b,5}$ = 5.4 Hz, $J_{6b,6a}$ = 15.4 Hz, H-6b), 1.77 (d, 3H, $J_{10,8}$ = 7.0 Hz, H-10); ^{13}C-NMR (125.8 MHz, D_2O): δ 177.0 (C-7), 176.6 (*C*OOMe), 132.3 (C-9), 131.7 (C-8), 101.7 (C-1'), 2 x 75.9 (C-3', C-5'), 73.2 (C-2'), 71.5 (C-1), 69.7 (C-4'), 61.8 (C-3), 60.9 (C-6'), 52.5 (O*C*H_3), 50.9 (C-4), 36.8 (C-6), 34.7 (C-5), 13.0 (C-10); ESI-MS *m/z* 407 ($[M-H]^-$, 29%); HR-ESI-MS *m/z* calculated for $C_{17}H_{27}O_{11}$ ($[M-H]^-$): 407.1559, found: 407.1562.

3-[3-Hydroxy-2-(methoxycarbonyl)propyl]-4-[2',3',4',6'-tetra-O-acetyl-β-D-glucopyranosyloxy)methyl]hex—enoate 2.

IR ν_{max} 2929, 2861, 1739 (str), 1570, 1424, 1372, 1219 (str), 1038 (str), 903 (weak), 837 (weak) cm^{-1}. ^{1}H-NMR (500.1 MHz, CD_3OD): δ 5.78 (q, 1H, $J_{8,10}$ = 6.9 Hz, H-8), 5.24 (t, 1H, $J_{3',2'}$ = $J_{3',4'}$ = 9.4 Hz, H-3'), 5.03 (dd, 1H, $J_{4',3'}$ = 9.4 Hz, $J_{4',5'}$ = 10.0 Hz, H-4'), 4.87 (dd, 1H, $J_{2',1'}$ = 8.0 Hz, $J_{2',3'}$ = 9.4 Hz, H-2'), 4.76 (d, 1H, $J_{1',2'}$ = 8.0, H-1'), 4.28 (dd, 1H, $J_{6a',5'}$ = 4.6 Hz, $J_{6a',6b'}$ = 12.3 Hz, H-6a'), 4.27 (d, 1H, $J_{1a,1b}$ = 11.0 Hz, H-1a), 4.24 (d, 1H, $J_{1a,1b}$ = 11.0 Hz, H-1b), 4.24 (dd, 1H, $J_{3a,4}$ = 4.0 Hz, $J_{3a,3b}$ = 10.5 Hz, H-3a), 4.15 (dd, 1H, $J_{6b',5'}$ = 2.3 Hz, $J_{6a',6b'}$ = 12.3 Hz, H-6b'), 4.02 (dd, 1H, $J_{3b,4}$ = 9.8 Hz, $J_{3a,3b}$ = 10.5 Hz, H-3b), 3.88 (ddd, 1H, $J_{5',6b'}$ = 2.4 Hz, $J_{5',6a'}$ = 4.6 Hz, $J_{5',4'}$ = 10.0 Hz,

H-5'), 3.71 (s, 3H, OCH_3), 3.49 (ddd, 1H, $J_{5,6b}$ = 5.1 Hz, $J_{5,6a}$ = 10.2 Hz, $J_{5,4}$ = 11.4 Hz, H-5), 2.86 (ddd, 1H, $J_{4,3a}$ = 4.0 Hz, $J_{4,3b}$ = 9.8 Hz, $J_{4,5}$ = 11.4 Hz, H-4), 2.37 (dd, 1H, $J_{6a,5}$ = 10.2 Hz, $J_{6a,6b}$ = 14.2 Hz, H-6a), 2.31 (dd, 1H, $J_{6b,5}$ = 5.1 Hz, $J_{6a,6b}$ = 14.2 Hz, H-6b), 2.08, 2.07, 2.01, 2.00, 1.96 (5s, 3H each, CH_3CO), 1.73 (d, 3H, $J_{10,8}$ = 6.9 Hz, H-10). ^{13}C-NMR (125.8 MHz, CD_3OD): δ 178.7 (C-7), 175.5 (*C*OOMe), 172.5, 172.4, 171.6, 171.3, 171.2 (5 CH_3OH), 134.9 (C-9), 131.5 (C-8), 101.1 (C-1'), 74.8 (C-3'), 73.7 (C-1), 73.0 (C-2'), 72.9 (C-5'), 69.9 (C-4'), 65.7 (C-3), 63.2 (C-6'), 52.3 (O*C*H_3), 50.0 (C-4), 40.2 (C-6), 37.0 (C-5), 20.9, 20.7, 20.6, 2 x 20.5 (5 *C*H_3CO), 14.0 (C-10). ESI-MS *m/z* 641 ($[M+Na]^+$, 17%), 214 (100%); HR-ESI-MS *m/z* calculated for $C_{27}H_{38}O_{16}Na$ ($[M+Na]^+$): 641.2052, found: 641.2051.

2.9. Acetylation of 1-glucosyl acyclodihydroelenolic acid.

A solution of 1-glucosyl acyclodihydroelenolic acid (42 mg, 0.103 mmol) in Ac_2O/Py (1:1, 4 mL) was stirred at room temperature for 24 h. Then, the mixture was concentrated to dryness under high vacuum and the residue was purified by column chromatography (AcOEt-cyclohexane 1:1 → AcOEt-MeOH 4:1). Yield: 45 mg, 70%.

3. Results and discussion.

3.1 Minor phenolic compounds present in the organic fraction.

We have previously characterized the phenolic extracts recovered with ethyl acetate from the liquid phase obtained from a new hydrothermal process of alperujo (Rubio-Senent et al., 2012). Among the compounds detected, hydroxytyrosol was

observed in the highest concentration, followed by elenolic acid derivatives and a polymeric phenolic fraction. In addition, almost 25 other compounds were identified using HPLC-MS, which demonstrated the great complexity of the steam-treated alperujo phenolic fraction. In this work, the organic ethyl acetate extract was further separated in order to recover and characterise other minor phenolic compounds from steam-treated alperujo, which have not been previously identified. The extract was subjected to fractionation by means of an atmospheric-pressure chromatographic system through a series of adsorbent polyamide resins followed by Amberlite XAD resin, and further desorption with water and methanol in order to separate the minor compounds from those present in greater proportions. Finally, to facilitate the identification of the minor compounds, they were concentrated and purified by solid phase extraction (SPE) using a reversed-phase C18 cartridge. This process allowed the collection of three main fractions PXS50/20, PXS75/20 and PXS75/40, enriched in minor compounds that were identified by HPLC-MS (**Figure 1; Table 1**).

The mass spectrum of the PXS50/20 fraction identified an oleoside methyl ester, characterized by a quasi-molecular ion in negative mode $[M-1]^-$ with *m/z* 403. In the PXS75/20 fraction, the dihydro-oleuropein was identified; its mass spectrum presented signals at *m/z* 543, 525 and 151, which were associated with $[M-H]^-$, $[M-H_2O]^-$, and with the remaining elenolic acid core (De Nino, Mazzotti, Morrone, Perri, Raffaelli, & Sindona, 1999), respectively. The PXS75/40 fraction was particularly enriched in minor phenolic compounds. Furthermore, in this fraction, dihydro oleuropein was also identified. A molecule with a quasi-molecular ion at *m/z* 701 was detected. This signal could be attributed to neo-nüzhenide or oleuropein diglucose since both species produced the fragment corresponding to the loss of a glucose molecule, *m/z* 539. Nüzhenide was previously found to be the major phenolic compound in olive seeds

(Servili, Baldioli, Selvaggini, Macchioni, & Montedoro, 1999). The alperujo sample, from which the phenolic extract was obtained, contains fragments of stone, and the molecule nüzhenide has been previously detected and quantified in alperujo extracts following hydrothermal treatment (Rubio-Senent et al., 2012). Therefore, it is possible that lower concentrations of neo-nüzhenide were present in minor concentrations. The species with m/z 540 was identified as a derivative of oleuropein, due to similarities in the absorbance profile and molecular weight, but the fragments obtained by MS are different from those described for oleuropein. Finally, another species with a molecular weight of 542 was detected; this molecule is identified as hydro oleuropein. The hydro oleuropein had a quasi-molecular ion at 541 and a fragment was observed at *m/z* 405, which could be attributed to $[M-HT-1]^{-}$; this species has not been detected to date, but the presence of oleuropein and dihydro oleuropein in the extract led us to consider it as an intermediate compound.

3.2 Secoiridoids and phenolic glycosides present in the aqueous fraction.

The aqueous fraction obtained after steam treatment of alperujo and subsequent ethyl acetate extraction was treated by a mild acid hydrolysis to increase the proportion of oligosaccharides of low molecular weight released from the cell wall material of olive pulp. Fractions eluted from adsorption chromatography were fractionated by size exclusion chromatography using Superdex-Peptide resin, purified, and characterized, in order to contribute to the knowledge of the more polar phenolic and secoiridoids compounds in steam-treated olive pomace. The identification of the compounds present was based on the search for quasi-molecular ions $[M-H]^{-}$ ions using electrospray

ionization mass spectrometry (ESI-MS) together with the interpretation of the different fragments formed (**Table 2**).

Spectroscopy data showed the presence of a predominant ion at *m/z* 407 which likely corresponds to the depronated molecule $[M-H]^-$, and the corresponding sodium $[M+Na]^+$adduct at *m/z* 431 was observed in the positive mode for a sample with R_t of 11.7 min. Also, the appearance of the peak at *m/z* 569, for another fraction from Superdex-Peptide, is consistent with the existence of second hexose unit, as suggested by the loss of 162 u from *m/z* 569. Although this ion, *m/z* 407, has previously been found in olive fruit (Cardoso, Guyot, Marnet, Lopes-Da-Silva, Renard, & Coimbra, 2005; Menéndez et al., 2008), the structure of the compound was unknown, although it appears consistently in the Superdex-Peptide fractions studied (Fernández-Bolaños et al., 2014). We named this molecule, with a molecular weight at 408, as Molecule **1.** Molecule 1 was purified exhaustibly, and we determined the structure for the first time, as detailed in the next section.

Analysis of the others fractions gave an ion $[M-H]^-$ at *m/z* 639 and another at *m/z* 623, which corresponded with the molecular ions of β-hydroxyverbascoside and verbascoside, respectively. In both cases, the characteristic fragments ions due to the loss of caffeic acid were also present, and the peak at *m/z* 161 would result from a proton transfer and the formation of an anionic ketene (Ryan, Antolovich, Herlt, Prenzler, Lavee, & Robards, 2002). Both acteosides are complex biophenols associated with beneficial health properties, such as an important antioxidant and antibacterial activity, a protection in the concentration of free radicals, and inhibition of the lipid perroxidation (Liu, & Wyman, 2003). The β-hydroxyverbascoside has been described in

315 previous work as a precursor of 3,4-dihydroxyphenylglicol (Lama-Muñoz, Rodríguez-
316 Gutiérrez, Rubio-Senent, Palacios-Díaz, & Fernández-Bolaños, 2013).

317 The flavonoids luteolin-7-*O*-glucoside and quercetin-arabinose were also
318 identified. While the first, a flavone, has previously been detected in the ethyl acetate
319 extract of steam-treated alperujo (Rubio-Senent et al., 2012), the flavonol, quercetin-
320 arabinose was detected for the first time in olive fruit or their by-products in this study.
321 The precursor of this last compound is probably the predominant flavonoid in olive
322 pomace, the quercetin-3-arabinose-glucoside (Bouaziz, Chamkha, & Sayadi, 2004;
323 Yahyaoui et al. 2014), since the loss of glucose during the steam-treatment would result
324 in the quercitin-arabinose detected. Its mass spectrum showed an ion at *m/z* 433 in
325 negative mode and the corresponding signal at *m/z* 301 due to the loss of the pentose
326 unit.

327 In addition, oleuropein, the major phenol in olive fruit, and other derivatives from
328 phenolic secoiridoids, which are closely correlated to oleuropein and ligstroside, were
329 also isolated by Superdex-Peptide chromatography from the aqueous fraction and
330 identified by spectral characterization. These last molecules were identified as
331 oleuropein aglycone derivatives, also identified in olive pulp (De Nino, Mazzotti, Perri,
332 Procopio, Raffaelli, & Sindona, 2000), and as an isomer of ligstroside aglycone
333 identified in olive oil (Fu, Segura-Carretero, Arráez-Román, Menéndez, De la Torre, &
334 Fernández-Gutiérrez, 2009); both were not detected in the ethyl acetate extract.

335

336 **3.3. Characterization of Molecule 1**

Molecule **1,** with a quasi-molecular ion at *m/z* 407 in negative mode, is not apparent from the HPLC chromatogram generated at 280 nm due to the relatively weak absorption of the compound at this wavelength. Molecule **1** was purified by HPLC using a C_{18} semi-preparative column and subsequent detection by electrospray ionization mass spectrometry (ESI-MS) (Supplementary data 1) (as described in Materials & Methods). The structure was elucidated using ^{1}H and ^{13}C NMR techniques (Figure 2A and Supplementary data 2, respectively), and chemical ionization mass spectrometry and infrared analysis. Structural analysis was confirmed using bi-dimensional NMR techniques, such as the COSY (^{1}H-^{1}H correlation) (Supplementary data 3), HSQC (^{1}H-^{13}C correlation) (Supplementary data 4), DEPT-45 and NOESY (through-space coupling) (data not show). In addition, because the signal overlapping in the ^{1}H-NMR spectrum, the sample was peracetylated to confirm the structure of Molecule **1** through the structural analysis of its peracetylated derivative, Molecule **2** (**Figure 3**).

The IR spectrum of the compound (Supplementary data 5) revealed a broad absorption band, 3650-3100 cm^{-1}, corresponding to O-H stretching vibrations of hydroxyl groups and the carboxylic group. The absorption band at 1717 cm^{-1} was attributed to the C=O stretching of the carboxylic and ester groups. Finally, the strong signal at 1019 cm^{-1} was associated with C-O tension of the ester. The ^{1}H-NMR data **(Figure 2A)** showed a quartet at δ 5.78 (J=6.9 Hz) coupled with a signal of methyl at δ 1.73, suggesting that the compound contains a vinyl moiety characteristic of the elenolic acid. ^{1}H-NMR also showed two geminal protons at δ 4.43 and 4.19 (2 doblets, $J_{1a,1b}$ = 11.7 Hz, H-1a, H-1b) that confirm the reduction of the C-1 of elenolic acid glucoside. COSY (Supplementary data 3) gave information about the correlation between H-3a/H3b/H-4, H-4/H-5, H-5/H-6a/H-6b. The HSQC (Supplementary data 4) correlation

of H-3a and H-3b with the carbon signal at 61.8 ppm, confirms the presence of the hydroxyl in position 3. Furthermore, the HSQC correlation of H-1a and H-1b with the carbon at 71.5 ppm confirms that presence of a glycosyloxy group at C-1. In fact, glycosylation of the hydroxyl group at C-1 induces a strong deshielding in the signal of the carbon attached to it (71.5 ppm for C-1), compared with C-3 (61.8 ppm) (Jiménez & Riguera, 1994). The *E* configuration for the olefin moiety was confirmed by 2D-NOESY, where cross peaks H-1a/=CH, H-1b/=CH and H-5/CH_3 indicate through-space coupling, and therefore spatial proximity (**Figure 4A**). Finally, we proved the molecular formula $C_{17}H_{28}O_{11}$ by High Resolution Electrospray Ionization Mass Spectrometry (HR-ESI-MS) on negative mode (Supplementary data 1).

Acetylation of 1-β-D-glucopyranosyl acyclodihydroelenolic acid was carried out in a mixture of acetic anhydride and pyridine (1:1) at room temperature and the compound (Molecule **2**) was purified by column chromatography using silica gel as stationary phase. The presence of five acetyl groups in the final compound was confirmed by ^{1}H-NMR (δ 2.08, 2.07, 2.01, 2.00, 1.96 ppm) (Figure 2B) and ^{13}C-NMR (δ 20.9, 20.7, 20.6, 20.5, 20.5 ppm) (Supplementary data 6). The deshielding of protons H-3a (4.24 ppm) and H-3b (4.02 ppm) for Molecule **2** compared with those protons for Molecule **1** (3.81-3.76 ppm) confirms the presence of a free hydroxyl group at C-3 of **1.** The *E* configuration was again confirmed by 2D-NOESY (Figure 4B, Supplementary data 7). On the basis of configuration of C-5 of oleuropein (Inouye, H., Yoshida, T., Tobita, S., Tanaka, K., & Nishioka, T,), and assuming a similar biosynthetic pathway from 7-ketologanin (Gutierrez-Rosales, Romero, Casanovas, Motilva, & Mínguez-Mosquera, 2010), the configuration of C-5 of the acyclodihydroelenolic acid **1** should be *R;* however, The configuration of C-4 could not be determined. A planar-zigzag conformation planar zigzag conformation for the 5-acetoxypentanoic acid moiety of **2** is

387 in agreement with vicinal H,H coupling constants ($J_{3a,4}$ = 4.0 Hz $J_{3b,4}$ = 9.8 Hz, $J_{4,5}$ =
388 11.4 Hz, $J_{5,6a}$ = 10.2 Hz, $J_{6b,5,6b}$ = 5.1 Hz).

389 The IR spectrum (Supplementary data 8) showed a weak absorption band at 2929
390 cm^{-1} that was attributed to the presence of the carboxylic group. The strong band at
391 1739 cm^{-1} corresponds to the C=O stretching of the carboxylic and ester groups. Finally,
392 the strong bands at 1220 and 1038 cm^{-1} were associated with C-O stretching vibrations
393 of the acetyl and the methyl ester groups, respectively. The molecular formula
394 $C_{27}H_{38}O_{16}$ was also corroborated by HR-ESI mass spectrometry (Supplementary data
395 9). The major fragment *m/z* 214 coincides with the loss of sugar, acetyl and methyl
396 moieties.

397 Taking into account the structural descriptions as detailed above, the compound
398 was identified as 1-β-D-glucopyranosyl acyclodihydroelenolic acid. This compound has
399 not previously been identified. However, another non-aldehydic open cycle secoiridoid
400 had been reported by Obied, Bedgood, Prenzler, & Robards (2007) hydroxytyrosyl
401 acyclodihydroelenolate. In this case, the hydroxytyrosol was linked to a non-
402 glycosylated secoiridoid nucleus. In both cases, the compounds with diol groups likely
403 originate from reduction reactions of elenolic acid, the terpenoid skeleton of oleuropein.
404 It is likely that the formation of **1** involves a glycosylation step after the reduction at C-1
405 has taken place.

406 **4. Conclusions.**

407 The results of this study show the presence of diverse compounds in the alperujo
408 extract (organic and aqueous) obtained from a new industrial liquid source by the steam
409 treatment of alperujo and subsequent extraction with ethyl acetate. In a first step, we

considered the phenolic compounds present in low concentrations in a phenolic extract obtained with ethyl acetate following purification procedures, which due to their minor concentration have not been previously studied. In a second step, the polar compounds present after the extraction with ethyl acetate were studied. Both steps led to the identification of numerous known compounds, including secoiridoids derivatives such as oleuropein and ligstroside derivatives, their hydrolytic derivatives (oleuropein and ligstroside aglycones), acteosides (verbascoside and hydroxyverbascoside) and flavonoids. In addition, we isolated, purified, and elucidated the structure of a new secoiridoid derivative with a molecular weight of 480. The study of spectroscopy allowed the identification of the molecule 1-β-D-glucopyranosyl acyclodihydro acid for the first time.

Acknowledgements.

Funding was received from the Spanish Ministry of Economy and Competitiveness and co-funded by European Social Fund (ESF) (Project AGL2013-48291-R). Dr. Rodríguez-Gutiérrez (RYC-2012-10456 contract) wishes to thank the "Ramón y Cajal" Program, and Ms Rubio-Senent (JAE-Pre104) the Spanish National Research Council (CSIC) program, for providing financial support, both co-funded by ESF.

References

Bouaziz, M., Chamkha, M., & Sayadi, S. (2004) Comparative Study on Phenolic Content and Antioxidant Activity during Maturation of the Olive Cultivar Chemlali from Tunisia. *Journal Agricultural and Food Chemistry. 52*, 5476−5481.

Cardoso, S. M., Guyot, S., Marnet, N., Lopes-Da-Silva, J. A., Renard, C. M., & Coimbra, M. A. (2005). Characterisation of phenolic extracts from olive pulp and olive pomace by electrospray mass spectrometry. *Journal of the Science of Food and Agriculture, 85*, 21–32.

Carlos Jiménez, C. & Riguera, R.(1994). Phenylethanoid Glycosides in Plants*:* Structure and Biological Activity. *Natural Products Reports, 11*, 591-606.

De Nino, A., Mazzotti, F., Morrone, S. P., Perri, E., Raffaelli, A., & Sindona, G. (1999). Characterization of Cassanese olive cultivar through the identification of new trace components by ionspray tandem mass spectrometry. *Journal Mass Spectrometry, 34*, 10−16.

De Nino, A., Mazzotti, F., Perri, E., Procopio, A., Raffaelli, A., & Sindona, G. (2000). Virtual freezing of the hemiacetal-aldehyde equilibrium of the aglycones of oleuropein and ligstroside present in olive oils from Carolea and Coratina cultivars by ionspray ionization tandem mass spectrometry. *Journal Mass Spectrometry, 35*, 461−467.

Estruch, R., Ros, E., Salas-Salvado, J., Covas, M.I., Corella, D., Aros, F., Gomez-Gracia, E., Ruiz-Gutierrez, V., Fiol, M., Lapetra, J., Lamuela-Raventos, R.M., Serra-Majem, L., Pinto, X., Basora, J., Munoz, M.A., Sorli, J.V., Martinez, J.A., & Martinez-Gonzalez, M.A. (2013). Primary Prevention of Cardiovascular Disease with a Mediterranean Diet. *New England Journal of Medicine, 368*, 1279-1290.

Fernández-Bolaños, J. Rodríguez-Gutiérrez, G. Lama-Muñoz, A. & Sánchez, P. (2010). *Dispositivo y procedimiento para el tratamiento de los subproductos de la obtención del aceite de oliva*. (Spanish Patent request No. P201031236).

Fernández-Bolaños, J., Rubio-Senent, F., Lama-Muñoz, A., García, A., & Rodríguez-Gutiérrez, G. (2014). Production of oligosaccharides with low molecular weights, secoiridoids and phenolic glycosides from thermally treated olive by-products. In: Oligosaccharides: Food Sources, Biological Roles and Health Implications. ISBN: 978-1-62948-328-3.

Fu, S., Segura-Carretero, A., Arráez-Román, D., Menéndez, J. A., De la Torre, A., & Fernández-Gutiérrez, A. (2009). Tentative characterization of novel phenolic compounds in extra virgin olive oils by rapid-resolution liquid chromatography coupled with mass spectrometry. *Journal of Agricultural and Food Chemistry*, *57*, 11140-11147.

Gutíerrez-Rosales, F., Romero M. P., Casanovas, M., Motilva, M. J., & Mínguez-Mosquera, M. I. (2010). Metabolites involved in oleuropein accumulation and degradation in fruits of *Olea europaea* L.: Hojiblanca and Arbequina varieties. *Journal of Agricultural and Food Chemistry*, *58*, 12924-1233.

Carlos Jiménez, C. & Riguera, R. (1994). Phenylethanoid glycosides in plants: structure and biological activity. Natural products Reports, 11, 591-606.

Kalogerakis, N., Polito, M., Foteinis, S., Chatzisymeon, E., & Mantzavinos, D. (2013). Recovery of antioxidants from olive mill wastewaters: A viable solution that promotes their overall sustainable management. *Journal of Environmental Management*, *128*, 749-758.

476 Lama-Muñoz, A., Rodríguez-Gutiérrez, G., Rubio-Senent, F., & Fernández-
477 Bolaños, J. (2012). Production, characterization and isolation of neutral and pectic
478 oligosaccharides with low molecular weights from olive by-products thermally treated.
479 *Food Hydrocolloids, 28*, 1-13.

480 Lama-Muñoz, A., Rodríguez-Gutiérrez, G., Rubio-Senent, F., Palacios-Díaz, R.,
481 & Fernández-Bolaños, J. (2013). A study of the precursors of the natural antioxidant
482 phenol 3,4-dihydroxyphenylglycol in olive oil waste. *Food Chemistry, 140*, 154-160.

483 Liu, C., & Wyman, C.E. 2003. The effect of flow rate of compressed hot water on
484 xylan, lignin and total mass removal from corn stover. *Industrial & Engineering*
485 *Chemistry Research, 42*, 5409– 5416.

486 Menéndez, J. A., Vázquez-Martín, A., García-Villalba, R., Carrasco-Pancorbo, A.,
487 Oliveras-Ferraros, C., Fernández-Gutiérrez, A., & Segura-Carretero, A. tabAnti-HER2
488 (erbB-2) oncogene effects of phenolic compounds directly isolated from commercial
489 extra-virgin olive oil (EVOO). *BCM Cancer*, *2008*, 8, 377.

490 Mulinacci, N., Innocenti, M., La Marca, G., Mercalli, E., Giaccherini, C., Romani,
491 A., et al. (2005). Solid olive residues: Insight into their phenolic composition. *Journal*
492 *of Agricultural and Food Chemistry, 53*, 8963–8969.

493 Obied, H. K., Bedgood, D. R. Jr., Prenzler, P. D., & Robards, K. (2007). Chemical
494 screening of olive biophenol extracts by hyphenated liquid chromatography. *Analytical*
495 *Chemical Acta, 603*, 176−189.

Obied, H. K., Prenzler, P. D., Ryan, D., Servili, M., Taticchi, A., Esposto, S., & Robards, K. (2008). Biosynthesis and biotransformations of phenol-conjugated oleosidic secoiridoids from Olea europaea L. *Natural Product Reports, 25*, 1167−1179.

Obied, H. K., Prenzler, P. D., Konczak, I, Rehman, A. U., & Robards, K. (2009). Chemistry and bioactivity of olive biophenols in some antioxidant and antiproliferative in vitro bioassays. *Chemistry and Research Toxicology, 22*, 227-234.

Owen, RW., Haubner, R., Mier, W., Giacosa, A., Hull, W. E ., Spiegelhalder, B., & Bartsch, H. (2003). Isolation, structure elucidation and antioxidant potential of the major phenolic and flavonoid compounds in brined olive drupes. *Food and Chemical Toxicology, 41*, 703–717.

Rubio-Senent, F., Lama-Muñoz, A., Rodríguez-Gutiérrez, G. & Fernández-Bolaños, J. (2013). Isolation and identification of phenolic glucosides from thermally treated olive oil byproducts. *Journal of Agricultural and Food Chemistry, 61*, 1235-1248.

Rubio-Senent, F., Rodríguez-Gutiérrez, G., Lama-Muñoz, A., & Fernández-Bolaños, J. (2012). New phenolic compounds hydrothermally extracted from the olive oil by-product alperujo and their antioxidative activities. *Journal of Agricultural and Food Chemistry, 60,* 1175–1186.

Rubio-Senent, F., Rodríguez-Gutiérrez, G., Lama-Munoz, A., García, A., & Fernández-Bolaños, J. (2015). Pectin extracted from thermally treated olive oil by-products: characterization physico-chemical properties, in vitro bile acid and glucose binding, *Food Hydrocolloids*, DOI: 10.1016/j.foodhyd.2014.06.001.

Ruiz-Canela, M., Estruch, R., Corella, D., Salas-Salvado, J., & Martinez-Gonzalez, MA. (2014). Association of Mediterranean diet with peripheral artery disease: The PREDIMED randomized trial. *Journal of the American Medical Association*, 311-415.

Ryan, D., Antolovich, M., Herlt, T., Prenzler, P.D., Lavee, S., & Robard, K. (2002). Identification of phenolic compounds in tissues of the novel olive cultivar Hardy's Mammoth. *Journal Agricultural and Food Chemistry*, 50, 6716-6724.

Savarese, M., De Marco, E., & Sacchi, R. (2007). Characterization of phenolic extracts from olives (Olea europaea cv. Pisciottana) by electrospray ionization mass spectrometry. *Food Chemistry*, 105, 761−770.

Servili, M., Baldioli, M., Selvaggini, R., Macchioni, A., & Montedoro, G. (1999). Phenolic compounds of olive fruit: One- and two-dimensional nuclear magnetic resonance characterization of nüzhenide and its distribution in the constitutive parts of fruit. *Journal Agricultural and Food Chemistry*, 47, 12–18.

Figure Captions.

Figure 1: Different derivatives from oleuropein. Molecular weight indicated in brackets.

Figure 2: A, ^{1}H-RMN Molecule 1. B, ^{1}H-RMN Molecule 2.

Figure 3: Acetylation of dihydroelenolic acid glucoside using acetic anhydride and pyridine at room temperature (rt).

539 Figure 4: Through-space correlation found in 2D-NOESY experiment for
540 Molecule 1 (a) and Molecule 2 (b).

541

Figure(s)

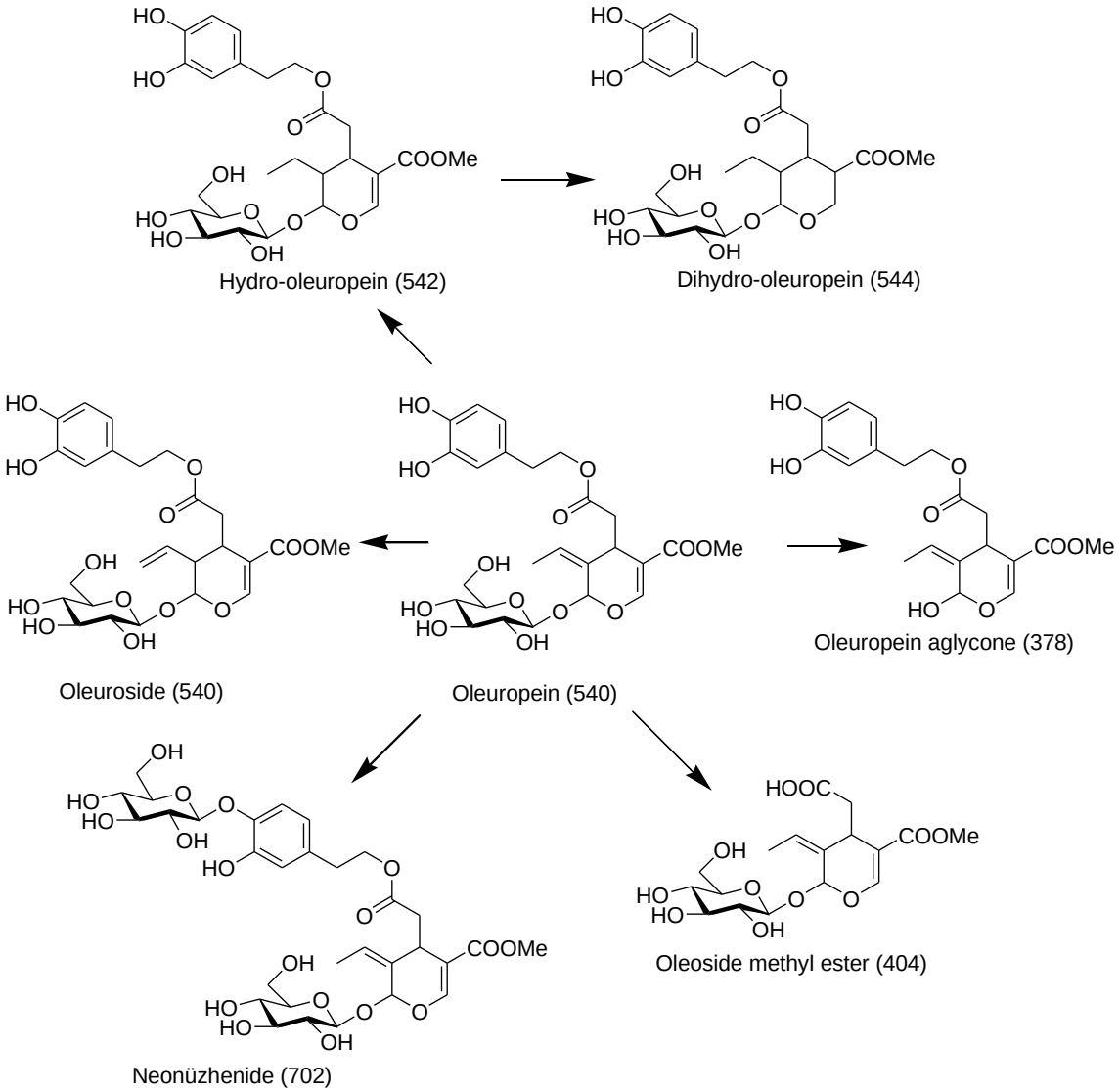
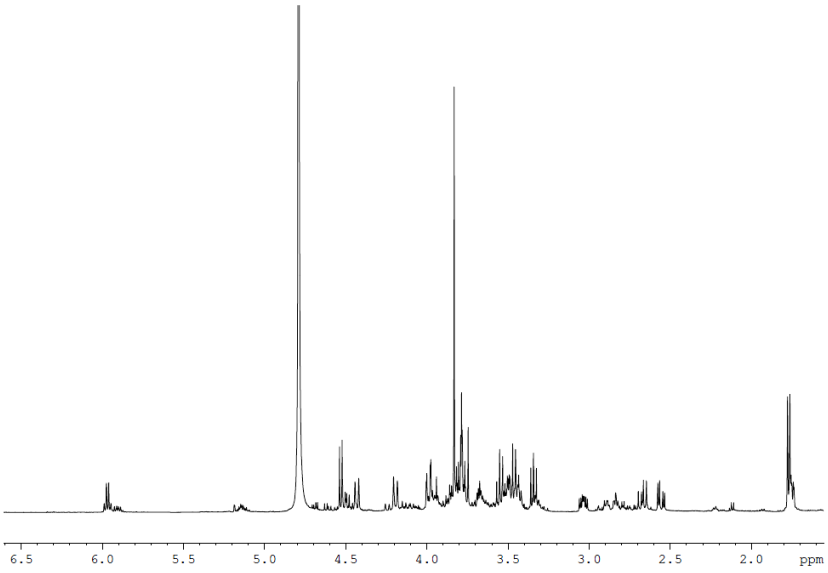


Figure 1

A



B

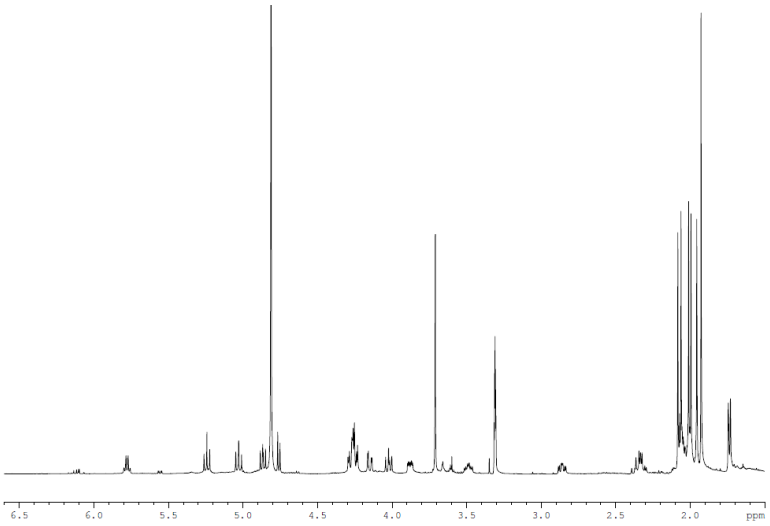


Figure 2

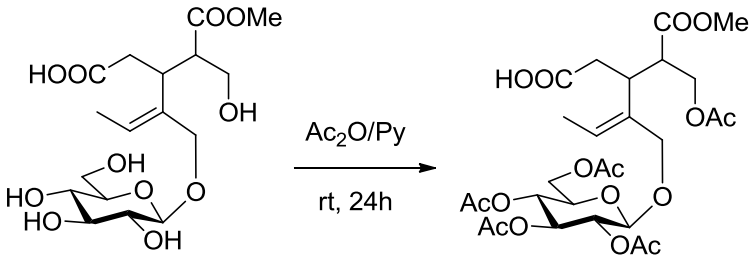

COOMe
HOOC
OH
OH
HO
O
O
HO
OH
Ac2O/Py
rt, 24h
COOMe
HOOC
OAc
OAc
AcO
O
O
AcO
OAc

Figure 3

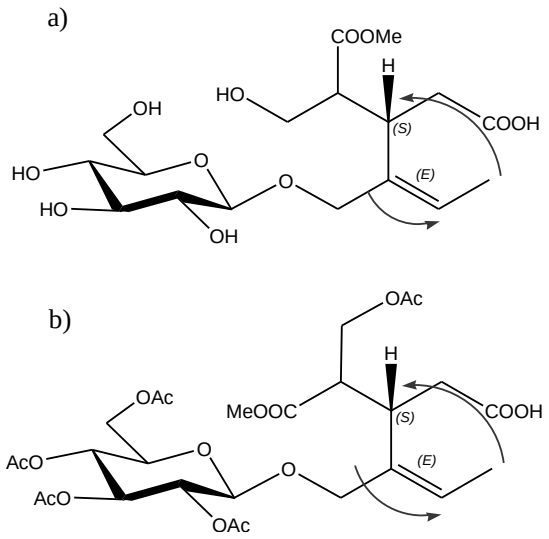

a)
COOMe
H
HO
OH
(S)
COOH
HO
O
(E)
O
HO
OH
b)
OAc
H
OAc
MeOOC
(S)
COOH
AcO
O
(E)
O
AcO
OAc

Figure 4.

Supplementary Material

Click here to download Supplementary Material: Supporting information.doc