



Salicin-7-sulfate: A new salicinoid from willow and implications for herbal medicine



Clarice Noleto-Dias, Jane L. Ward, Alice Bellisai, Charlotte Lomax, Michael H. Beale*

Department of Computational and Analytical Sciences, Rothamsted Research, West Common, Harpenden, Hertfordshire AL5 2JQ, UK

ARTICLE INFO

Keywords:

Salicin
Salicin-7-sulfate
Salix
Willow

ABSTRACT

Willow (*Salix* sp.) is a historically well-known herbal medicine that provided the lead compound (salicin) for the discovery of aspirin, one of the most successful plant derived drugs in human medicine. During a metabolomics screen of 86 *Salix* species contained in the UK National Willow Collection, we have discovered, isolated and fully characterised a new natural salicinoid – salicin-7-sulfate. This molecule may have important human pharmacological actions that need to be considered in determining the efficacy and safety of willow herbal medicines.

1. Introduction

Comminuted or powdered barks from *Salix* (willow) species, especially *S. alba*, *S. nigra*, *S. purpurea*, *S. daphnoides* and *S. fragilis* are well-known phyto-medicines with a history of ethno-medical use that stretches back to ancient Greek, Assyrian and Egyptian civilisations. The story of the identification of salicin **1** as an active analgesic from willow, and the introduction of the synthetic analogue aspirin (acetyl salicylate) **2**, which was to become a huge pharmaceutical success, has been well documented [1,2]. Both salicin and aspirin act as pro-drugs, being metabolised in humans to salicylate - the active pharmacophore that competitively inhibits cyclooxygenase [3,4], whilst aspirin itself also has a more direct action on cyclooxygenase, via irreversible acetylation of the active site [3].

The broad Salicaceae woody plant family contains several hundred species of *Salix* as well as the smaller *Populus* genus (e.g. poplar, aspen, cottonwoods). The family is characterised by the presence of phenolic glycosides, including, in many cases, the salicinoid sub-group of which salicin **1** represents the basic structure in a modular array of more complex analogues [5]. Possibly because of the success of aspirin, the potential for the discovery of new pharmacologically active compounds in the Salicaceae has been largely unexplored, although it has been suggested recently that the bioactivity of herbal extracts of willow cannot be accounted for by the levels of salicin alone [6]. In addition to pain relief, the use of aspirin in mitigation of thrombo-embolism is also well established and, more recently, both salicin and aspirin have been investigated for the prevention of cancer [7–10].

As part of a programme dedicated to high value products from plants we have focussed on novel phytochemistry in the Salicaceae, in

particular, those species contained in the 1500+ National Willow Collection (NWC), maintained as a short-rotation coppice plantation at Rothamsted Research. Taking a metabolomics approach [11] to polar extracts using NMR and high mass accuracy LC-MS-MS we have constructed a large annotated data-resource that spans the NWC and the full breadth of phenolic glycoside diversity. In this paper, we report on the discovery and structure determination of salicin-7-sulfate **3** (Fig. 1), a close analogue of salicin **1**, that potentially has a different metabolic fate in humans and thus requires further investigation in the context of efficacy and safety of the herbal materials.

2. Materials and methods

2.1. General experimental procedures

¹H-1D and ¹H-¹H & ¹H-¹³C 2D-NMR spectra were acquired on a Bruker Avance 600 MHz NMR spectrometer (Bruker Biospin, Germany), operating at 600.05 MHz for ¹H and 150.9 MHz for ¹³C NMR spectra, using a 5 mm selective inverse probe. 1D ¹H spectra were collected using 128 scans and by using the zgpr pulse sequence with a 90° angle. The residual HOD signal was suppressed by pre-saturation during a 5 s delay. Spectra consisted of 64,000 data points with a spectral width of 12 ppm. FIDs were automatically Fourier transformed using an exponential window function with a line broadening of 0.5 Hz. Phasing and baseline correction were carried out within the instrument software. 2D COSY, HSQC and HMBC spectra were collected using standard Bruker parameter sets and acquisition details are given in Supporting Information. All spectra were collected at 300 °K in D₂O:CD₃OD (8:2) and chemical shifts are given in δ, relative to d₄-TSP [(trimethylsilyl)]

* Corresponding author.

E-mail address: mike.beale@rothamsted.ac.uk (M.H. Beale).

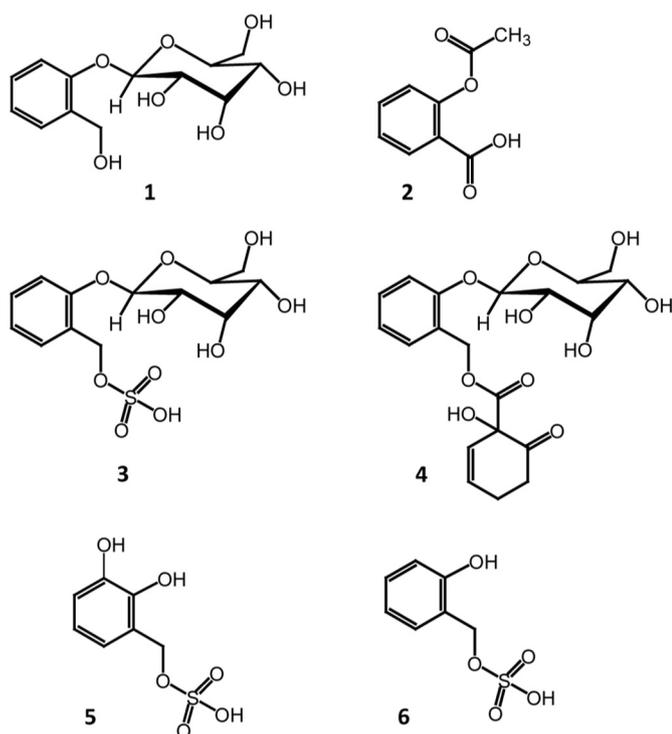


Fig. 1. Chemical structures described in this paper.

propionic acid, 0.01 % w/v] added as a chemical shift reference standard. NMR data was processed using TOPSPIN v. 2.1 (Bruker Biospin, Germany), MestReNova v. 6.0.2 (Mestrelab Research SL, Spain) and ACD NMR Processor (ACD Labs, Toronto, Canada)

UHPLC–MS were recorded with a Dionex UltiMate 3000 RS UHPLC system, equipped with a DAD-3000 photodiode array detector, coupled to an LTQ-Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Germany). UHPLC separation was carried out using a reversed-phase Hypersil GOLD™ column (1.9 μm, 30 × 2.1 mm i.d. Thermo Fisher Scientific, Germany) which was maintained at 35 °C. The solvent system consisted of water/0.1% formic acid (A) and acetonitrile/0.1% formic acid (B), both Optima™ grade (Thermo Fisher Scientific, Germany). Separation was carried out for 40 min under the following conditions: 0–5 min, 0% B; 5–27 min, 31.6% B; 27–34 min, 45% B; 34–37.5 min, 75% B. The flow rate was 0.3 mL/min, and the injection volume was 10 μL. Mass spectra were collected using an LTQ-Orbitrap Elite with a heated ESI source (Thermo Fisher Scientific, Germany). Mass spectra were acquired in negative mode with a resolution of 120,000 over m/z 50–1500. The source voltage, sheath gas, auxiliary gas, sweep gas and capillary temperature were set to 2.5 kV, 35 (arbitrary units), 10 (arbitrary units), 0.0 (arbitrary units) and 350 °C, respectively. Default values were used for other acquisition parameters. Automatic MS–MS was performed on the four most abundant ions and an isolation width of m/z 2 was used. Ions were fragmented using high-energy C-trap dissociation with a normalised collision energy of 65 and an activation time of 0.1 ms. Data was collected and inspected using Xcalibur v. 2.2 (Thermo Fisher Scientific, Germany). Data were analysed with the SIEVE™ 2.0 software (Thermo Fisher Scientific) using the Chromatographic Alignment and Framing algorithm. Frames were calculated from 0 to 40 minutes, between m/z 50 and 1500. Framing parameters were set at frame width of 2.5 minutes and m/z of 100 ppm, and peak intensity threshold of 2682520.

Compound isolation was carried out using an HPLC system (Dionex UltiMate 3000, Thermo Fisher Scientific) equipped with an Ascentis C-18 column (5 μm, 5 × 250 mm i.d., Supelco, UK) maintained at 25 °C. The chromatographic separation was performed by using a constant flow rate of 1 ml/min of the mobile phases water (A) and acetonitrile

(B), both containing 0.1% formic acid. The binary gradient was: 10 min, isocratic of 2% B; 10 to 30 min, linear from 2 to 5% B, followed by 15 min of 5% B. Peaks were detected using wavelengths of 210 to 310 nm and the peak corresponding to salicin-7-sulfate was collected, in automation, by time (37.5–42 min) into glass tubes. Eighteen injections (100 μL each) were performed and fractions from repeated runs were combined and the solvent evaporated using a Speedvac concentrator (Genevac, Suffolk, UK).

Optical rotation was measured in water on an Anton Paar MCP-100 polarimeter using a 100 mm sample cell.

2.2. Plant material

Multiple dormant stems were harvested in February 2015 from the National Willow Collection (NWC) maintained at Rothamsted Research, Harpenden, UK (RRes), UK. Plants had been previously coppiced in February 2014. Each plot of the collection contains 10 plants that were generated from separate stem cuttings. Stem tissue portions (10–15 cm) were harvested from the top of each plant in a plot and combined to give a single sample. Tissue was kept at –80 °C prior to freeze-drying to remove residual water. After lyophilisation plant material was milled to a fine power (Ultra Centrifugal Mill ZM200, Retsch, UK). Milled tissue was maintained at –80 °C until analysis. Voucher specimens of lyophilised material have been retained and are available on request.

2.3. Metabolite extraction and isolation

For initial metabolite profiling by NMR and UHPLC–MS triplicate aliquots of milled freeze-dried willow stem powder (30 mg) were extracted as previously described [11]. Separate extractions were made for each analytical method. For compound isolation freeze-dried, milled, *Salix koriyanagi* (NWC1038) powder (270 mg) was extracted at 50 °C (10 min) in H₂O: MeOH (80:20, 5 mL). The sample was centrifuged (5 min) and the supernatant transferred to a new tube and heated at 90 °C (2 min). After cooling and centrifugation the supernatant (3.0 mL) was removed to a glass HPLC vial for purification by HPLC peak collection.

2.4. Spectroscopic data

Salicin-7-sulfate 3: Yellowish amorphous powder (0.9 mg), $[\alpha]_{25}^D - 32.8$ (c 0.0367, water), UHPLC–MS: RT 9.90 min, UV λ_{max} 210, 271 nm; m/z 365.0549 [M–H][–] calc'd for C₁₃H₁₇O₁₀S, 365.0542. ¹H NMR [600 MHz, (D₂O):CD₃OD = 8:2] δ 3.52 (1H, m, H-4'), 3.59–3.63 (2H, m, H-3',5'), 3.65 (1H, dd, $J = 9.3, 7.7$ Hz, H-2'), 3.76 (1H, dd, $J = 12.5, 5.7$ Hz, H-6' β), 3.93 (1H, dd, $J = 12.5, 2.2$ Hz, H-6' α), 5.12 (1H, d, $J = 11.2$ Hz, H-7), 5.22 (1H, d, $J = 11.2$ Hz, H-7), 5.10 (1H, d, $J = 7.6$ Hz, H-1'), 7.17 (1H, td, $J = 7.5, 1.0$ Hz, H-4), 7.24 (1H, d, $J = 8.0$ Hz, H-6), 7.43 (1H, td, $J = 8.3, 1.6$ Hz, H-5), 7.48 (1H, dd, $J = 7.6, 1.6$ Hz, H-3). See Table 2 for ¹³C NMR and Supplementary information file for 2D spectra.

3. Results and discussion

Data from a standardised ¹H-NMR fingerprinting method for aqueous methanolic extracts of willow [11], that gives quantitative data on a mixture of primary and secondary metabolites, was mined to examine the variation in salicin concentration in stem tissue samples, across 86 pure (i.e. non-hybrid) *Salix* genotypes in the NWC, harvested at the dormant stage (February), a time-point when biomass willows are generally cropped. Quantitative data derived from ¹H-NMR via integration of the distinctive and isolated benzylic hydrogens of salicin 1 (84.74 and 84.69) against internal d₄-trimethylsilylpropionate standard, are given in Table 1. Salicin levels varied from 2.85 (*S. maccaliana*) to 57.6 (*S. acutifolia* Willd.) mg/g dry weight, (i.e. 0.29% to 5.8% dry weight) of whole stem tissue. Interestingly, the *Salix* sp. (*alba*,

Table 1

Concentration of salicin, **1**, in 86 accessions from the National Willow Collection (NWC) held at Rothamsted Research (RRes). Data is obtained from ¹H-NMR analysis (600 MHz) of a D₂O:CD₃OD (4:1) extract of dormant stem tissue.

ID	NWC plot code	RRes no	Species	Variety	Salicin, 1 mg/g d.w.
1	M30	1165	<i>S. arbusculoides</i>		15.35 ± 0.08
2	M130	1043	<i>S. wilsonii</i>		7.88 ± 1.00
3	M46	1062	<i>S. rosmarinifolia</i>		45.04 ± 0.97
4	M147	1236	<i>S. rhamnifolia</i> Pall.		8.51 ± 2.41
5	M28	-	<i>S. waldsteiniana</i> Willd.		2.88 ± 0.83
6	NWC742	-	<i>S. bebbiana</i>		8.14 ± 0.87
7	GH1346	-	<i>S. kalarica</i>		11.48 ± 0.14
8	GH1239	1239	<i>S. saposchnikovii</i>		6.24 ± 0.68
9	M56	415	<i>S. magnifica</i> Hemsl.		8.18 ± 0.21
10	M36	500319	<i>S. balfourii</i>		10.90 ± 1.55
11	M164	500340	<i>S. humilis</i>	microphylla	19.40 ± 0.26
12	M60	746	<i>S. gracilistyla</i> Miq.	Neko-Yanagi	6.83 ± 0.17
13	NWC1011	791	<i>S. alberti</i>		12.72 ± 1.98
14	M38	823	<i>S. caesia</i> Vill.	Misurina. Belluna	18.15 ± 0.54
15	M39	830	<i>S. kochiana</i> Traut.		18.79 ± 0.04
16	M113	888	<i>S. pchnostachya</i>		9.01 ± 0.14
17	NWC1096	889	<i>S. suchowensis</i>		13.23 ± 0.52
18	NWC1037	828	<i>S. integra</i> Thunb.		15.13 ± 0.39
19	M121	984	<i>S. alaxensis</i> Anders.		18.55 ± 1.80
20	M129	-	<i>S. elaeagnos</i> Scop.		11.68 ± 1.08
21	M27	-	<i>S. nakamuraana</i>		4.32 ± 0.22
22	M47	-	<i>S. hastata</i> L.		26.87 ± 2.32
23	M54	-	<i>S. commutata</i> Bebb		40.93 ± 3.31
24	NWC1231	-	<i>S. excelsa</i>		11.07 ± 0.00
25	NWC1306	1317	<i>S. lasiocarpa</i>	Musgroves Orange	26.11 ± 0.00
26	NWC1308	1339	<i>S. patula</i>		3.92 ± 0.00
27	NWC1309	-	<i>S. tetrapla</i>		7.32 ± 0.00
28	NWC1315	-	<i>S. discolor</i>		6.88 ± 0.00
29	NWC1316	-	<i>S. ehrhartiana</i>		25.13 ± 0.00
30	NWC1317	-	<i>S. exigua</i>		9.95 ± 0.00
31	NWC1318	-	<i>S. lasiocarpa</i>		10.95 ± 0.00
32	NWC1319	-	<i>S. meyeriana</i>		28.63 ± 0.00
33	NWC1320	-	<i>S. pellita</i>		25.63 ± 0.00
34	NWC1321	-	<i>S. pendulina</i>		11.64 ± 0.00
35	NWC1322	-	<i>S. wardiana</i>		11.17 ± 0.00
36	NWC695	500579	<i>S. pyralifolia</i>		8.79 ± 0.00
37	NWC1195	1017	<i>S. phyticifolia</i> L.		4.12 ± 1.01
38	NWC1196	1018	<i>S. reinii</i> Franch. et Savat.		15.17 ± 0.19
39	NWC1165	1024	<i>S. hookeriana</i> Barratt ex Hook.		28.16 ± 1.12
40	NWC1202	1038	<i>S. sitchensis</i> Sanson ex Bong.		24.40 ± 4.03
41	NWC285	11	<i>S. lucida</i> Muhl.		38.53 ± 5.82
42	NWC1214	1130	<i>S. maccaliana</i>		2.85 ± 0.05
43	NWC1215	1131	<i>S. wimmeriana</i>		21.35 ± 0.01
44	NWC1216	1155	<i>S. acutifolia</i> Willd.		57.57 ± 1.88
45	NWC1219	1157	<i>S. adhenophylla</i> Hook.		6.95 ± 0.60
46	NWC722	1184	<i>S. cinerea</i> L.		8.50 ± 1.39
47	M137	1187	<i>S. coesia</i> Vill.		14.45 ± 0.63
48	NWC1236	1208	<i>S. irrorata</i> Anders.		22.91 ± 4.30
49	NWC1237	1209	<i>S. kangensis</i> Nakai		37.39 ± 1.45
50	NWC1241	1218	<i>S. nipponica</i> (Franch. et Sav.) A.Skvorts.		10.13 ± 1.65
51	NWC1245	1229	<i>S. pierotii</i> Miq.		12.73 ± 1.62
52	NWC1301	1285	<i>S. udensis</i> Trautv. et Mey.		4.96 ± 0.35
53	NWC1302	1301	<i>S. vinogradovii</i> A.Skvorts.		29.87 ± 2.12
54	M162	1308	<i>S. repens</i> L.	Argentea	6.13 ± 0.93
55	M93	14	<i>S. pentandra</i> L.	Dark French	21.19 ± 3.45

Table 1 (continued)

ID	NWC plot code	RRes no	Species	Variety	Salicin, 1 mg/g d.w.
56	NWC470	191	<i>S. alba</i> L.	Kew	8.40 ± 0.57
57	NWC279	2	<i>S. nigra</i> Marsh.	SN3 Primrose Hill	28.90 ± 4.79
58	NWC488	210	<i>S. alba</i> L. var. <i>coerulea</i>	Wantage Hall	12.21 ± 0.62
59	NWC295	23	<i>S. amygdaloides</i> Anders.		20.07 ± 3.95
60	M55	261	<i>S. babylonica</i> L.	Annularis	7.96 ± 0.09
61	NWC647	386	<i>S. fragilis</i> L.	Cox	19.06 ± 0.65
62	M69	420	<i>S. daphnoides</i> Vill.	Ruberrima	24.70 ± 5.93
63	NWC688	441	<i>S. rorida</i> Lacksch.		48.26 ± 2.30
64	NWC692	446	<i>S. aegyptiaca</i> L.		9.25 ± 0.11
65	NWC698	449	<i>S. appendiculata</i> Vill.	Venzonassa	3.57 ± 0.64
66	NWC701	452	<i>S. apennina</i> A.Skvorts.	Pescara	26.87 ± 1.47
67	NWC704	455	<i>S. aurita</i> L.	Innis Moor	6.67 ± 0.24
68	NWC708	459	<i>S. caprea</i> L.	Smithiana	6.31 ± 0.97
69	NWC733	481	<i>S. scouleriana</i> Barrat ex Hook.		32.45 ± 0.70
70	NWC326	56	<i>S. triandra</i> L. f. <i>concolor</i>	Baldwin	4.24 ± 0.41
71	NWC821	577	<i>S. dasyclados</i> Wimm.		6.68 ± 0.01
72	NWC841	607	<i>S. rehderiana</i> C.K. Schneider		10.54 ± 3.14
73	NWC849	615	<i>S. schwerinii</i> E.Wolf	K3 Hilliers	7.04 ± 0.29
74	NWC890	672	<i>S. viminalis</i> L.	Bowles Hybrid	5.01 ± 0.11
75	NWC963	741	<i>S. turanica</i> Nasarov		4.76 ± 0.06
76	NWC1032	820	<i>S. amplexicaulis</i>	Bory	6.61 ± 1.03
77	NWC1034	824	<i>S. gilgiana</i> Seemen		15.12 ± 2.25
78	NWC1038	831	<i>S. koriyanagi</i> Kimura ex Goerz		24.16 ± 0.00
79	NWC1053	844	<i>S. purpurea</i> L.	Uralensis	29.25 ± 1.56
80	NWC1095	886	<i>S. caspica</i> Pall.		14.14 ± 1.89
81	NWC1097	890	<i>S. tenuijulis</i> Ledeb.		13.99 ± 2.93
82	NWC1046	941	<i>S. miyabeana</i> Seemen	Purpurescens	28.11 ± 1.02
83	NWC1141	956	<i>S. eriocephala</i> Michx.	Mawdesley	12.80 ± 0.24
84	NWC1170	988	<i>S. drummondiana</i> Barratt ex Hook.		36.28 ± 1.84
85	NWC1174	993	<i>S. laggerii</i> Wimm.		7.03 ± 0.23
86	NWC1175	994	<i>S. mielichhoferii</i> Saut.	Seiseralp	11.70 ± 0.74

purpurea, *fragilis* and *daphnoides*), that are documented for medicinal use by the Herbal Medicinal Product Committee of European Medicines Agency [12], were not the highest in salicin content, ranging from 0.84 to 3% dry weight. The highest salicin contents were found in *S. acutifolia* (5.76%) and *S. rorida* (4.83%). *S. acutifolia* has previously been found to contain mainly salicin **1** and salicortin **4** in the emerging green shoots during the rapid growth season (May) [13] and this agrees with the finding here that much of the salicin remains in the matured stem tissue in the dormant season. High-resolution LC-MS data (negative ionisation mode) was also collected on comparative extracts of all samples for direct comparison with the ¹H-NMR data. The peak corresponding to salicin **1** appeared at 11.27 min (Fig. 2) and gave ions at *m/z* 285.0976 (C₁₃H₁₇O₇) corresponding to [M-H]⁻ and at *m/z* 331.1030 (C₁₄H₁₉O₈) which corresponded to the formate adduct (Fig. 3).

MSMS of *m/z* 331.1030 gave a single fragment ion at *m/z* 123.0455 corresponding to the C₇H₇O₂ salicyl alcohol aglycone moiety (Fig. 4). In many lines, a further peak corresponding to **3** was present in the same region of the Total Ion Chromatogram appearing at 9.90 min (Fig. 2). The mass spectrum of **3** contained an ion at *m/z* 365.0549 and a formula corresponding to C₁₃H₁₇O₁₀S (Fig. 3). The presence of sulfur in

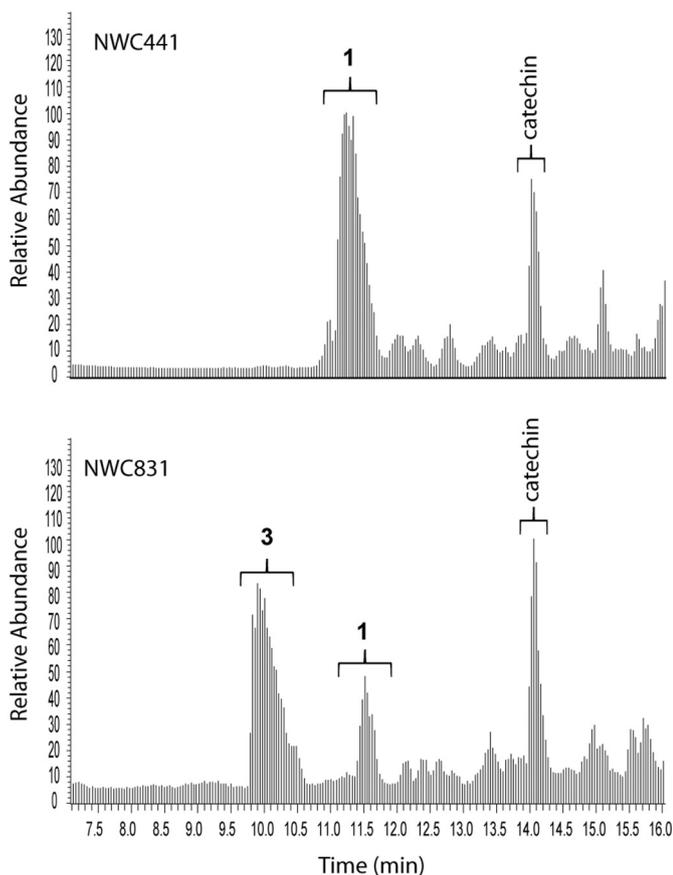


Fig. 2. Total Ion Chromatograms (negative ion mode) of two accessions (NWC441 and NWC 831) from the natural willow collection. 1: salicin, 3: salicin-7-sulfate.

the molecule was confirmed via inspection of the $M + 2$ isotope region. In addition to the ions at m/z 367.0580 and m/z 367.0635,

corresponding the ^{18}O ($\text{C}_{13}\text{H}_{17}\text{O}_9^{18}\text{OS}$) and ^{13}C ($\text{C}_{11}^{13}\text{C}_2\text{H}_{17}\text{O}_{10}\text{S}$) isotopes respectively, an ion at m/z 367.0500 was present and corresponded to an entity with molecular formula $\text{C}_{13}\text{H}_{17}\text{O}_{10}^{34}\text{S}$. These $M + 2$ ions and their relative intensities, are consistent for S-containing metabolites when data is collected on MS instruments with a high resolving power (such as FT-ICR-MS or Orbitrap) and have previously been suggested to confirm the molecular formulae of sulfur bearing metabolites [14].

The MSMS spectrum (Fig. 4) showed a base peak at m/z 96.9607 $[\text{SO}_4\text{H}]^-$. Other peaks at m/z 347.0441 ($\text{C}_{13}\text{H}_{15}\text{O}_9\text{S}$), 203.0028 ($\text{C}_7\text{H}_7\text{O}_5\text{S}$) and 123.0455 ($\text{C}_7\text{H}_7\text{O}_2$) corresponded to $[\text{M}-\text{H}_2\text{O}]^-$, $[\text{M}-\text{glucose}]^-$ and $[\text{salicyl}]^-$ respectively. The MS data thus indicated that 3 is a sulfated derivative of salicin. The presence of a further ion at m/z 259.0140 ($\text{C}_6\text{H}_{11}\text{O}_9\text{S}$), corresponding to sulfated glucose, seemed to indicate that the sulfate group was possibly located on the glucose moiety of salicin. However, this ion could also arise from a rearrangement and neutral loss of orthoquinone methide from a 7-sulfate as shown in Fig. S1, Supporting information. Small ions arising from an analogous neutral loss are present in the published MS data of other salicinoids, e.g. a $423 \rightarrow 317$ transition in salicortin 4 that have not, to date, been rationalised [15], but can be explained by a similar loss of orthoquinone methide and transfer of the 7-ester group to glucose. Thus, from MS data alone it was impossible to assign the position of the sulfate group with certainty. The ions obtained in the MS-MS spectrum suggested three possible structures, that of salicin-7-sulfate, isosalicin-1-sulfate (i.e. 7-glucosylsalicyl alcohol-1-sulfate) or salicin-2'/6'-sulfate. The final structure was determined via isolation using repeated HPLC injections and structural characterisation by 1D and 2D-NMR.

The $^1\text{H-NMR}$ (Table 2 and Fig. S2, Supporting information) was compared to that of salicin 1 to determine the position of sulfation. A clear 0.44 ppm downfield shift of the two $J = 11$ Hz doublet signals relating to the methylene group at C-7 were observed. These signals now appeared at δ 5.216 and δ 5.123 in contrast to those observed in salicin (δ 4.734 and δ 4.681). Signals relating to the glucoside moiety were largely unchanged as were those of the aromatic salicyl unit. Thus, the $^1\text{H-NMR}$ data suggested a structure of salicin-7-sulfate 3. ^{13}C NMR data was obtained from HSQC and HMBC data and is presented in

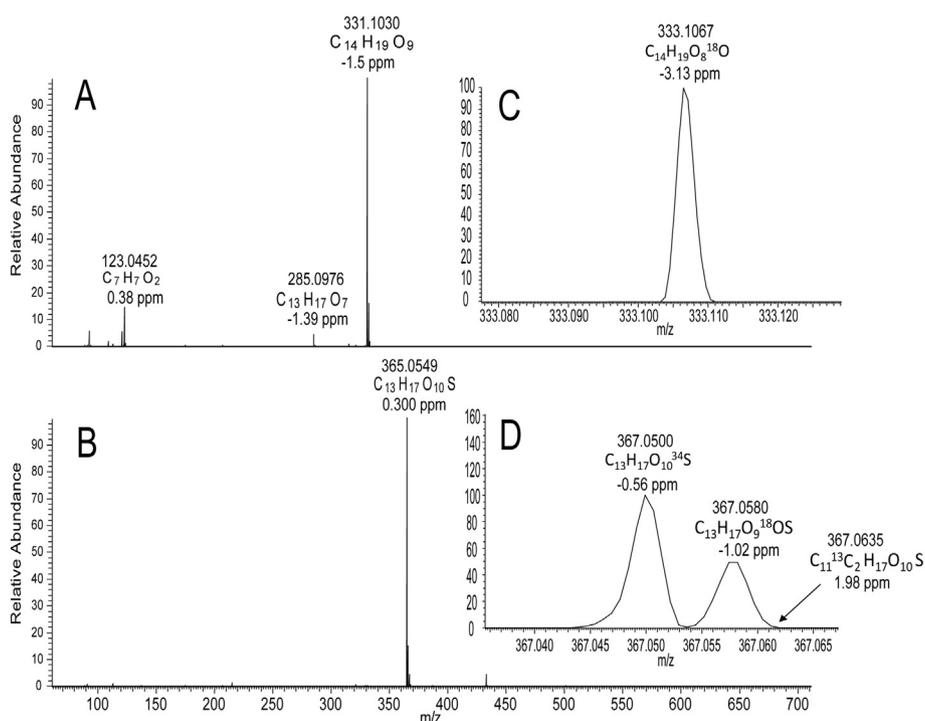


Fig. 3. A: MS spectrum of 1 (m/z 331, 11.27 min); B: MS spectrum of 3 (m/z 365, 9.90 min); C: Enlarged region showing m/z 333 ion $[(M+2)+\text{formate-H}]$ of 1; D: Enlarged region of m/z 367 ions relating to $[(M+2)-\text{H}]$ ions of 3. Data was collected in negative ion mode.

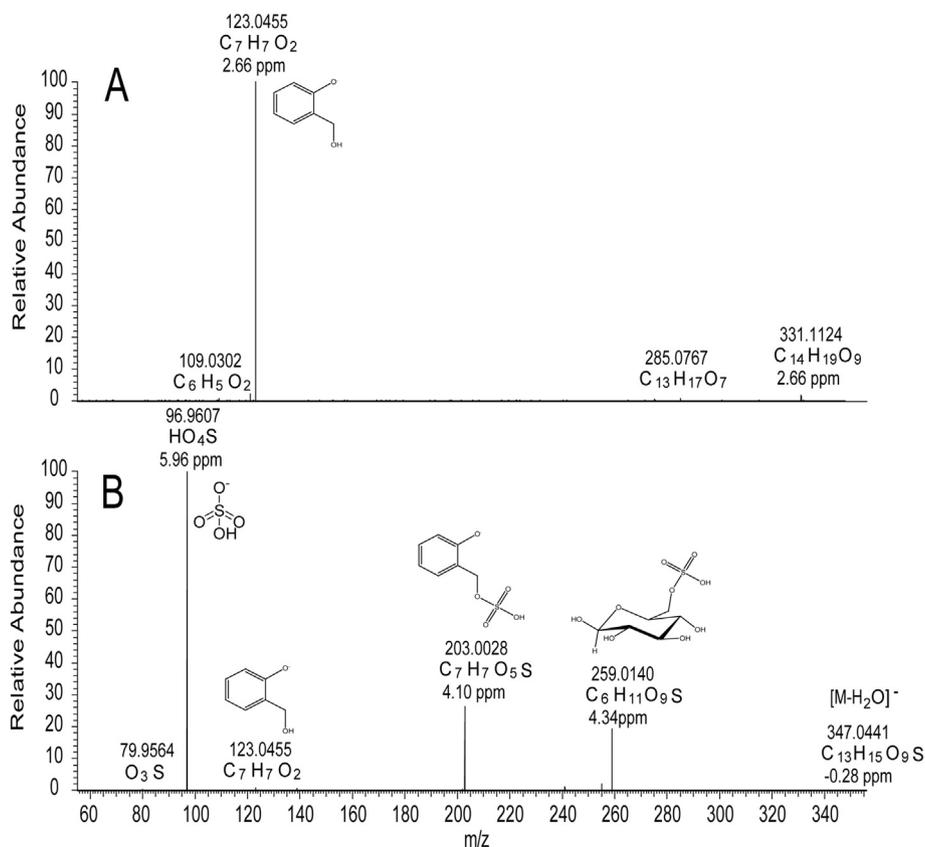


Fig. 4. A: MSMS spectrum of **1** (m/z 331, 11.27 min); B: MSMS spectrum of **3** (m/z 365, 9.90 min). Data was collected in negative ion mode.

Table 2 and Figs. S4 and S5, supporting information. The most significant difference was a 6.8 ppm downfield shift of the signal relating to C-7 which appeared at 68.8 ppm in comparison to the equivalent carbon in salicin (62.0 ppm). An associated upfield shift of 6.4 ppm was observed for the signal corresponding to C-2 which now appeared at 127.5 ppm. Salicin-7-sulfate **3** has not previously been reported in the literature. However, ^1H and ^{13}C NMR data is available for other sulfated natural products. The structurally related idesin hydrogen sulfate **5** isolated from the fruits of *Idesia polycarpa* Maxim. (Flacourtiaceae) showed comparable changes in its NMR spectra (Table 2) when

compared to the non-sulfated compound, i.e. a downfield shift of 0.54 and 5.7 ppm for the protons and carbon, respectively, attached to the sulfate group [16]. Similarly, the position of a sulfate group at C-6 of glucose in a triterpene glycoside, isolated from the whole plant of *Bacopa monnieri* (L.) Wettst. (Scrophulariaceae), was also confirmed based on a downfield shift of 3.4 ppm in the carbon directly attached to the sulfate group [17]. Although sulfation of natural products is not rare, and occurs often in mammalian metabolism, most examples from the plant world concern sulfated flavonoids [18]. Other instances from plant pathways include a sulfated anthraquinone [19] and a sulfate of

Table 2
Chemical shift data of salicin **1**, salicin-7-sulfate **3** and idesin hydrogen sulfate **5**.

Position	1 ^a			3 ^a			5 ^{b,c}		
	δ_{C}	δ_{H}	$J_{\text{H-H}}$ (Hz); multiplicity ^d	δ_{C}	δ_{H}	$J_{\text{H-H}}$ (Hz); multiplicity ^d	δ_{C}	δ_{H}	$J_{\text{H-H}}$ (Hz); multiplicity ^d
1	157.6	-	-	157.8	-	-	144.1	-	-
2	133.9	-	-	127.5	-	-	132.0	-	-
3	132.1	7.40	7.5; 1.5; <i>dd</i>	133.7	7.48	7.6; 1.6; <i>dd</i>	121.1	6.96	7.5; 2.0; <i>dd</i>
4	126.0	7.15	7.5; 1.0; <i>td</i>	126.1	7.17	7.5; 1.0; <i>td</i>	126.6	7.01	7.5; <i>t</i>
5	132.3	7.37	8.3; 1.7; <i>td</i>	133.7	7.43	8.3; 1.6; <i>td</i>	117.6	6.86	7.5; 2.0; <i>dd</i>
6	118.0	7.21	8.0; <i>d</i>	118.1	7.24	8.0; <i>d</i>	150.6	-	-
7	62.0	4.68	12.7; <i>d</i>	68.8	5.12	11.2; <i>d</i>	66.2	5.32 5.19	12.0; <i>d</i>
		4.73	12.7; <i>d</i>		5.22	11.2; <i>d</i>			12.0; <i>d</i>
1'	103.4	5.08	7.4; <i>d</i>	103.7	5.10	7.6; <i>d</i>	106.9	4.63	7.5; <i>d</i>
2'	75.7	3.56-3.63	Overlapped	76.2	3.65	7.7; 9.3; <i>dd</i>	75.3	3.55	<i>m</i>
3'	78.6	3.56-3.63	Overlapped	78.7	3.59-3.63	<i>m</i>	77.7	3.45-3.36	<i>m</i>
4'	72.1	3.50	<i>m</i>	72.5	3.52	9.0, 9.8, <i>dd</i>	70.9	3.45-3.36	<i>m</i>
5'	78.6	3.56-3.63	<i>m</i>	78.7	3.59-3.63	<i>m</i>	78.3	3.32	<i>m</i>
6'	63.3	3.75 3.91	12.4; 5.7; <i>dd</i>	63.9	3.76 3.93	12.5; 5.7; <i>dd</i> 12.5; 2.2; <i>dd</i>	62.1	3.88 3.75	12.3; 2.3; <i>dd</i>
			12.4; 2.2; <i>dd</i>						12.3; 5.0; <i>dd</i>

^a Data collected in 80:20 D₂O:CD₃OD (4:1). Spectra were referenced to d₄-TSP at δ 0.00.

^b Data collected in d₄-MeOH. Spectra were referenced to d₄-MeOH.

^c Chou et al. [16].

^d *d* doublet; *dd* double doublet; *m* multiplet; *dt* doublet of triplets; *t* triplet.

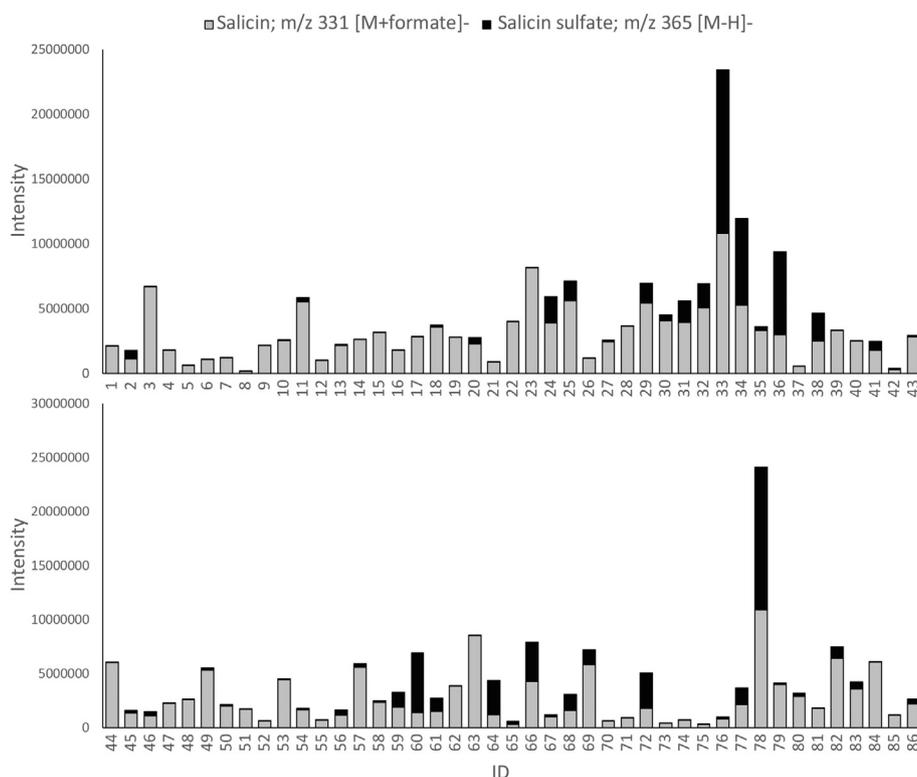


Fig. 5. Ion intensities of salicin **1** (m/z 331, grey) and salicin-7-sulfate **3** (m/z 365, black) from negative mode LC-MS data. ID numbers refer to entries in Table 1.

deoxylactucin – a sesquiterpene from lettuce [20]. This report and that of the related idesin hydrogen sulfate [16] are the first examples from simple phenolic metabolism although there are examples of sulfated lignans (e.g. [21,22]).

Unlike salicin, it was not possible, due to overlapping signals, to quantify salicin-7-sulfate in the NMR spectra. However, relative quantitation within the original LC-MS datasets arising from the 86 genotypes screened, was used to gain insight into the abundance of the sulfated form (Fig. 5). It can be seen, by correlating the NMR-quantified levels of salicin shown in Table 1 with the relative intensities of the salicin peak in the output data table from Sieve™ software processing of the LC-MS data (see Fig. S6, supporting information for the correlation), that the LC-MS ‘quantitation’ in general shows good correlation with the NMR data, but there are several significant outliers such as lines 33 and 78 where LC-MS data is high, and lines 41 and 44, where LC-MS is giving low readings. The vagaries of differential ionisation and ion suppression are known influences on LC-MS quantitation and thus the data presented in Fig. 5 provides relative levels rather than absolute quantitation provided by NMR. It also should be noted that the sulfated form of salicin ionises much more easily than salicin and thus Fig. 5 serves as a guide to relative levels of the sulfate across samples rather than accurate quantitation. Nevertheless, the LC-MS data indicated that the ratio of salicin: salicin-7-sulfate was not fixed. Many of willow species produced only trace levels of the sulfated form. The highest amounts were observed in *S. pellita* (ID = 33, NWC1320) and *S. koriyana* (ID = 78, NWC1038). However, as a general rule across the dataset, the amount of salicin produced does not correlate with the amount of the sulfated form (as observed by LC-MS). For example, *S. commutata* (ID = 23, M54) and *S. rorida* (ID = 63, NWC688) both contained appreciable amounts of salicin, yet only trace amounts of salicin-7-sulfate. In contrast, *S. babylonica* L. var. *Annularis* (ID = 60, M55) showed low levels of salicin and a much higher proportion of salicin-7-sulfate.

Although, in many cases the levels of salicin-7-sulfate are very low, the presence of this compound in the varieties that are in both

traditional and commercial use as herbal medicines is of concern. Of the above mentioned traditional varieties used in herbal products, *S. alba* and *fragilis*, contained significant amounts of (2) with respect to salicin. The presence of the sulfate group in **3**, when metabolised by humans, is likely to lead to the formation of salicyl alcohol-7-sulfate **6** (Fig. 6) that is unlikely to be further metabolised to salicylate, but more likely to form orthoquinone methide, a reactive entity. There are four pharmacological aspects to consider – (i) that **6** is a close analogue of salicylate and thus itself is a cyclooxygenase inhibitor; (ii) that having the sulfate group gives **6** different anti-coagulant properties (c.f. heparin) to salicylate; (iii) that **6** is a much stronger acid than salicylate and thus potentially more harmful in gastro-intestinal bleeding side effects and (iv) that the elimination of sulfuric acid from **6** to form orthoquinone methide *in vivo* may result in co-valent binding to enzyme active sites. Although herbal extracts of *Salix* species also contain other more complex salicinoids, many of these break down to salicin and thus can be considered as further sources of active salicylate. However, the blocking acidic sulfate group means that **3** needs further investigation as this is likely to have direct effects in humans, and may partly explain the different efficacies of herbal willow and synthetic aspirin. Also, from a safety of herbal medicines perspective, the pharmacology of salicin-7-sulfate **3** warrants further investigation, and our efforts are now focussed on obtaining larger quantities of this new analogue for this purpose.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

The authors acknowledge the assistance of Mr. William Macalpine, Dr. Claudia Harflett and Mr Gianluca Ruvo who assisted in plant selection, tissue harvesting and NMR data collection respectively. Rothamsted Research receives grant-aided support from the

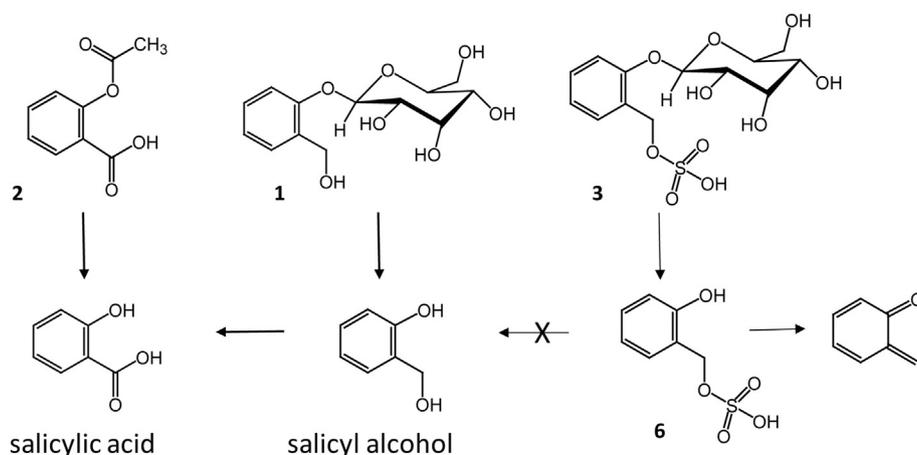


Fig 6. Human metabolism of salicin 1 and aspirin 2 to active salicylic acid and the likely metabolism of salicin-7-sulfate 3 to salicylate analogue 6 and orthoquinone methide.

Biotechnology and Biological Sciences Research Council (BBSRC) of the UK. C. Noleto-Dias is funded via a grant awarded by the Engineering and Physical Sciences Research Council [EP/K014676/1] and C. Lomax received funding from a grant awarded by the High Value Chemicals from Plants Network, a BBSRC Network in Industrial Biotechnology and Bioenergy (NIBB)[†]. Grant Ref: BB/L013665/1. The work described in this study was also supported by Institute Strategic Programme grants Cropping Carbon [BBS/E/C/00005199] and Tailoring Plant Metabolism [BB/E/C/00010410], funded by the BBSRC.

Appendix A. Supplementary data

Supplementary data related to this article including 1D and 2D NMR spectra for compound 3 can be found online at <https://doi.org/10.1016/j.fitote.2018.02.009>.

References

- [1] M.J.R. Desborough, D.M. Keeling, The aspirin story – from willow to wonder drug, *Br. J. Haematol.* 177 (2017) 674–683.
- [2] J.G. Mahdi, Medicinal potential of willow: a chemical perspective of aspirin discovery, *J. Saudi Chem. Soc.* 14 (2010) 317–322.
- [3] G.A. Higgs, J.A. Salmon, B. Henderson, J.R. Vane, Pharmacokinetics of aspirin and salicylate in relation to Inhibition of arachidonate cyclooxygenase and anti-inflammatory activity, *Proc. Natl. Acad. Sci. USA* 84 (1987) 1417–1420.
- [4] T. Akao, T. Yoshino, K. Kobashi, H. Hattori, Evaluation of salicin as an antipyretic prodrug that does not cause gastric injury, *Planta Med.* 68 (2002) 714–718.
- [5] G.A. Boeckler, J. Gershenzon, S.B. Unsicker, Phenolic glycosides of the Salicaceae and their role as anti-herbivore defenses, *Phytochemistry* 72 (2011) 1497–1509.
- [6] G.A. Bonterra, E.U. Heinrich, O. Kelber, D. Weiser, J. Metz, R. Kinscherf, Anti-inflammatory effects of the willow bark extract STW 33-1 (Proaktiv) in LPS-activated human monocytes and differentiated macrophages, *Phytomedicine* 17 (2010) 1106–1113.
- [7] C.-S. Kong, K.-H. Kim, J.-S. Choi, J.-E. Kim, C. Park, J.-K. Jeong, Salicin, an extract from white willow bark, inhibits angiogenesis by blocking the ROS-ERK pathways, *Phytother. Res.* 28 (2014) 1246–1251.
- [8] U. Wölfle, B. Haarhus, A. Kersten, B. Fiebich, M.J. Hug, C.M. Schempp, Salicin from willow bark can modulate neurite outgrowth in human neuroblastoma SH-SY5Y cells, *Phytother. Res.* 29 (2015) 1494–1500.
- [9] P.M. Rothwell, M. Wilson, C.E. Elwin, B. Norrving, A. Algra, C.P. Warlow, T.W. Meade, Long-term effect of aspirin on colorectal cancer incidence and mortality: 20-year follow-up of five randomised trials, *Lancet* 376 (2010) 1741–1750.
- [10] P.M. Rothwell, F.G.R. Fowkes, J.F.F. Belch, H. Ogawa, C.P. Warlow, T.W. Meade, Effect of daily aspirin on long-term risk of death due to cancer: analysis of individual patient data from randomised trials, *Lancet* 377 (2010) 31–41.
- [11] D.I. Corol, C. Harflett, M.H. Beale, J.L. Ward, An efficient high throughput metabolite typing platform for screening of biomass willows, *Metabolites* 4 (2014) 946–976.
- [12] European Medicines Agency, Assessment report on *Salix* [various species including *S. purpurea* L., *S. daphnoides* Vill., *S. fragilis* L.], cortex, http://www.ema.europa.eu/docs/en_GB/document_library/Herbal_-HMPC_assessment_report/2017/07/WC500230918.pdf, (2017).
- [13] Y. Wu, D. Dobermann, M.H. Beale, J.L. Ward, Acutifoliside, a novel benzoic acid glucoside from *Salix acutifolia*, *Nat. Prod. Res.* 30 (2015) 1731–1739.
- [14] R. Nakabayashi, K. Saito, Ultrahigh resolution metabolomics for S-containing metabolites, *Curr. Opin. Biotechnol.* 43 (2017) 8–16.
- [15] I.N. Abreu, M. Ahnlund, T. Moritz, B.R. Albrechtsen, UHPLC-ESI/TOFMS Determination of salicylate-like phenolic glycosides in *Populus tremula* leaves, *J. Chem. Ecol.* 37 (2011) 857–870.
- [16] C.-J. Chou, L.-C. Lin, W.-J. Tsai, S.-Y. Hsu, L.-K. Ho, Phenyl β-d-glucopyranoside derivatives from the fruits of *Idesia polycarpa*, *J. Nat. Prod.* 60 (1997) 375–377.
- [17] Y. Zhou, Y.-H. Shen, C. Zhang, J. Su, R.-H. Liu, W.-D. Zhang, Triterpene saponins from *Bacopa monnieri* and their antidepressant effects in two mice models, *J. Nat. Prod.* 70 (2007) 652–655.
- [18] L. Varin, Flavonoid sulfation: phytochemistry, enzymology and molecular biology, in: H.A. Stafford, R.K. Ibrahim (Eds.), *Phenolic Metabolism in Plants*, Plenum Press, New York, 1992.
- [19] L. Krenn, A. Presser, R. Pradhan, B. Bahr, D.H. Paper, K.K. Mayer, B. Kopp, Sulfemodin 8-O-β-D-glucoside, a new sulfated anthraquinone glycoside, and anti-oxidant phenolic compounds from *Rheum emodi*, *J. Nat. Prod.* 66 (2003) 1107–1109.
- [20] R.A. Sessa, M.H. Bennett, M.J. Lewis, J.W. Mansfield, M.H. Beale, Metabolite profiling of sesquiterpene lactones from *Lactuca* species, *J. Biol. Chem.* 275 (2000) 26877–26884.
- [21] X.-N. Zhong, T. Ide, H. Otsuka, E. Hirata, Y. Takeda, (+)-Isolarisiresinol 3a-O-sulphate from leaves of *Myrsine seguinii*, *Phytochemistry* 40 (1998) 1777–1778.
- [22] H.H. Kim, M.H. Oh, K.J. Park, J.H. Heo, M.W. Lee, Anti-inflammatory activity of sulfate-containing phenolic compounds isolated from leaves of *Myrica rubra*, *Fitoterapia* 92 (2014) 188–193.