PERGAMON

Phytochemistry 55 (2000) 419–428

PHYTOCHEMISTRY

www.elsevier.com/locate/phytochem

Structural studies of triterpenoid saponins with new acyl components from *Quillaja saponaria* Molina

Shengjun Guo, Lennart Kenne *

Department of Chemistry, Swedish University of Agricultural Sciences, PO Box 7015, SE-750 07 Uppsala, Sweden

Received 24 May 2000; received in revised form 3 August 2000

Abstract

Eight new triterpenoid saponins were isolated from a bark extract of *Quillaja saponaria* Molina by silica and reverse phase chromatography. The saponins were characterized by spectroscopic data and chemical methods as phytolaccagenic acid, 22β-hydroxy-quillaic acid, and echinocystic acid substituted with different oligosaccharides at C-3 and C-28. The O-4 of the fucosyl residue in the 28-*O*-oligosaccharide was substituted with either acetyl, (*S*)-2-methylbutanoyl, or (3*S*,4*S*)-3-hydroxy-4-methylhexanoyl groups.

Keywords: Structural analysis; Triterpene; *Quillaja* saponins; Phytolaccagenic acid; 22β-Hydroxyquillaic acid; Echinocystic acid; Acetyl group; (*S*)-2-Methylbutanoyl group; (3*S*,4*S*)-3-Hydroxy-4-methylhexanoyl group

1. Introduction

Quillaja saponins obtained from bark extracts of the South American tree *Quillaja saponaria* Molina (Rosaceae) are known to be a physiologically active triterpenoid saponin mixture. It shows strong adjuvant activity, immuno-enhancing effects and it is widely applied in animal and human vaccines (Dalsgaard, 1974; Wu et al., 1992; Kensil et al., 1996). The *Quillaja* saponins consist of a large number of components and several of the major saponins have been purified and identified. All these identified structures have in common the triterpene quillaic acid with either a di- or one of two trisaccharides attached to C-3 and several oligosaccharides at C-28. The latter oligosaccharide is substituted with two C_9 aliphatic acyl groups (Higuchi et al., 1988; Kensil et al., 1995; Jacobsen et al., 1996; Nord and Kenne, 1999; Nyberg et al., 2000) or an *O*-acetyl group (van Setten et al., 1998; Guo et al., 2000a; Guo and Kenne, 2000). Using HPLC in combination with multiple stage mass spectrometry on an ion-trap instrument, a large number of quillaic acid saponins was found but also the occurrence of some non-quillaic acid triterpenes was indicated as the molecular masses of the triterpene part differed from those of quillaic acid (van Setten et al., 1998).

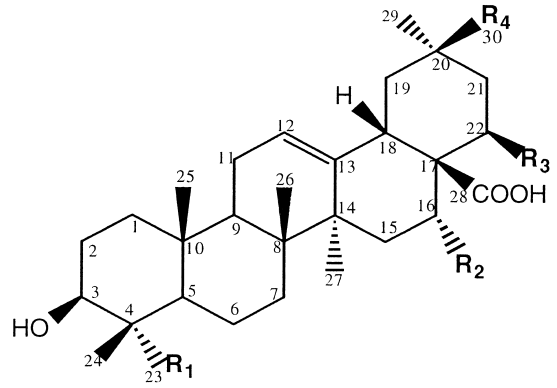
This paper describes the isolation and structure determination of eight novel saponins (**19–23**) isolated from a more hydrophilic fraction of a bark extract of *Q. saponaria* Molina from which 26 components were identified previously (Guo et al., 1998, 2000a; Guo and Kenne, 2000). These saponins consist of different non-quillaic acid triterpenes glycosated at both C-3 and C-28 positions (Fig. 1). The oligosaccharide at C-3 consists either of a di- or a trisaccharide with the same structures as those previously identified in *Quillaja* saponins. A tri-, tetra- or pentasaccharide with different acyl groups at O-4 of the fucosyl residue is linked to C-28 (Fig. 2). Two new acyl groups were found and these were identified as (*S*)-2-methylbutanoyl for compounds **19–21** and (3*S*,4*S*)-3-hydroxy-4-methylhexanoyl for compound **22**. Compound **23** contains an acetyl group at O-4 of the fucosyl residue.

2. Results and discussion

A commercial bark extract was first separated by SPE on a C-18 column, which was eluated with a stepwise gradient of aq. 10–80% MeOH. The saponin mixture eluted from 60 and 70% MeOH was collected and further

* Corresponding author. Tel.: +46-18-671-573; fax: +46-18-673-477.
E-mail address: lennart.kenne@kemi.slu.se (L. Kenne).

0031-9422/00/$ - see front matter
PII: S0031-9422(00)00340-X



	R_1	R_2	R_3	R_4
19	CH_2OH	H	H	$COOCH_3$
20a,b-22a,b	CHO	OH	OH	CH_3
23	CH_3	OH	H	CH_3

Fig. 1. Triterpene moiety of **19–23**.

fractionated by column chromatography on silica gel and the eluate monitored by TLC, MALDI-TOF MS and NMR spectroscopy. Most of the major components in this eluate were identified as quillaic acid saponins in previous studies and these were designated **1–18** (Guo et al., 1998, 2000a; Guo and Kenne, 2000). In addition to these major saponins, several new components were detected and MALDI-TOF MS and NMR indicated saponins with triterpenes different from quillaic acid. One such component was relatively fast moving with an $[M+Na]^+$ ion of m/z 1415.2, and seven more minor components were detected and MALDI-TOF MS showed $[M+Na]^+$ ions of m/z 1665.6, 1679.6, 1797.7, 1811.7, 1988.7, 2002.8 and 1755.4, respectively. The spectra showed the $[M+Na]^+$ ion as the dominating ion but also $[M+H]^+$, $[M+K]^+$ and $[M-H+2Na]^+$ ions were observed. All these components were further purified by reverse phase HPLC using an ammonium acetate buffer of pH 6.8, a system which separates the saponins due to the different structures of the oligosaccharide at C-28 (Nord and Kenne, 1999; Nyberg et al., 2000; Guo et al., 2000a). These chromatographic conditions gave a pure compound **19** from the first relatively fast moving fraction and separated from other fractions three pairs of components, **20a,b–22a,b**. MALDI-TOF mass spectra showed a molecular mass difference of 14 Da for each pair. The ^1H NMR spectra of the pairs showed that the saponins had two different trisaccharides at C-3 with either a terminal rhamnose or a terminal xylose group, as the NMR data (Table 2) of the C-3 substituent were similar to those previously found (Guo et al., 1998, 2000a; Guo and Kenne, 2000) and as the signals from the two anomeric protons together integrated to one proton.

By a second reverse phase separation step using a phosphate buffer of pH 2.8 on the same HPLC system, minor component **23** was separated from a fraction containing the pair of components **14a** and **14b**, previously studied (Guo and Kenne, 2000).

Analysis of the monosaccharides, released by acidic hydrolysis, as alditol acetates by GC (Sawardeker et al., 1965) showed the presence of galactose, glucose, xylose, apiose, fucose and rhamnose in **19–23** and the relative proportions of the different sugar residues are given in Table 1. The absolute configurations of the monosaccharides were assumed to be the same as those found in previously identified *Quillaja* saponins (van Setten and van de Werken, 1996; Guo et al., 1998, 2000a), thus giving the L-configuration for rhamnose and D-configuration for the other monosaccharides. This assumption was supported by the similar ^1H and ^{13}C chemical shifts for signals of the same disaccharide elements as those for previously identified *Quillaja* saponins.

NMR spectroscopy provides structural information on the triterpene moiety and the sugar residues including anomeric configuration, sequence and acyl groups and is thus the main tool used for the structure elucidation in this study. All NMR signals from the different sugar residues and the aglycone moiety in **19–23** were assigned by different 2D experiments and the chemical shifts are given in Tables 2 and 3. First all proton spin-systems were determined using different $^1H,^1H$ COSY, TOCSY and NOESY experiments and then the ^{13}C signals could be assigned by the one- and three-bond heteronuclear connectivities observed in the HSQC and HMBC spectra. The cross-peaks observed in NOESY and HMBC spectra could be used to assign the sequence of sugar residues and the overlapped signals. From comparison of the chemical shifts and the pattern of the cross-peaks with those of corresponding monosaccharides

Table 1
Molecular masses and neutral monosaccharide residues[a] in saponins **19–23** of *Quillaja saponaria* Molina

Compound	M^b_W	M^c_W	D-Gal	D-Glc	D-Xyl	D-Fuc	L-Rha	D-Api
19	1392.2	1392.7	1	1	–	1	1	–
20a	1656.6	1656.7	1	1	1	1	2	–
20b	1642.6	1642.7	1	1	2	1	1	–
21a	1788.7	1788.8	1	1	1	1	2	1
21b	1774.7	1774.8	1	1	2	1	1	1
22a	1979.8	1978.8	1	1	1	1	3	1
22b	1965.7	1964.8	1	1	2	1	2	1
23	1732.4	1732.7	1	2	2	1	1	–

[a] The number of neutral monosaccharide residues was estimated from the results of the monosaccharide analyses and the integrals of the signals for anomeric protons in the ^1H NMR spectra.

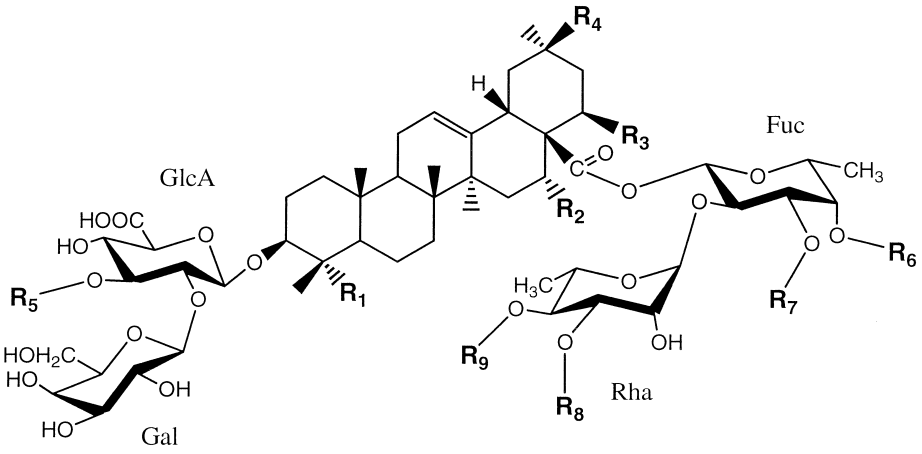
[b] Monoisotopic molecular mass determined by MALDI-TOF mass spectrometry.

[c] Monoisotopic molecular mass of the assigned structures.

and methyl glycosides (Jansson et al., 1989; Agrawal, 1992) and previously identified *Quillaja* saponins (Jacobsen et al., 1996; Guo et al., 1998; Nord and Kenne, 1999; Nyberg et al., 2000; Guo et al., 2000a) each sugar and its anomeric configuration could be identified. Observed intra-residue NOEs supported the anomeric configuration whereas the inter-residue NOEs established the glycosidic linkages. Information on the glycosidic linkages was also obtained from the three-bond heteronuclear correlation over the glycosidic bond observed as cross-peaks in the HMBC spectra.

2.1. Saponin **19**

The NMR spectra of compound **19** showed that it is not a quillaic acid saponin as there are no signals for an aldehyde group at C-23. Instead signals for two carbonyl carbons (δ 177.6 and 178.9) are present in the ^{13}C NMR spectrum. The ^{1}H NMR spectrum showed resonances for only five methyl groups on tertiary carbons at δ 0.73, 0.81, 0.98, 1.14 and 1.19 (s, 3 H each). A three proton singlet at δ 3.73 with a one-bond heteronuclear correlation to the carbon resonance at δ 52.3 and a



	R_5	R_6	R_7	R_8	R_9
19	H	AcylI	Glc	H	H
20a,b	Rha/Xyl	AcylI	H	GlcI	Xyl
21a,b	Rha/Xyl	AcylI	H	GlcI	Api-Xyl
22a,b	Rha/Xyl	AcylII	RhaI	GlcI	Api-Xyl
23	Xyl	Ac	Glc	GlcI	Xyl

Rha = α-L-Rha*p*
Xyl = β-D-Xyl*p*
Glc = β-D-Glc*p*
Api-Xyl = β-D-Api*f*-(1→3)-β-D-Xyl*p*

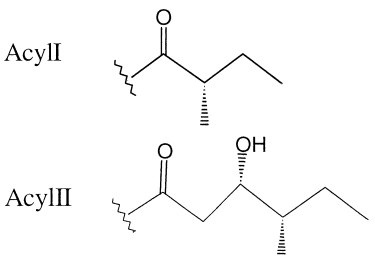


Fig. 2. Structures of saponins **19–23**.

Table 2
The ^{1}H and ^{13}C NMR chemical shifts (ppm) for the triterpene moiety and the oligosaccharide in its 3-position of saponins **19–23**

	19		**20a**		**21b**		**22b**		**23**	
Atoms No	^{1}H	^{13}C	^{1}H	^{13}C	^{1}H	^{13}C	^{1}H	^{13}C	^{1}H	^{13}C
1	0.99, 1.64	39.4	1.11, 1.71	39.2	1.11, 1.71	39.0	1.12, 1.73	39.4	1.00, 1.63	40.0
2	1.93, 1.77	26.0	2.05, 1.79	25.2	2.02, 1.80	25.4	2.04, 1.79	25.5	1.94, 1.73	26.5
3	3.64	83.6	3.87	85.7	3.88	85.4	3.89	85.4	3.19	91.8
4		44.2		56.1		56.2		56.8		40.5
5	1.23	48.2	1.31	49.1	1.33	49.1	1.33	49.1	0.77	57.2
6	1.36, 1.51	18.7	0.96, 1.48	21.3	0.97, 1.48	21.2	0.97, 1.48	21.4	1.39, 1.57	19.3
7	1.61, 1.36	33.3	1.54, 1.29	33.6	1.56, 1.41	33.5	1.55, 1.34	33.7	1.60, 1.37	33.5
8		40.7		40.9		41.2		40.9		40.9
9	1.64	48.9	1.75	47.8	1.76	48.0	1.74	48.0	1.63	48.2
10		37.6		36.9		37.1		37.1		36.5
11	1.93, 1.92	24.5	1.94, 1.94	24.4	1.93, 1.94	24.2	1.94, 1.94	24.6	1.91, 1.91	24.5
12	5.34	124.0	5.36	122.9	5.36	122.9	5.38	123.1	5.33	122.9
13		144.5		144.8		144.8		144.7		144.8
14		43.2		42.8		42.8		42.9		42.8
15	1.65, 1.21	28.9	1.65, 1.47	36.3	1.64, 1.47	36.2	1.63, 1.48	36.3	1.74, 1.43	36.4
16	2.08, 1.71	24.2	4.56	71.5	4.55	71.5	4.54	71.1	4.43	74.2
17		47.4		49.7		50.1		49.6		49.9
18	2.74	43.4	3.09	40.0	3.11	39.8	3.09	40.0	2.92	42.4
19	1.95, 1.72	43.2	2.36, 1.08	47.8	2.37, 1.08	47.9	2.39, 1.09	48.0	2.27, 1.05	48.2
20		44.9		31.0		30.8		31.0		31.2
21	2.03, 1.40	31.2	2.03, 1.57	43.8	2.04, 1.56	44.0	2.04, 1.56	43.7	1.90, 1.19	36.3
22	1.67, 1.67	34.1	4.25	73.3	4.25	73.6	4.26	73.8	1.90, 1.72	31.7
23	3.27, 3.75	64.9	9.46	211.5	9.46	211.4	9.46	211.0	1.09	28.3
24	0.73	13.2	1.18	10.7	1.17	10.7	1.18	10.6	0.87	16.9
25	0.98	24.8	1.01	16.2	1.03	16.2	1.02	16.4	0.96	16.2
26	0.81	17.8	0.80	17.7	0.80	17.9	0.81	17.5	0.84	17.9
27	1.19	26.0	1.41	27.0	1.41	27.0	1.42	26.9	1.38	27.1
28		177.6		174.8		–[a]		174.5		–[a]
29	1.14	28.4	0.88	33.5	0.87	33.1	0.88	33.6	0.88	33.2
30		178.9	1.11	28.0	1.10	27.9	1.10	27.9	0.98	25.2
OMe	3.73	52.3								
GlcA1	4.57	104.4	4.37	104.0	4.37	103.9	4.41	104.0	4.52	105.5
GlcA2	3.55	82.4	3.63	78.2	3.65	78.4	3.64	78.3	3.79	79.3
GlcA3	3.64	78.1	3.64	85.9	3.68	86.4	3.63	86.1	3.73	86.5
GlcA4	3.57	73.5	3.52	72.2	3.53	71.7	3.55	71.9	3.60	71.7
GlcA5	3.72	76.8	3.59	77.5	3.59	77.5	3.69	76.8	3.69	78.7
GlcA6		–[a]		172.6		172.7		174.3		–[a]
Gal1	4.60	105.4	4.47	103.9	4.80	103.4	4.80	103.5	4.85	104.2
Gal2	3.59	73.5	3.46	73.4	3.47	73.2	3.50	73.2	3.52	73.4
Gal3	3.49	74.5	3.49	74.6	3.46	75.1	3.46	75.1	3.46	74.9
Gal4	3.85	69.8	3.83	70.5	3.84	70.5	3.84	70.6	3.83	70.0
Gal5	3.48	76.6	3.48	76.5	3.50	76.3	3.50	76.5	3.47	76.7
Gal6	3.71, 3.73	61.8	3.73,3.79	61.9	3.73,3.77	62.0	3.73,3.77	62.0	3.73,3.78	61.9
Xyl1					4.62	104.5	4.62	104.6	4.63	105.2
Xyl2					3.23	75.2	3.24	75.1	3.26	75.1
Xyl3					3.32	78.0	3.32	77.8	3.32	78.0
Xyl4					3.51	71.0	3.51	71.0	3.52	71.1
Xyl5					3.23,3.90	66.9	3.23,3.90	66.7	3.24,3.91	67.1
Rha1			5.06	103.0						
Rha2			4.01	71.8						
Rha3			3.67	72.1						
Rha4			3.38	73.7						
Rha5			4.01	70.2						
Rha6			1.24	17.4						

[a] Signal not detected.

three-bond correlation to the carbonyl signal at δ 178.9 observed in the HSQC and HMBC spectrum, respectively, indicated the presence of a methyl ester in the aglycone. Additionally, an oxymethylene group (δ 3.75 and 3.23) with the corresponding carbon resonance at δ 64.9 demonstrated that the triterpene in **19** contains a primary alcohol group and the connectivities observed in the HMBC spectrum demonstrated the primary alcohol to be at C-23. Due to the effects of the primary alcohol located at C-23, the C-3 (δ 83.6) and C-4 (δ 44.2) signals were shifted upfield ~2 and ~12 ppm, respectively, compared to those of quillaic acid saponins (Jacobsen et al., 1996; Guo et al., 1998, 2000a; Nord and Kenne, 1999; Nyberg et al., 2000). The identified structural elements and the ^{1}H and ^{13}C chemical shifts of the triterpene in **19** were similar to those of a previously found triterpene (Bandara et al., 1990; Rastrelli et al., 1996) and could thus be identified as 3β,23-dihydroxy-30-methylcarbonylolean-12-en-28-oic acid, phytolaccagenic acid. This triterpene was previously found

Table 3
The ^{1}H and ^{13}C NMR chemical shifts (ppm) for the oligosaccharide in the 28-position of saponins **19–23**

	19		**20**		**21**		**22**		**23**	
Atoms No	^{1}H	^{13}C	^{1}H	^{13}C	^{1}H	^{13}C	^{1}H	^{13}C	^{1}H	^{13}C
Fuc1	5.43	95.1	5.44	95.1	5.47	95.0	5.60	94.6	5.41	95.3
Fuc2	3.88	74.7	3.77	74.9	3.77	74.9	3.95	75.2	3.91	74.5
Fuc3	4.01	82.7	3.91	73.9	3.90	74.0	4.01	82.7	4.03	82.4
Fuc4	5.41	74.6	5.09	74.6	5.09	74.7	5.23	74.4	5.36	74.9
Fuc5	3.87	70.7	3.90	71.2	3.91	71.2	3.97	71.8	3.86	70.9
Fuc6	1.05	16.5	1.08	16.3	1.07	16.2	1.09	16.8	1.06	16.3
Rha1	5.28	102.0	5.27	101.5	5.29	101.5	5.12	101.5	5.34	101.3
Rha2	3.96	71.7	4.24	70.9	4.23	71.1	4.12	71.5	4.24	71.0
Rha3	3.65	71.9	3.93	82.8	3.90	82.7	3.84	82.8	3.91	82.8
Rha4	3.38	73.4	3.68	78.7	3.67	78.9	3.66	79.5	3.69	78.7
Rha5	3.72	70.2	3.86	68.9	3.84	68.7	3.80	69.1	3.87	68.9
Rha6	1.24	18.2	1.30	18.4	1.30	18.3	1.32	19.1	1.29	18.1
RhaI1							4.92	104.8		
RhaI2							3.88	71.8		
RhaI3							3.60	71.9		
RhaI4							3.36	73.5		
RhaI5							3.63	70.4		
RhaI6							1.23	17.5		
Xyl1			4.69	104.8	4.69	104.9	4.68	105.1	4.68	105.2
Xyl2			3.10	75.4	3.21	75.0	3.19	75.1	3.14	75.4
Xyl3			3.28	78.3	3.36	85.7	3.35	85.9	3.29	78.2
Xyl4			3.47	71.2	3.49	70.4	3.48	70.0	3.48	71.3
Xyl5			3.17, 3.85	66.8	3.19, 3.87	66.8	3.18, 3.85	66.8	3.17, 3.85	66.8
Api1					5.29	111.0	5.29	111.0		
Api2					4.03	77.7	4.05	77.7		
Api3					–	80.3	–	80.3		
Api4					3.81, 4.16	74.8	3.82, 4.17	74.8		
Api5					3.66	65.1	3.67	65.2		
Glc1	4.50	105.6							4.49	105.6
Glc2	3.14	75.1							3.15	75.2
Glc3	3.33	78.0							3.36	78.0
Glc4	3.19	71.1							3.23	71.2
Glc5	3.28	77.7							3.29	78.2
Glc6	3.61, 3.86	62.7							3.62, 3.85	62.8
GlcI1			4.52	105.0	4.51	105.1	4.54	105.0	4.58	104.8
GlcI2			3.30	75.0	3.31	75.0	3.31	75.1	3.31	75.1
GlcI3			3.31	77.9	3.33	78.0	3.34	77.8	3.38	77.8
GlcI4			3.34	70.8	3.37	70.7	3.36	70.9	3.36	71.0
GlcI5			3.27	78.3	3.26	77.7	3.27	78.0	3.32	78.1
GlcI6			3.70, 3.82	62.3	3.68, 3.83	62.0	3.71, 3.86	61.9	3.71, 3.86	62.4
Acyl-1		178.6		178.5		178.0		173.8		172.0
Acyl-2	2.53	42.5	2.51	42.3	2.51	42.4	2.47, 2.64	40.0	2.09	20.9
Acyl-3	1.75, 1.56	27.8	1.74, 1.53	27.7	1.74, 1.53	27.6	3.92	73.2		
Acyl-4	1.01	11.8	0.98	11.7	0.98	11.8	1.53	41.8		
Acyl-5	1.20	17.4	1.20	17.2	1.20	17.2	1.19, 1.58	25.6		
Acyl-6							0.93	11.7		
Acyl-7							0.96	14.9		

in saponins from *Phytolacca americana* (Kang and Woo, 1987) and other plants (Bandara et al., 1990; Rastrelli et al., 1996) but is new in the *Quillaja* saponins. However, it was recently found in a more lipophilic fraction of the bark extract from *Quillaja saponaria* (Nord and Kenne, 2000).

The 1D and 2D NMR spectra of **19** showed inter alia two spin-systems including the resonances for anomeric protons at δ 4.57 (β-D-Glc*p*A) and 4.60 (β-D-Gal*p*). These systems could be attributed to the disaccharide β-D-Gal*p*-(1→2)-β-D-Glc*p*A attached at C-3 of the triterpene since NMR data are similar to those of corresponding disaccharide in previously identified saponins (Guo et al., 1998, 2000a; Nord and Kenne, 1999). Some deviations from the signals of GlcA H-1/C-1 and Gal H-1/C-1 were observed, which could be due to the primary alcohol at C-23 in **19** and an aldehyde group in quillaic acid. The connectivities between H-1 of GlcA and C-3 or H-3 of the triterpene (δ 4.57/83.6 and δ 4.57/3.64), and between H-1 of Gal and C-2 or H-2 of GlcA (δ 4.60/82.4 and δ 4.60/3.55), observed in HMBC and NOESY spectra, confirmed the structure of the disaccharide and also that it is substituted to C-3 of the triterpene.

The oligosaccharide at C-28 is a trisaccharide since in addition to the signals from the disaccharide at C-3, three signals from anomeric protons at δ 5.43 (β-D-Fuc*p*), 5.28 (α-L-Rha*p*) and 4.50 (β-D-Glc*p*) were observed in the ^{1}H NMR spectrum. The complete assignment of signals from this part of **19** (Table 3) was achieved by different 2D NMR experiments and by comparison with data of previously identified *Quillaja* saponins. The connectivity between H-1 of the Fuc residue and the carbonyl carbon, C-28, of the aglycone (δ 5.43/177.6) observed by HMBC confirmed that the Fuc residue is linked to C-28 of the aglycone. From connectivities found by HMBC and NOESY experiments (Table 4), the Fuc residue is substituted at C-2 with a terminal α-L-Rha and at C-3 with the terminal β-D-Glc.

The ^{13}C NMR spectrum showed a total of 66 signals, of which 31 were assigned to the triterpenoid moiety and 35 to the saccharide part. As the five sugar residues in **19** only account for 30 carbon signals, it suggests the presence of an additional 5 carbon element in **19**. According to COSY and HSQC spectra, a CH multiplet at δ 2.53 showed connectivities to a methyl group at δ 1.20/17.4 (3H, d) and two methylene protons at δ 1.56 and 1.75 and the latter two protons to another methyl group at δ 1.01/11.8 (3H, t). The ^{1}H and ^{13}C NMR data (Table 3) identified this element as a 2-methylbutanoyl group. The signal from H-4 of the Fuc residue at δ 5.41 has a high chemical shift and a three-bond heteronuclear connectivity to the carbonyl carbon of the 2-methylbutanoyl group (δ 5.41/178.6), observed in the HMBC spectrum, demonstrating that the 2-methylbutanoic acid is ester linked to O-4 of Fuc. Thus the structure of the oligosaccharide at C-28 is β-D-Glc*p*-(1→3)-[α-L-Rha*p*-(1→2)]-4-*O*-acyl-β-D-Fuc*p*. The (*S*)-configuration of the 2-methylbutanoic acid was determined by GC–MS on a capillary column of fused-silica coated with derivatised β-cyclodextrin (Guo et al., 2000b).

2.2. Saponin *20a,b*

All ^{1}H and ^{13}C chemical shifts for compound **20a,b** (Tables 2 and 3) were assigned by 1D- and different 2D-experiments. Resonances for the triterpene moiety of **20a,b** observed in the ^{1}H NMR spectrum suggested the presence of six methyl groups on tertiary carbons (3H, s, at δ 0.80, 0.88, 1.01, 1.11, 1.18 and 1.41), an olefinic group (δ 5.36 /122.9) and an aldehyde function (δ 9.46). All ^{1}H and ^{13}C chemical shifts from the atoms in the A–D rings of the triterpene moiety were very similar to those of quillaic acid, indicating structural similarity of this part with quillaic acid. However, the high chemical shifts of the H-22 (δ 4.25, 1H) and C-22 (δ 73.3) signals indicated the presence of an additional secondary hydroxyl group. The HMBC spectrum showed two- and three-bond heteronuclear connectivities between H-22 (δ 4.25) and C-21 (δ 43.8), C-20 (δ 31.0) and C-18 (δ 41.0) and the NOESY spectrum showed NOEs between this proton and H-16 (δ 4.56) and H-21 (δ 2.03, 1.57). The aglycone in **20a,b** could be identified as 22β-hydroxyquillaic acid (3β,16α,22β-trihydroxyolean-12-en-28-oic acid).

The NMR data showed that compounds **20a** and **b** are substituted at both C-3 and C-28. The oligosaccharide at C-3 is a trisaccharide which consists either of β-D-Gal*p*-(1→2)-[α-L-Rha*p*-(1→3]-β-D-Glc*p*A (**20a**, 72%) or β-D-Gal*p*-(1→2)-[β-D-Xyl*p*-(1→3)]-β-D-Glc*p*A (**20b**, 28%). The structures of the trisaccharides were consistent with those of previously indentified saponins (Guo et al., 1998, 2000a; Nord and Kenne, 1999; Nyberg et al., 2000). Each component in such a pair could be identified by high resolution NMR spectroscopy without further purification as described previously (Guo and Kenne, 2000). The sequence of sugar residues and the linkage to C-3 were corroborated by the inter-residue three-bond heteronuclear connectivities, observed as cross-peaks in the HMBC spectrum.

The oligosaccharide at C-28 consisted of a tetrasaccharide according to the result of the monosaccharide analysis (Table 1) and the number and types of spin-systems that were observed in the NMR spectra. The β-D-Fuc residue was substituted with an α-L-Rha at C-2 and this residue was then substituted with a β-D-Glc at C-3 and a β-D-Xyl at C-4, shown by inter-residue connectivities observed in the NOESY and HMBC spectra (Table 4). The β-D-Fuc residue was also substituted at O-4 with an (*S*)-2-methylbutanoyl group, as in **19** since almost identical NMR data were obtained

for the 4-*O*-acyl substitutent of Fuc in both compounds (Table 3). Thus the C-28 oligosaccharide has the structure β-D-Glc*p*-(1→3)-[β-D-Xyl*p*-(1→4)]-α-L-Rha*p*-(1→2)-4-*O*-acyl-β-D-Fuc*p*. The 28-O-oligosaccharide in **20a,b** has the same basic structure as that of previously identified saponins **13a,b** (Guo and Kenne, 2000) and **B1** and **B2** (Nyberg et al., 2000). However, in **13a,b** the 4-*O*-acyl group on the Fuc residue is an acetyl group and in **B1** and **B2** the dimeric C_9 acyl group terminated with an arabinofuranosyl group instead of the (*S*)-2-methylbutanoyl group in **20a,b**. When the 3 position of the Fuc residue was unsubstituted the 3-*O*-acyl substituted isomers of the β-D-Fuc residue were formed by migration of the *O*-acyl substituent (Jacobsen et al., 1996; Nord and Kenne, 1999; Nyberg et al., 2000). By comparison of the 1D and 2D NMR spectra of **20a,b** with those of the 3-*O*-acyl isomers **B1a** and **B2a**, signals from a 3-*O*-acyl Fuc residue were also observed, demonstrating the occurrence of these isomers in **20a,b**. The amount of the 3-*O*-acetylated isomers was estimated to ~20% according to the relative intensity of the signals in ^{1}H NMR spectrum.

2.3. Saponin *21a,b*

The pair **21a** and **b** has the Rha form of the trisaccharide at C-3 as the major component (73%), contrary to **20a,b**, as evident from the NMR spectra.

Table 4
Observed inter-residue NOE and three-bond heteronuclear connectivities for the anomeric protons of the sugar residues in the 28-O-oligosaccharide of saponin **19–23**[a]

Compound	Residue	Anomeric ^{1}H (ppm)	Connectivities to δ_C	δ_H	Inter-residue Residue	Atom
19	α-L-Rha-(1→	5.28	74.7		→2,3)-β-D-Fuc-(1→	C-2
		5.28		*3.88*	*→2,3)-β-D-Fuc-(1→*	*H-2*
	β-D-Glc-(1→	4.50	82.7		→2,3)-β-D-Fuc-(1→	C-3
		4.50		*4.01*	*→2,3)-β-D-Fuc-(1→*	*H-3*
	→2,3)-β-D-Fuc-(1→	5.43	177.6		Aglycone	C-28
20a,b	β-D-Xyl-(1→	4.69	78.7		→3,4)-α-L-Rha-(1→	C-4
		4.69		*3.68*	*→3,4)-α-L-Rha-(1→*	*H-4*
	β-D-Glc-(1→	4.52	82.8		→3,4)-α-L-Rha-(1→	C-3
		4.52		*3.93*	*→3,4)-α-L-Rha-(1→*	*H-3*
	→3,4)-α-L-Rha-(1→	5.27	74.9		→2)-β-D-Fuc-(1→	C-2
		5.27		*3.77*	*→2)-β-D-Fuc-(1→*	*H-2*
	→2)-β-D-Fuc-(1→	5.44	174.8		Aglycone	C-28
21a,b	β-D-Api-(1→	5.29	85.7		→3)-β-D-Xyl-(1→	C-3
		5.29		*3.36*	*→3)-β-D-Xyl-(1→*	*H-3*
	→3)-β-D-Xyl-(1→	4.69	78.9		→3,4)-α-L-Rha-(1→	C-4
		4.69		*3.67*	*→3,4)-α-L-Rha-(1→*	*H-4*
	β-D-Glc-(1→	4.51	82.7		→3,4)-α-L-Rha-(1→	C-3
		4.51		*3.90*	*→3,4)-α-L-Rha-(1→*	*H-3*
	→3,4)-α-L-Rha-(1→	5.29	74.9		→2-β-D-Fuc-(1→	C-2
		5.29		*3.77*	*→2-β-D-Fuc-(1→*	*H-2*
22a,b	β-D-Api-(1→	5.29	85.9		→3)-β-D-Xyl-(1→	C-3
		5.29		*3.35*	*→3)-β-D-Xyl-(1→*	*H-3*
	→3)-β-D-Xyl-(1→	4.68	79.5		→3,4)-α-L-Rha-(1→	C-4
		4.68		*3.66*	*→3,4)-α-L-Rha-(1→*	*H-4*
	β-D-Glc-(1→	4.54	82.8		→3,4)-α-L-Rha-(1→	C-3
		4.54		*3.84*	*→3,4)-α-L-Rha-(1→*	*H-3*
	α-L-Rha-(1→	4.92	82.7		→2,3)-β-D-Fuc-(1→	C-3
		4.92		*4.01*	*→2,3)-β-D-Fuc-(1→*	*H-3*
	→3,4)-α-L-Rha-(1→	5.12	75.2		→2,3)-β-D-Fuc-(1→	C-2
		5.12		*3.95*	*→2,3)-β-D-Fuc-(1→*	*H-2*
	→2,3)-β-D-Fuc-(1→	5.60	174.5		Aglycone	C-28
23	β-D-Xyl-(1→	4.68	78.7		→3,4)-α-L-Rha-(1→	C-4
		4.68		*3.69*	*→3,4)-α-L-Rha-(1→*	*H-4*
	β-D-Glc-(1→	4.58	82.8		→3,4)-α-L-Rha-→	C-3
		4.58		3.91	*→3,4)-α-L-Rha-(1→*	*H-3*
	→3,4)-α-L-Rha-(1→	5.34	74.5		→2,3-β-D-Fuc-(1→	C-2
		5.34		*3.91*	*→2,3)-β-D-Fuc-(1→*	*H-2*
	β-D-Glc-(1→	4.49	82.4		→2,3)-β-D-Fuc-(1→	C-3
		4.49		*4.03*	*→2,3)-β-D-Fuc-(1→*	*H-3*

[a] The connectivities were observed as cross-peaks in NOESY and HMBC spectra. NOE connectivities are showed in italics.

Monosaccharide analysis of compound **21a,b** showed the same sugar residues as those found in **20a,b** but also one Api residue (Table 1). For both the triterpene moiety and the sugar residues in compound **21a,b** the ^{1}H and ^{13}C NMR spectra exhibited signals almost identical to those of **20a,b** (Table 2 and 3). However, signals for an additional β-D-Api*f* are present with the signals for anomeric atoms at δ 5.29 ppm, $J_{1,2}=2.8$ Hz and δ 111.0 ppm. According to the downfield shifts of the C-3 signal of the Xyl residue and the inter-residue connectivities (Table 4) between the anomeric proton of β-D-Api*f* group and the linkage atoms of the Xyl residue the β-D-Api*f* group is linked to C-3 of the Xyl residue. The same acyl group, namely (*S*)-2-methylbutanoic acid linked to O-4 of β-D-Fuc, was found in **21a,b** as almost identical chemical shifts, coupling pattern and cross-peaks in NOESY and HMBC spectra from this part were observed for both **20a,b** and **21a,b** (Table 3). Thus the structures of the saponin **21a** and **b**, which is 22β-hydroxyquillaic acid 3-*O*-{β-D-Gal*p*-(1→2)-[α-L-Rha*p*/β-D-Xyl*p*-(1→3)]-β-D-Glc*p*A}-28-O-{β-D-Api*f*-(1→3)-β-D-Xyl*p*-(1→4)-[β-D-Glc*p*-(1→3)]-α-L-Rha*p*-(1→2)-4-*O*-acyl-β-D-Fuc*p*}, are the same as those of **20a** and **b** but with an extra β-D-Api at C-3 of the Xyl residue (Fig. 2).

The structure of 28-*O*-oligosaccharide in **21a,b** is the same as those previously identified saponins **B3** and **B4** (Nyberg et al., 2000). The difference between them is that the 4-*O*-acyl group on the Fuc residue in **21a,b** is the (*S*)-2-methylbutanoyl group but in **B3** and **B4** the dimeric C_9 acyl group terminated with an arabinofuranosyl group. The occurrence of the 3-*O*-acyl isomer of the Fuc residue was also observed in **21a,b**, since similar chemical shifts from this isomer with those of the 3-*O*-acyl **B3a** and **B4a** were obtained. The amount of the 3-*O*-acylated isomers was estimated to ~30% according to the relative intensity of the ^{1}H NMR signals.

2.4. Saponin *22a,b*

Compound **22a,b** has the highest molecular mass of the saponins isolated from the 60% MeOH eluate. The ^{1}H NMR spectrum showed similarity with that of **21a,b**, especially for the signals from the triterpene moiety. The resonances for the six tertiary methyl groups, two hydroxyl groups, one olefinic and aldehyde group and the connectivities observed in NOESY and HMBC spectra established that the aglycone in **22a,b** is 22β-hydroxy-quillaic acid, the same as that in **21a,b** and **20a,b** (Fig. 1 and Table 2).

According to monosaccharide analysis, MALDI-TOF MS and NMR data, **22a,b** consists of a pair of components substituted with the two different trisaccharides at C-3 with the Xyl form dominating (75%) over the Rha form. Assignment of signals for the 28-*O*-oligosaccharide was achieved by comparison of NMR data of **22a,b** with those of **21a,b** and previously identified saponins. The results showed that the Fuc, linked to C-28 of the aglycone, was substituted with a terminal Rha at C-3 and a 3,4-disubstituted Rha at C-2, and the latter Rha was substituted with a terminal Glc at C-3 and a β-D-Api*f*-(1→3)-β-D-Xyl*p* at C-4. This is the same pattern as in **21a,b**, except that an additional Rha is linked to C-3 of the Fuc residue. The linkages were supported by the inter-residue connectivities observed in HMBC and NOESY spectra (Table 4).

The relative high chemical shift at C-4 of Fuc (δ 5.23/74.4) and a three-bond connectivity between H-4 and the carbonyl carbon at δ 173.8 demonstrated that an acyl substituent was linked to O-4 of Fuc. A COSY spectrum showed that the resonances observed for methylene protons at δ 2.64, 2.47/40.0 (Acyl-2) is linked to a secondary hydroxyl group at δ 3.79/73.2 and both signals were correlated to the carbonyl carbon (δ 173.8) in the HMBC spectrum. The hydroxyl group is then connected to a methine proton at δ 1.53/41.8 (Acyl-4) and the latter is in turn linked to a methylene group at δ 1.58, 1.19 /25.6 (Acyl-5) and a methyl group at δ 0.96/14.9 (doublet, Acyl-7). Acyl-5 is substituted by another methyl group at δ 0.93/11.7 (triplet, Acyl-6). Thus the acyl substituent is a 3-hydroxy-4-methylhexanoyl group (Table 3). The (3*S*,4*S*)-configuration of this acyl group has been determined recently (Guo et al., 2000b).

2.5. Saponin *23*

In a fraction from the first HPLC separation step compound **23** was found as a minor component together with two quillaic acid saponins **14a** and **b**, studied in our previous work (Guo and Kenne, 2000). However, compound **23** could be separated by a second HPLC step using a pH 2.8 buffer. The ^{1}H NMR spectrum of **23** showed seven singlets from methyl groups at δ 0.84, 0.87, 0.88, 0.96, 0.98, 1.09, 1.38 and no aldehyde signal, suggesting that **23** has an additional methyl group instead of an aldehyde group found in quillaic acid. The NMR data from ring B–E of triterpene was in good agreement with those of the quillaic acid. The cross-peaks observed in HMBC and NOESY spectra demonstrated that C-23 is a methyl group. Due to the effects of the C-23 methyl group, the C-3 signal (δ 91.8) was shifted ~6 ppm downfield and the C-4 signal (δ 40.5) 15.5 ppm upfield, compared to those of quillaic acid saponins. Thus the aglycone of **23** was identified as 3β,16α-dihydroxyolean-12-en-28-oic acid (echinocystic acid), a component of a previously identified saponin (Shao et al., 1996).

NMR data showed that the trisaccharide at C-3 of the aglycone is the xylose form, β-D-Gal*p*-(1→2)-[β-D-Xyl*p*-(1→3)]-β-D-Glc*p*A. As identical NMR data for the 28-*O*-oligosaccharide with those of **14a** and **14b** (Guo and Kenne, 2000) were obtained, the structure of the oligosaccharide at C-28 was identified as β-D-Xyl*p*-(1→4)-[β-

D-Glc*p*-(1→3)]-α-L-Rha*p*-(1→2)-[β-D-Glc*p*-(1→3)]-4-*O*-acetyl-β-D-Fuc*p*.

The structures of the saponins **19–23** (Figs. 1 and 2) are different from those of previously identified saponins isolated from bark extracts of *Quillaja saponaria* Molina (Higuchi et al., 1987, 1988; Jacobsen et al., 1996; Guo et al., 1998; Nord and Kenne, 1999; Nyberg et al., 2000; Guo et al., 2000a; Guo and Kenne, 2000). In this study we have completely characterised the structures of **19–23** showing that they possess different triterpenes instead of quillaic acid and that the fucosyl residue, attached to C-28 of the aglycone, is substituted with either acetyl, (*S*)-2-methylbutanoyl or (3*S*,4*S*)-3-hydroxy-4-methylhexanoyl groups. The absolute configuration of these acyl groups has recently been determined by chemical methods, NMR and X-ray crystallography (Guo et al., 2000b).

3. Experimental

3.1. Materials

The *Quillaja* saponin bark extract was obtained from Berghausen (Cincinnati, OH, USA). This material (portions of 200 mg) was dissolved in 1 ml aq. 10% MeOH and the solution applied to an Isolute SPE column [C-18 (EC), 10 g] for a coarse separation. The column was eluated with a stepwise gradient of aq. 10–80% MeOH (10 ml each step). The eluates from 60% and 70% MeOH were collected and used for further separation.

3.2. Isolation of compounds *19–23*

A sample obtained from the coarse separation was fractionated on a column (5×45 cm) of silica gel 60 (0.04–0.063 mm, Merck) using a mixture of $CHCl_3$, MeOH, H_2O and HOAc (24:17.5:3:0.1) as solvent. The elution of saponins was monitored by TLC, MALDI-TOF MS and ^1H NMR spectroscopy. Some fractions contained components with Rf-values between 0.3–0.5 which showed data that indicated non-quillaic acid saponins. One fraction contained one major component with molecular mass 1392.2 Da and three other fractions contained saponins of molecular masses between 1640 and 1980 Da. The four fractions were concentrated to dryness, yielding 38, 200, 106, and 85 mg, respectively. All four fractions were further separated on a semipreparative HPLC column (Kromasil 100-5 C18, 2×15 cm). Saponin components, detected by the absorbance at 205 nm, were eluted isocratically in a mixture of MeCN and aq 0.02 M ammonium acetate buffer in 20% MeCN, pH 6.8 (8.5:91.5) at a flow rate of 9 ml min^{-1}. The fractions collected were first evaporated to remove the MeCN, then diluted with H_2O and loaded on Isolute SPE columns [C-18, (EC) 10 g]. The columns were washed with H_2O to remove salts and the saponins eluted with MeOH (~20 ml) and individually evaporated to dryness. By this procedure, **19** was isolated as a pure compound and **20a,b–22a,b** were obtained as pairs of compounds with 14 Da difference in molecular mass.

One fraction collected from the fourth fraction was further separated on second HPLC separation with the same system but using a mixture of MeCN and aq. 0.01 M phosphate buffer, pH 2.8 (30:70) as mobile phase at the flow rate 8.5 ml min^{-1}. The pure saponin **23**, in addition to the pair of components (see previously determined saponins **14a** and **14b**, Guo and Kenne, 2000), was obtained and isolated by SPE in the same way as described for the first separation.

Monosaccharide analysis was performed as previously described (Guo et al., 1998, 2000a).

3.3. Mass spectrometry

The MALDI-TOF mass spectra were recorded on Bruker Reflex III spectrometer using a 337 nm nitrogen laser and 2,5-dihydroxybenzoic acid as matrix.

3.4. NMR spectroscopy

NMR spectra were recorded for samples in CD_3OD on a Bruker DRX-600 spectrometer with a proton frequency of 600 MHz equipped with a 5-mm triple-resonance inverse probe or a 2.5-mm microprobe. All spectra were acquired at 30° without spinning the sample. Chemical shifts are reported in ppm using the solvent peak as a reference (δ_H 3.31 and δ_C 49.0). DQF-COSY, TOCSY, NOESY (mixing time 300 ms), HSQC-DEPT and HMBC (delay times of 50 or 70 ms) experiments were performed according to standard pulse sequences.

Acknowledgements

This work was supported by grants from the Swedish Natural Science Research Council and the Swedish Council for Forestry and Agricultural Research.

References

Agrawal, P.K., 1992. NMR spectroscopy in the structural elucidation of oligosaccharides and glycosides. Phytochemistry 31, 3307–3330.

Bandara, B.M.R., Jayasinghe, U.L.B., Karunaratne, V., Wannigama, G.P., Bokel, M., Sotheeswaran, S., 1990. Triterpenoidal constituents of *Diploclisia glaucescens*. Planta Medica 56, 290–292.

Dalsgaard, K., 1974. Saponin adjuvants. III. Isolation of a substance from *Quillaja saponaria* Molina with adjuvant activity in foot-and-mouth vaccines. Arch Gesamte Virusforsch 44, 243–254.

Guo, S., Kenne, L., 2000. Characterization of some *O*-acetylated saponins from *Quillaja saponaria*. Phytochemistry 54, 615–623.

Guo, S., Kenne, L., Lundgren, L.N., Rönnberg, B., Sundquist, B.G., 1998. Triterpenoid saponins from *Quillaja saponaria*. Phytochemistry 48, 175–180.

Guo, S., Falk, E., Kenne, L., Rönnberg, B., Sundquist, B.G., 2000a. Triterpenoid saponins containing an acetylated branched D-fucosyl residue from *Quillaja saponaria* Molina. Phytochemistry 53, 861–868.

Guo, S., Kristiansson, O., Nord, L., Kenne, L., 2000b. Identification of (*S*)-2-methylbutanoic acid, (3*S*,4*S*)-3-hydroxy-4-methylhexanoic acid and (3*S*,5*S*,6*S*)-3,5-dihydroxy-6-methyloctanoic acid as components of saponins from *Quillaja saponaria* Molina. Tetrahedron Asymmetry, in manuscript.

Higuchi, R., Tokimitsu, Y., Fujioka, T., Komori, T., Kawasaki, T., Oakenful, D.G., 1987. Structure of desacylsaponins obtained from the bark of *Quillaja saponaria*. Phytochemistry 26, 229–235.

Higuchi, R., Tokimitsu, Y., Komori, T., 1988. An acylated triterpenoid saponin from *Quillaja saponaria*. Phytochemistry 27, 1165–1168, (and references therein).

Jacobsen, N.E., Fairbrother, W.J., Kensil, C.R., Lim, A., Wheeler, D.A., Powell, M.F., 1996. Structure of the saponin adjuvant QS-21 and its base-catalysed isomerization product by ^{1}H and natural abundance ^{13}C NMR spectroscopy. Carbohydrate Research 280, 1–14.

Jansson, P.-E., Kenne, L., Widmalm, G., 1989. Computer-assisted structural analysis of polysaccharides with an extended version of CASPER using ^{1}H and ^{13}C NMR data. Carbohydrate Research 188, 169–191.

Kang, S.S., Woo, W.S., 1987. Two new saponins from *Phytolacca Americana*. Planta Medica 53, 338–340.

Kensil, C.R., Wu, J.-Y., Soltysik, S., 1995. Structural and immunological characterization of the vaccine adjuvant QS-21. In: Powell, M.F., Newman, M.J. (Eds.), Vaccine Design: The Subunit Adjuvant Approach, Chapter 22. Plenum Press, New York.

Kensil, C.R., Soltysik, S., Marciani, D., 1996. Enhancing drug delivery with modified saponins. US Patent 52732965.

Nord, L.I., Kenne, L., 1999. Separation and structural analysis of saponins in a bark extract from *Quillaja saponaria* Molina. Carbohydrate Research 320, 71–81.

Nord, L.I., Kenne, L., 2000. Novel acetylated triterpenoid saponin in a chromatographic fraction from *Quillaia saponaria* Molina. Carbohydrate Research 329, 77–89.

Nyberg, N.T., Kenne, L., Rönnberg, B., Sundquist, B.G., 2000. Separation and structural analysis of some saponins from *Quillaja saponaria* Molina. Carbohydrate Research 323, 87–97.

Rastrelli, L., De Simone, F., Schettino, O., Dini, A., 1996. Constituents of *Chenopodium pallidicaule* (Cañihua) seeds: isolation and characterization of new triterpene saponins. Journal of Agriculture and Food Chemistry 44, 3528–3533.

Sawardeker, J.S., Sloneker, J.H., Jeanes, A.R., 1965. Quantitative determination of monosaccharides as their alditol acetates by gas liquid chromatography. Analytical Chemistry 37, 1602–1604.

Shao, Y., Poobrasert, O., Ho, C.-T., Chin, C.-K., Cordell, G.A., 1996. An echinocystic acid saponin derivative from *Kalimeris shimadae*. Phytochemistry 43, 195–200.

van Setten, D.C., van de Werken, G., 1996. Molecular structures of saponins from *Quillaja saponaria* Molina. Advances in Experimental Medicine and Biology 404, 185–193.

van Setten, D.C., ten Hove, G.J., Wiertz, E.J.H.J., Kamerling, J.P., van de Werken, G., 1998. Multiple-stage tandem mass spectrometry for structural characterization of saponins. Analytical Chemistry 70, 4401–4409.

Wu, J.-Y., Gardner, B., Murphy, C., Seals, J., Kensil, C.R., Recchia, J., 1992. Saponin adjuvant enhancement of antigen-specific immune responses to an experimental HIV-1 vaccine. Journal of Immunology 148, 1519–1525.