Tetrahedron Letters, Vol. 22, No. 16, pp 1529 - 1532,1981
Printed in Great Britain

0040-4039/81/161529-04$02.00/0

THE CONFIGURATION AND CONFORMATION OF THE ARABINOSE MOIETY IN PLATYCODINS, SAPONINS ISOLATED FROM PLATYCODON GRANDIFLORUM, AND MI-SAPONINS FROM MADHUCA LONGIFOLIA BASED ON CARBON-13 AND HYDROGEN-1 NMR SPECTROSCOPIC EVIDENCE: THE TOTAL STRUCTURES OF THE SAPONINS

Hiroshi Ishii, Isao Kitagawa,+ Kazuhiro Matsushita,‡ Kiyoharu Shirakawa,+
Kazuo Tori,* Takehiko Tozyo, Masayuki Yoshikawa,+ and Yohko Yoshimura

Shionogi Research Laboratories, Shionogi & Co., Ltd., Fukushima-ku, Osaka, 553 Japan,
+Faculty of Pharmaceutical Sciences, Osaka University, Yamada-kami, Suita, 565 Japan, and
‡Application Center, JEOL Ltd., Nakagami, Akishima, Tokyo, 196 Japan

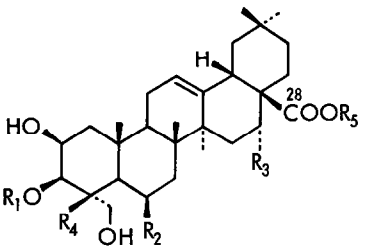
Abstract. ^{13}C and ^{1}H FT NMR spectroscopy provided confirmatory evidence for the anomeric α configuration and the predominance of the $^{1}C_4$ conformation of the L-arabinopyranose moiety in all saponins hitherto isolated from the title plants.

Shibata and coworkers[1] first isolated platycodin-D (1) from the root of Platycodon grandiflorum A. De Candolle (Japanese name, Kikyo) and determined its structure except for the anomeric configuration of the L-arabinopyranose (Ara) moiety. Recently, Shoji and coworkers[2] determined the configuration to be α, by measuring the ^{1}J(C-1,H-1)(Ara) value of 1 to be 168 Hz in its ^{1}H non-decoupled 25-MHz ^{13}C FT NMR spectrum in pyridine-d_5 (C_5D_5N). However, the Shionogi group[3] suggested the β-configuration, tentatively assigning all 15-MHz ^{13}C NMR signals of 1, platycodin-D_2 (2) and -D_3 (3),[4] deapioplatycodin-D (4)[4] and -D_3 (5),[4] polygalacin-D (6) and -D_2 (7), platyconic acid-A (8),[4] and several mono-O-acetyl derivatives in their L-rhamnopyranose (Rha) moiety, because the Ara carbon signals indicated the β-L nature as the C-5 signals appear at relatively higher fields (see the TABLE) when the Ara-ring assumes the $^{4}C_1$ conformation.

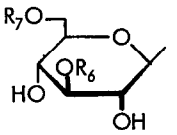
On the other hand, the Osaka University group[5] elucidated the structures of Mi-saponin-A (9) and -B (10) isolated from the seed kernels of Madhuca longifolia (L.) Macbride; the anomeric α configuration of the L-Ara moiety was determined by the ^{3}J(H-1,H-2)(Ara) value of 5.5 Hz obtained from the anomeric ^{1}H doublet signal (100 MHz) of an enzymatic hydrolysis product (16) and its hexa-O-acetyl derivative (16a).[5]

In order to solve the controversial problem of the anomeric configuration, we investigated the ^{13}C and ^{1}H FT NMR spectra of platycodins (1-4), Mi-saponins (9 and 10), crude-hesperidinase hydrolysis[5] products from platycodins (11 and 13), those from Mi-saponins (14 and 16), taka-diastase hydrolysis products from 11 and 14 (12 and 15, respectively), and their peracetates (1a, 9a, and 11a-16a).[6] We report here confirmatory evidence for the configuration and solution conformations of the Ara-rings in these saponins.

Previous tentative ^{13}C signal assignments[3,4] for the C-28 side-chain sugar carbons were first re-examined using the complete ^{1}H-decoupled 25-MHz ^{13}C FT and PRFT spectra of 1-4 in C_5D_5N at 100°C. The inversion recovery rates of the sugar CH signals were found to be in the order of Ara < Rha < D-xylopyranose (Xyl) [< D-apiofuranose (Api)] in 1-3 [4], and the same orders were also found for their ring-CH_2 signals. These observations are consistent with the



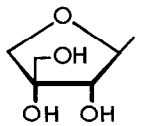
	R_1	R_2	R_3	R_4	R_5
1	β-D-Glc	H	OH	CH_2OH	S_1: Platycodin-D
2	β-Lam	H	OH	CH_2OH	S_1: Platycodin-D_2
3	β-Gen	H	OH	CH_2OH	S_1: Platycodin-D_3
4	β-D-Glc	H	OH	CH_2OH	S_2: Deapioplatycodin-D
5	β-Gen	H	OH	CH_2OH	S_2: Deapioplatycodin-D_3
6	β-D-Glc	H	OH	Me	S_1: Polygalacin-D
7	β-Lam	H	OH	Me	S_1: Polygalacin-D_2
8	β-D-Glc	H	OH	CO_2H	S_1: Platyconic acid-A
9	β-D-Glc	OH	H	Me	S_3: Mi-saponin-A
10	β-D-Glc	OH	H	Me	S_4: Mi-saponin-B
11	H	H	OH	CH_2OH	S_2
12	H	H	OH	CH_2OH	L-Ara—L-Rha
13	H	H	OH	CH_2OH	L-Ara
14	H	OH	H	Me	S_2
15	H	OH	H	Me	L-Ara—L-Rha
16	H	OH	H	Me	L-Ara
	H	H	OH	CH_2OH	H: Platycodigenin (Pla)
	H	OH	H	Me	H: Protobassic acid (Pro)



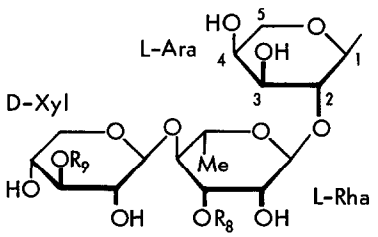
β-D-Glc: $R_6 = R_7 = H$

β-Lam: R_6 = β-D-Glc, $R_7 = H$

β-Gen: $R_6 = H$, R_7 = β-D-Glc



β-D-Api



S_1: $R_8 = H$, R_9 = β-D-Api

S_2: $R_8 = R_9 = H$

S_3: $R_8 = H$, R_9 = α-L-Rha

S_4: R_8 = β-D-Api, R_9 = α-L-Rha

orders of magnitudes of T_1 values of the carbon signals expected from the sugar sequence.[7] Thus, the tentative ^{13}C signal assignments for 1-4 previously reported[3,4] were shown to be almost valid. In a similar manner, the ^{13}C signals of the sugar side-chain in 9 and 10 were assigned in comparison with those of 1-4 and with the aid of the order of the inversion recovery rates of the CH and ring-CH_2 signals; i.e., Ara < Rha < Xyl < (Api) < Rha. The chemical shift δ_C data for the Ara moieties are listed in the TABLE. Thus, ^{13}C signal assignments of sugar moieties of 11, 12, 14, and 15 were straightforward, and the δ_C values for Ara are also listed in the TABLE. In addition, ^3J(H-1,H-2)(Ara) values observed (100 MHz) were about 2-3 Hz for the compounds mentioned here (see the TABLE). Therefore, in all the cases examined, the Ara seemed to be β-L rather than α-L, when the Ara-rings adopt the $^4C_1$ conformation.

Next, we examined ^1H non-decoupled ^{13}C spectra of the above-mentioned compounds in C_5D_5N at 100°C, from which ^1J(C-1,H-1)(Ara) values were obtained (see the TABLE). These values were about 171-173 Hz, which differ from that (168 Hz) reported by Shoji and coworkers,[2] and may be taken as evidence for either the equatorial or axial conformation of the anomeric Ara proton as compared with the values for tetra-O-acetyl-α- (17a, 168 Hz) and tetra-O-acetyl-β-L-arabinopyranoses (18a, 176 Hz).[8] Consequently, these ^1J(C-1,H-1)(Ara) values should not be taken as clear evidence for the anomeric α configuration. Note here that all ^{13}C spectral data concerning the Ara moieties are almost identical among the compounds examined.

TABLE. ^{13}C and ^{1}H NMR Spectroscopic Data for L-Arabinose Moieties of Platycodins, Mi-Saponins, and Their Derivatives[a]

	Platycodigenin series								Protobassate series									
	1-3	4,11	12	13	1a	11a	12a	13a	9,10	14	15	16	9a	14a	15a	16a	17a	18a
	C_5D_5N				$CDCl_3$				C_5D_5N				$CDCl_3$				$CDCl_3$	
δ_C C-1	93.7	93.7	93.7	95.7	93.2	92.7	93.0	91.9	93.5	93.5	93.5	95.8	93.2	93.1	93.4	91.9	92.5	90.6
C-2	75.6	75.8	75.8	71.4	74.2	73.8[f]	73.8[f]	68.7	75.4	75.5	75.4	71.5	73.4[f]	73.1[f]	72.9[f]	68.6	68.9	67.5
C-3	70.4[c]	70.3[c]	70.4	73.5	68.3	68.1[f]	67.8[f]	68.7	70.5	70.9	70.8	73.6	67.4[f]	67.7[f]	67.1[f]	69.6	70.0	68.9
C-4	65.9[c]	66.0[c]	66.3	67.4	68.3	68.3[f]	68.8[f]	66.5	66.1	66.3	66.2	67.5	67.8[f]	67.9[f]	67.6[f]	67.1	67.5	67.2
C-5	62.9	63.0	62.8	65.3	60.4	60.2	60.7	61.9	63.1	63.3	63.3	65.5	62.1	62.5	63.7	62.9	63.7	62.9
^{1}J(C-1,H-1)[d]	172	172	172	165	173	173	173	168	173	171	171	167	169	169	168	169	168 (168	177 176)[g]
δ_H H-1	6.29	6.27	6.36	6.18	5.85	5.84	5.85	5.63	6.29	6.30	6.31	6.19	5.63	5.61	5.54	5.59	5.67	6.36
H-2	b	b	b	b	3.78	3.80	3.85	5.16	b	b	b	b	3.99	4.01	4.82	4.70	5.30	5.37
^{3}J(H-1,H-2)[d]	2.8	3.0	3.0	4.5	2.9	2.9	3.3	5.0	3.0	3.0	3.3	4.9	5.5	5.8	5.2	6.2	6.5	2.3
^{3}J(H-2,H-3)[d]	b	b	b	b	3.5	4.0	4.9	6.5	b	b	b	b	6.8	7.6	9.0	8.5	8.5	b
^{3}J(H-4,H-5)[d]	b	b	b	b	7.5	7.3	7.0	5.5	b	b	b	b	b	b	2.5	4.5	3.5	1.8
Conformation[e]	C	C	C	B	C	C	C	B	C	C	C	B	B	B	B	A	A	A

[a] ^{13}C FT and PRFT NMR spectra were recorded on a JEOL FX-100 FT NMR spectrometer at 25.05 MHz in 10-mm spinning spherical tubes with TMS as an internal reference (δC 0) at 100°C in C_5D_5N and at 60°C in $CDCl_3$. Complete ^{1}H-decoupled and ^{1}H non-decoupled (in the gated mode with NOE) FT NMR measurement conditions were: spectral width, 5 KHz; pulse width, 14 μs (flipping angle 90°); pulse repetition time 1 s (2 s for the gated mode); number of data points, 8K. ^{1}H FT NMR spectra were recorded on a Varian XL-100-12A and/or a Bruker WH-270 FT NMR spectrometer in 5-mm spinning tubes with TMS as an internal reference (δH 0) at 100°C in C_5D_5N at 100.058 MHz for saponins and at 23°C in $CDCl_3$ at 100 and 270 MHz for acetates. ^{1}H FT NMR measurement conditions at 100 and 270 MHz (in parentheses) were: spectral width, 2 KHz (3.012 KHz); pulse width, 3 μs [flipping angle 10°] (10 μs [75°]); acquisition time, 5.0 s (2.7 s), numbers of data points, 20K (16K).

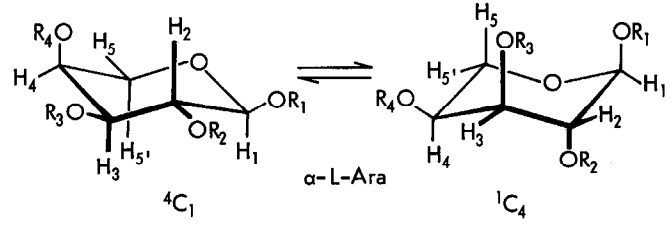
[b] Not observed because of signal overlappings.

[c] These assignments were reversed in refs. 3 and 4.

[d] Predicted values for ^{1}J(C-1,H-1), ^{3}J(H-1,H-2), ^{3}J(H-2,H-3), ^{3}J(H-4,H-5) are: ~165, 7-8, 7-10, 2-3 Hz for α-L-Ara in $^{4}C_1$; ~177, 2-3, 2-3, 7-10 Hz for α-L in $^{1}C_4$; ~177, 3-5, 7-10, 2-3 Hz for β-L-Ara in $^{4}C_1$; ~165, 3-5, 2-3, 7-10 Hz for β-L in $^{1}C_4$, respectively.

[e] Conformations assigned for α-L-Ara: A, $^{4}C_1 >> {}^{1}C_4$; B, $^{4}C_1 > {}^{1}C_4$; C, $^{4}C_1 < {}^{1}C_4$.

[f] Assignments may be reversed.



R_1 = Pla, Pro, or Ac

R_2 = H, Ac, Rha, Rha—Xyl, Rha—Xyl—Api, or Rha—Xyl—Rha

$R_3 = R_4$ = H or Ac

Next, we examined ^{13}C spectra of the acetates (1a, 9a, 11a, 12a, 14a, and 15a) in $CDCl_3$. The data obtained for Ara are shown in the TABLE. Unexpectedly, both δ_C and ^1J(C-1,H-1)(Ara) values for the protobassate (Pro) series (9a, 14a, and 15a) differed from those for the platycodigenin (Pla) series (1a, 11a, and 12a). Also, the ^3J(H-1,H-2)(Ara) values at 100 MHz were about 3 and 5.5 Hz for the Pla and Pro series, respectively. The data observed for the Pro series indicate the α-L feature for Ara in contrast with the Pla series.

Thus, we further investigated the ^{13}C spectra of 13, 16,[5] and their peracetates[6] (13a and 16a[5]). In these spectra, the Ara signals appear to arise from those of α-L- rather than β-L-Ara as compared with those of 17a (see the TABLE). Moreover, their ^3J(H-1,H-2)(Ara) values were found to be about 5-6 Hz as reported earlier.[5] These findings suggest that the anomeric configuration is α, as already pointed out.[5] Rather smaller ^3J(H-1,H-2)(Ara) values for ^3J(ax, ax) may be due to the considerable contribution of the $^1C_4$ conformation.

The discrepancy of the above results seemed to originate from the conformational difference in their Ara rings in solution. We measured the 270- and/or 100-MHz ^1H FT NMR spectra of the compounds. The Ara ^1H signals were assigned by ^1H-decoupling experiments and by comparisons of the spectra of Pla, Pro, and their peracetates.[6] The confirmatory evidence for the anomeric configuration together with the Ara-ring conformation was thus provided by ^3J(H-1,H-2), ^3J(H-2, H-3), ^3J(H-4,H-5), and ^1J(C-1,H-1) values obtained for Ara (see the TABLE for the observed and predicted values). In view of these values, we concluded that the anomeric configuration of Ara is α in all the saponins examined here, and that the Ara-ring predominantly adopts the $^1C_4$ conformation in saponins 1-4, 9-12, 14, and 15, and peracetates 1a, 11a, and 12a, whereas in saponins 13 and 16, and acetates 9a and 13a-16a, the Ara-ring predominantly adopts the $^4C_1$ conformation. Thus, the total structures of all saponins from the title plants have been elucidated.

These results suggest that the O-Rha substitution at C-2 in Ara increases the population of the $^1C_4$ conformation. O-Acetylation slightly stabilizes the $^4C_1$ conformation in the Pro series, whereas acetylation in the Pla series may lead to the 16α-OAc group hindering this conformational change by a non-bonded interaction with the Ara moiety. These trends can also be seen in the saponins in each Pro and Pla series; in the latter, the $^1C_4$ population seems to increase a little by the interaction of the 16α-OH group.

Although the present platycodin case might be the most sophisticated one, much caution should be exercised in determining the configuration of sugars such as L-Ara. Variable-temperature NMR studies as well as more detailed data will be reported in a full paper.

Acknowledgement. We thank Prof. T. Miyazawa of The University of Tokyo for obtaining the 270-MHz ^1H FT NMR spectra.

REFERENCES

1. A. Tada, Y. Kaneiwa, J. Shoji, and S. Shibata, Chem. Pharm. Bull. 20, 1957 (1972).
2. T. Konishi, A. Tada, J. Shoji, R. Kasai, and O. Tanaka, Ibid. 26, 669 (1978).
3. H. Ishii, K. Tori, T. Tozyo, and Y. Yoshimura, Ibid. 26, 674 (1978).
4. H. Ishii, K. Tori, T. Tozyo, and Y. Yoshimura, Chemistry Lett. 719 (1978).
5. I. Kitagawa, A. Inada, and I. Yosioka, Chem. Pharm. Bull. 23, 2268 (1975).
6. In 9a and 14a-16a, the C-6 OH were not acetylated. The properties of the new compounds are: 9a, 11, 11a, 12, 12a, 13, 13a, 14a, and 15a: mp 172-175°, 218-222°, 154-157°, 206-209°, 149-152°, 188-193°, 135-138°, 167-171°, and 169-172°; [α]D, -11.9°, -21.7°, -36.9°, -23.8°, -14.6°, not determined, -3.1°, -14.6°, and +13.8°, respectively.
7. A. Neszmélyi, K. Tori, and G. Lukacs, J.C.S. Chem. Comm. 613 (1977), and references therein.
8. K. Bock and C. Pedersen, Acta Chem. Scand. B29, 258 (1975).

(Received in Japan 16 January 1981)