Pergamon

Tetrahedron Letters, Vol. 35, No. 4, pp. 583-586, 1994
Elsevier Science Ltd
Printed in Great Britain
0040-4039/94 $6.00+0.00

0040-4039(93)E0279-S

Shikometabolins A, B, C and D, Novel Dimeric Naphthoquinone Metabolites Obtained from Shikonin by Human Intestinal Bacteria

Meselhy R. Meselhy,[a] Shigetoshi Kadota,[a,*] Koji Tsubono,[b] Akihiko Kusai,[c]
Masao Hattori [a] and Tsuneo Namba[a]

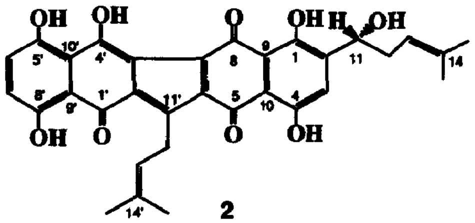
Research Institute for Wakan-Yaku (*Traditional Sino-Japanese Medicines*),[a] Toyama Medical and Pharmaceutical University, 2630 Sugitani, Toyama, 930-01, Analytical Instruments Training and Application Center, JEOL Datum Ltd.[b] and Scientific Instrument Division, JEOL Ltd.,[c] Musashino, Akishima, Tokyo 196, Japan

Summary: Shikometabolins A (2), B (3), C (4) and D (5), four novel naphthoquinone dimers, have been isolated after anaerobic incubation of shikonin (1) with *Bacteroides fragillis* subsp. *thetaotus*, and their structures have been determined by means of 2-D NMR spectroscopy including INADEQUATE experiments.

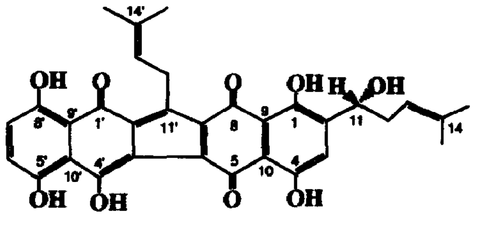
Shikonin (1) is a naphthoquinone pigment found in the root of *Lithospermum erythrorhizon* Sieb. et Zucc. It is produced commercially from *Lithospermum* cell cultures, and used as a natural coloring agent for food, drugs and cosmetics.1) In the course of our studies on the biotransformation of natural pigments by human intestinal flora,2) we found four novel naphthoquinone dimers, shikometabolins A (2), B (3), C (4) and D (5), which were produced by anaerobic incubation of shikonin (1) with *Bacteroides fragilis* subsp. *thetaotus*, one of the predominant bacteria in the human feces. In this letter we report the isolation and structure determination of these four naphthoquinone dimers.

After anaerobic incubation of shikonin (1, 20 g) with a *B. fragilis* suspension (in 20 *l* phosphate buffer, pH 7.3) for three days, the mixture was acidified and extracted with EtOAc. Through rigorous purification of the EtOAc extract using Sephadex LH-20 and preparative TLC, four dimeric metabolites were isolated, and their structures were determined as follows.

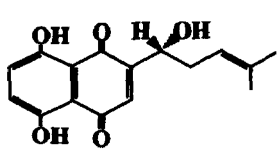
Shikometabolin A (2, 200 mg), dark violet needles from acetone-hexane, showed mp > 300 °C. CD: $[\theta]_{234}$ - 4200 (MeOH).3) The IR spectrum (KBr) showed absorptions for OH (3350 cm^{-1}) and a conjugated ketone (1620 cm^{-1}). The UV spectrum showed λ_{max} MeOH (log ε) at 280 (2.97), 420 (2.52), and 575 (2.78) nm for an extended naphthoquinone. HR-FABMS (negative ion mode) displayed the molecular ion peak at *m/z* 555.1650 $[M-H]^-$ (Calcd for $C_{32}H_{27}O_9$: 555.1658).4) The ^{1}H- and ^{13}C-NMR and ^{1}H-^{1}H COSY (DMSO-d_6) spectra indicated 32 carbon atoms with 20 sp^2 quaternary carbons and suggested the presence of partial structures (I), (II) and (III) (Chart 1). The HMBC and COLOC experiments showed long-range correlations between C-1 and proton signals 11-H and 3-H, the latter of which was correlated with C-4, C-10 and C-11. On the other hand, C-9 and C-10 were respectively correlated with 1-OH and 4-OH, supporting the partial structure (I). Long-range correlations between C-5' and 7'-H and 5'-OH, and between C-8' and 6'-H and 8'-OH, suggested partial structure (II). The COLOC spectrum also showed correlation between 12'-H and C-13', and three sp^2 quaternary carbons, one being assigned to C-14', and the others to C-6 and C-2'. The sequences of most of the quaternary carbons were ultimately confirmed by 2D INADEQUATE.5) It confirmed the connectivities of carbon A (C-1') with carbon X (C-9') and carbon M (C-2'), and carbon B (C-5) with carbon N (C-6) and carbon Y (C-10). Moreover, connectivities of carbon M (C-2') and carbon R (C-3'), carbon N (C-6) and carbon Q (C-7), and carbon J (C-11') and both of M and N were established, partial structure (III) being



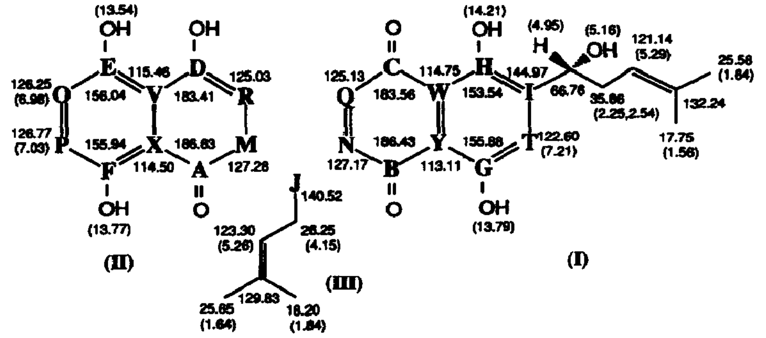
2



3



1



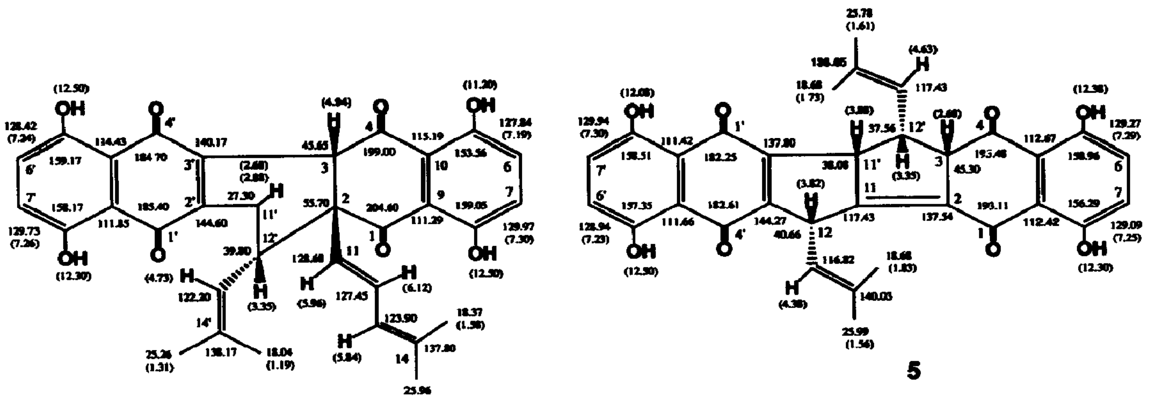
* Values in parentheses indicate ^{1}H-NMR spectral data for 2.

Chart 1

consequently attached to the opposite side of the side chain in (I). The connectivities between carbon C (C-8), W (C-9) and Q (C-7) and between carbon D (C-4'), V (C-10') and R (C-3'), were solved by the application of a ^{13}C-{^{1}H}NOE experiment.$^{6)}$ Irradiation of the hydroxyl proton (1-OH, δ_H 14.21) enhanced signals of carbons C (C-8, δ_c 183.56), H (C-1, δ_c 153.54) and W (C-9, δ_c 114.75), while irradiation of the hydroxyl proton (4-OH, δ_H 13.79) increased the intensity of carbons B (C-5, δ_c 186.43), G (C-4, δ_c 155.88), and Y (C-10, δ_c 113.11). Similarly, the signals of carbons A (C-1', δ_c 186.63), X (C-9', δ_c 114.50) and F (C-8', δ_c 155.94) were enhanced by irradiation of 8'-OH (δ_H 13.77), while irradition of (5'-OH, δ_H 13.54) enhanced the signals of carbons D (C-4', δ_c 183.41), E (C-5', δ_c 156.04) and V (C-9', δ_c 115.46). In view of the molecular formula, a carbon bond directly connecting carbon Q and R was proposed. From these findings, the gross structure of shikometabolin A (**2**) was postulated as shown in formula **2**.

Shikometabolin B (**3**, 170 mg) was obtained as dark violet radiating plates from acetone-hexane, mp > 300 °C. CD: $[\theta]_{234}$ + 11400 (MeOH).$^{7)}$ The two metabolites A (**2**) and B (**3**) were obtained in almost equi-amounts. The IR, UV, ^{1}H-and ^{13}C-NMR spectra of **2** and **3** were quite similar.$^{7)}$ Analysis of the spectroscopic data suggested the same partial structures (I-III) for **3** (Chart 1). However, HMBC, COLOC and ^{13}C-{^{1}H} NOE experiments established the connectivity of C-11' with C-7 and C-2', suggesting the structure **3** with partial structure (III) attached to the same side of the side chain in (I) for shikometabolin B.

Shikometabolin C (**4**, 20 mg), red needles from $CHCl_3$, showed mp 256-258 °C. CD: $[\theta]_{439}$ - 295 (MeOH).$^{8)}$ HRMS displayed the molecular ion peak at *m/z* 540.1809 [M$^+$] (Calcd for $C_{32}H_{28}O_8$: 540.1784). The ^{1}H-^{1}H and ^{1}H-^{13}C COSY ($CDCl_3$) spectra$^{8)}$ showed patterns different from those of **2**, with two allylic groups (δ_H 4.94 and 3.35; δ_c 45.65 and 39.80), four olefinic groups (δ_H 4.73, 5.84, 5.96 and 6.12; δ_c 122.20,



4

5

* Values in parentheses indicate ^{1}H-NMR spectral data.

Chart 2

123.90, 128.68 and 127.45), a cluster of four AB- type aromatic protons at δ_H 7.19, 7.24, 7.26 and 7.30, four *peri*-hydroxyl groups at δ_H 11.20, 12.30 and 12.50, and four carbonyls at δ_c 204.60 (C-1), 199.00 (C-4), 185.40 (C-1') and 184.70 (C-4'). These spectral data led us to speculate structure 4 (Chart 2). The HMBC spectrum showed long-range correlation between the carbonyl carbons C-1 and C-4 and the allylic proton at δ_H 4.94 (3-H), which in turn was correlated with the carbon signals at δ_c 140.17 (C-3'), 55.70 (C-2) and 39.80 (C-12'), indicating connectivity of C-3 and C-3'. Connectivity of C-2 and C-12' was confirmed by the long-range correlation observed between C-2 and the proton signals at δ_H 2.68 (11'-H$_\beta$), 3.35 (12'-H) and 6.12 (12-H), while the carbon signal at δ_c 144.60 (C-2') was correlated 11'-H$_\beta$ and 12'-H. Thus the planar structure of 4 was proved. The relative stereochemistry of 4 was predicted on the basis of the coupling constants of respective protons and NOE experiments. Irradiation at 3-H increased the signal intensity of 12-H, and irradiation at 12'-H enhanced the signal intensities of 11'-H$_\beta$, 11-H and 3-H. Therefore, we concluded that the structure of shikometabolin C was 4.

Shikometabolin D (5, 12 mg) was isolated as a red powder. CD: $[q]_{500}$ - 252 (MeOH).[9] HRMS showed the molecular ion peak at *m/z* 538.1613 [M$^+$] (Calcd for $C_{32}H_{26}O_8$: 538.1626). Extensive analysis of the ^{1}H- and ^{13}C-NMR ($CDCl_3$) spectra[9] indicated patterns similar in part to those of 4, suggesting the formula 5. Long-range correlations observed in the HMBC spectrum, confirmed the connectivities of C-3 with C-12', C-11 with C-11', and C-3' with C-12. Three bond correlation from 3-H (δ_H 2.68) to the carbonyl carbon C-1 (δ_c 193.11), the olefinic carbon C-13' (δ_c 117.43) and the allylic carbon C-11' (δ_c 38.08), connecting C-3 and C-12'. The connectivity of C-11 and C-11' was established by the correlation between 12-H (δ_H 3.82) and the carbon signals C-4' (δ_c 182.61), C-2' (δ_c 137.80), and C-11' (δ_c 38.08). Moreover, 11'-H (δ_H 3.88) was correlated with C-1'(δ_c 182.25), C-3' (δ_c 144.27), C-2 (δ_c 137.54) and C-3 (δ_c 45.30), confirming the connectivity of C-3' and C-12. Consequently, the planar structure was concluded to be 5. The relative stereochemistry of 5 was deduced from the coupling constants and NOE difference spectra. Irradiation at 3-H enhanced the signal intensities of 12'-H and 13'-H, and measurable NOE effects of 3-H, 13'-H, 12-H and 11'-H were observed after irradiation at 12'-H. Irradiation at 11'-H intensified 12-H and 12'-H, and irradiation at 12-H increased the signals intensity of 11'-H and 13-H. From the foregoing evidence, we concluded that shikometabolin D was 5.

Our present results provide the first dimeric naphthoquinone metabolites isolated after anaerobic incubation of shikonin (1) with a human intestinal bacterial strain. Among these dimers, shikometabolin A (2) inhibited reverse transcriptase, one of the essential enzymes of human immunodefficency virus (HIV-1 RT) with an IC_{50} of 0.71 mM. Under the same conditions, adriamycin, used as a positive control, had an IC_{50} = 0.12 mM. Further studies of shikonin biotransformation will be reported elsewhere.

REFERENCES AND NOTES:

1. a) Inouye, H.; Matsumura, H.; Kawasaki, M.; Inoue, K.; Tsukada, M.; Tabata, M. *Phytochemistry* **1981**, *20*, 1701-1705; b) Mizukami, H.; Konoshima, M.; Tabata, M. *ibid.* **1978** , *17*, 95-97; c) Tabata, M.; Mizukami, H.; Hiraoka, N.; Konoshima, M. *ibid.* **1974** , *13*, 927-932.
2. Meselhy, M. R.; Kadota, S.; Hattori, M.; Namba, T. *J. Nat. Prod.* **1993**, *56*, 39-45.
3. **2**: CD (*c* = 0.67 mM, MeOH): $[\theta]_{234}$ - 4200, $[\theta]_{276}$ +1800, $[\theta]_{304}$ -1800. ^{1}H-NMR (DMSO-d_6) δ_H :1.56 (3H, d, *J* = 2 Hz, 15-CH_3), 1.63 (3H, s, 16'-CH_3), 1.67 (3H, s, 16-CH_3), 1.84 (3H, d, *J* = 2 Hz, 15'-CH_3), 2.25 (1H, dt, *J* = 12, 7 Hz, 12-H_β), 2.54 (1H, dd, *J* = 12, 3.5 Hz, 12-H_α), 4.15 (2H, d, *J* = 7 Hz, 12'-H), 4.95 (1H, dd, *J* = 7, 3.5 Hz, 11-H), 5.16 (1H, s, 11-OH), 5.26 (1H, dd, *J* = 7, 2 Hz, 13'-H), 5.29 (1H, dd, *J* = 7, 2 Hz, 13-H), 6.98 (1H, d, *J* = 9 Hz, 6'-H), 7.03 (1H, d, *J* = 9 Hz, 7'-H), 7.21 (1H, s, 3-H), 13.54 (1H, s, 5'-OH), 13.77 (1H, s, 8'-OH), 13.79 (1H, s, 4-OH), 14.21 (1H, s, 1-OH).
4. Field desorption (FD) and Electrospray Ionization (ESI) mass spectra of **2** indicated the same molecular ion peak.
5. Bax, A. "*2D NMR in Liquids*, " D. Reidel Publishing Co., Dordrecht, Holland, 1982, pp. 155-174; Turner, D. L. *J. Magn. Res.* **1982**, *49*, 175-178.
6. Ford, J. J.; Gibbons, W. A.; Niccorai, N. *J. Magn. Res.* **1982**, *47*, 522-527.
7. **3**: HR-FABMS showed the molecular ion peak *m/z* 555.1650 (Calcd for $C_{32}H_{27}O_9$: 555.1658). CD (*c* = 0.67 mM, MeOH): $[\theta]_{218}$ + 15600, $[\theta]_{234}$ +11400, $[\theta]_{270}$ -3000, $[\theta]_{302}$ -3000. ^{1}H-NMR (DMSO-d_6) δ_H : 1.55 (3H, s, *J* = 1.5 Hz, 15-CH_3), 1.63 (3H, d, *J* = 1.5 Hz, 15'-CH_3), 1.68 (3H, s, 16-CH_3), 1.82 (3H, s, 16'-CH_3), 2.20 (1H, dt, *J* = 13, 7 Hz, 12-H_β), 2.50 (1H, m, 12-H_α), 4.14 (2H, d, *J* = 7 Hz, 12'-H), 4.91 (1H, br t, 11-H), 5.25 (1H, dd, *J* = 7, 1.5 Hz, 13'-H), 5.29 (1H, dd, *J* = 7, 1.5 Hz, 13-H), 7.06 (1H, d, *J* = 9 Hz, 7'-H), 7.09 (1H, d, *J* = 9 Hz, 6'-H), 7.17 (1H, s, 3-H), 13.76 (2H, s, 4-OH and 5'-OH), 13.80 (1H, s, 8'-OH), 14.34 (1H, s, 1-OH). ^{13}C-NMR (DMSO-d_6) δ_c: 17.51 (C-15), 17.94 (C-16'), 25.32 (C-16), 25.39 (C-15'), 26.12 (C-12'), 35.52 (C-12), 113.70 (C-9), 113.98 (C-10), 114.44 (C-9'), 115.38 (C-10'), 120.97 (C-13), 122.90 (C-3), 123.08 (C-13'), 124.72 (C-6), 125.16 (C-3'), 125.98 (C-7'), 126.49 (C-6'), 126.83 (C-7), 127.04 (C-2'), 129.62 (C-14'), 132.06 (C-14), 140.59 (C-11'), 144.23 (C-2), 153.29 (C-1), 155.71 (C-8), 155.85 (C-4 and C-5'), 183.00 (C-5), 183.15 (C-4'), 186.70 (C-1'), 187.04 (C-8).
8. **4**: CD (*c* = 3.71 mM, MeOH): $[\theta]_{439}$ - 295, $[\theta]_{486}$ +269, $[\theta]_{520}$ -269. UV λ_{max} $CHCl_3$ (log ε): 250 (3.20), 410 (2.73), 520 (2.66), 555 (2.43) nm. ^{1}H-NMR ($CDCl_3$) δ_H: 1.19 (3H, d, *J* = 1.5 Hz, 15'-CH_3), 1.31 (3H, d, *J* = 1.0 Hz, 16'-CH_3), 1.58 (3H, d, *J* = 1.5 Hz, 15-CH_3), 1.73 (3H, d, *J* = 1 Hz, 16-CH_3), 2.68 (1H, ddd, *J* = 20.0, 5.5, 1 Hz, 11'-H_β), 2.88 (1H, dd, *J* = 20.0, 1.5 Hz, 11'-H_α), 3.35 (1H, dd, *J* = 10.5, 5.5 Hz, 12'-H), 4.73 (1H, ddt, *J* = 10.5, 1.5, 1.0 Hz, 13'-H), 4.94 (1H, d, *J* = 1.0 Hz, 3-H), 5.84 (1H, ddt, *J* = 10.5, 1.5, 1.0 Hz, 13-H), 5.96 (1H, d, *J* = 15.5 Hz, 11-H), 6.12 (1H, dd, *J* = 15.5, 10.5 Hz, 12-H), 7.19 (1H, d, *J* = 9 Hz, 6-H), 7.24 (1H, d, *J* = 9 Hz, 6'-H), 7.26 (1H, d, *J* = 9 Hz, 7'-H), 7.30 (1H, d, *J* = 9 Hz, 7-H), 11.20 (1H, s, 5-OH), 12.30 (1H, s, 8'-OH), 12.50 (2H, s, 8-OH and 5'-OH).
9. **5**: CD (*c* = 1 mM, MeOH): $[\theta]_{280}$ - 252, $[\theta]_{500}$ -252. UV λ_{max} $CHCl_3$ (log ε): 220 (2.92), 420 (2.57), 520 (2.47) nm. ^{1}H-NMR ($CDCl_3$) δ_H: 1.56 (3H, d, *J* = 1.0 Hz, 16-CH_3), 1.61 (3H, d, *J* = 1.0 Hz, 16'-CH_3), 1.73 (3H, d, *J* = 2 Hz, 15'-CH_3), 1.83 (3H, d, *J* = 2 Hz, 15-CH_3), 2.68 (1H, dd, *J* = 3, 1 Hz, 3- H), 3.35 (1H, ddd, *J* = 7.5, 4.5, 3 Hz, 12'-H), 3.82 (1H, dd, *J* = 9.0, 4.5 Hz, 12-H), 3.88 (1H, td, *J* = 4.5, 1 Hz, 11-H), 4.38 (1H, ddd, *J* = 9.0, 2.0 Hz, 13-H), 4.63 (1H, ddd, *J* = 7.5, 2.0 Hz, 13'-H),7.23 (1H, d, *J* = 9.0 Hz, 6'-H), 7.25 (1H, d, *J* = 9.0 Hz, 7-H), 7.29 (1H, d, *J* = 9.0 Hz, 6-H), 7.30 (1H, d, *J* = 9.0 Hz, 7'-H), 12.08 (1H, s, 8'-OH), 12.30 (1H, s, 8-OH), 12.38 (1H, s, 5-OH), 12.50 (1H, s, 5'-OH).

(*Received in Japan* 21 *August* 1993; *accepted* 27 *September* 1993)