ORGANIC LETTERS

2010
Vol. 12, No. 7
1504−1507

Aromatized C-2 Oxygenated Triterpenoids as Indicators for a New Transformation Pathway in the Environment

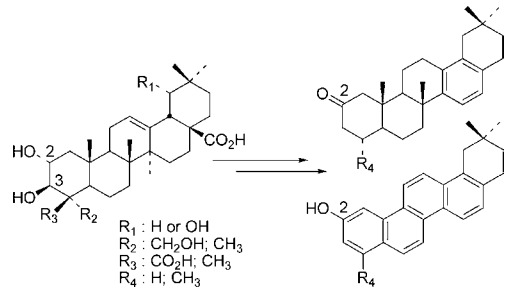
Claude Le Milbeau,[†] Philippe Schaeffer, Jacques Connan, Pierre Albrecht, and Pierre Adam*

Laboratoire de Biogéochimie Moléculaire, Université de Strasbourg, Institut de Chimie de Strasbourg UMR7177 CNRS, 25 rue Becquerel, 67200 Strasbourg, France

padam@unistra.fr

Received January 28, 2010

ABSTRACT



Four oleanane-related aromatic triterpenoids isolated from an archeological piece of oak wood have been identified by NMR studies. Their structures give clue to a novel diagenetic transformation pathway of 2,3-oxygenated higher plant triterpenoids in the environment.

Structural transformations undergone by C-3-oxygenated triterpenoids from decaying higher plants in the environment have been extensively investigated. In particular, the aromatization processes are among the most important microbially mediated alterations undergone by higher plant triterpenoid skeletons at the earliest stages of diagenesis.[1] It may start with the loss of the ring A of the triterpenoids, leading to the formation of mono- and polyaromatic tetracyclic skeletons (Scheme 1). Alternatively, ring A may be aromatized, triggered by the loss of the functionality at C-3, with the aromatization proceeding gradually from ring A to ring D (or E; Scheme 1). We report here the identification by extensive NMR studies of four novel triterpenoid phenols and ketones (**1**−**4**; Figure 1) from the oleanane series detected in three archeological pieces of oak wood collected in sediments of the Saône river (France). These samples originate from two bridges and a boat built during the Roman epoch and dating back to the first and third century AD.[2] The identified triterpenoids are characterized by the rather unusual presence of a functionality located at C-2 of the triterpene skeleton. We have also detected these novel structures in various freshwater recent sediments. These novel compounds give clues to the transformations undergone in the subsurface by triterpenoids with oxygenated functions at C-2 and C-3, a class of triterpenes widely distributed in higher plants.[3]

Analysis of the organic extract of the archaeological oak wood samples by gas chromatography−mass spectrometry (GC−MS) showed very similar distributions of lipidic constituents comprising linear alcohols and fatty acids

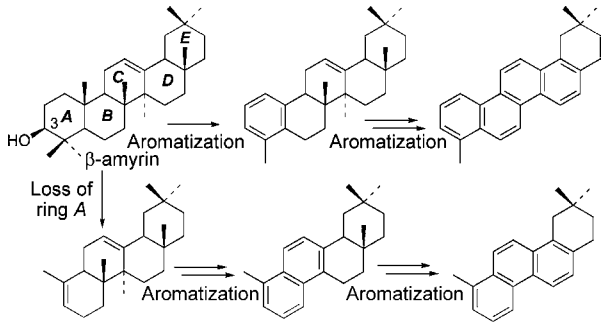
[†] Present address: Plate-forme Analytique Géochimie Organique et Minérale, Institut des Sciences de la Terre d'Orléans, Université d'Orléans, 1A, rue de la Férollerie, 45071 Orléans, France.

(1) (a) Spyckerelle, C.; Greiner, A. C.; Albrecht, P.; Ourisson, G. *J. Chem. Res., Synop.* **1977**, 330. (b) Wolff, G. A.; Trendel, J. M.; Albrecht, P. *Tetrahedron* **1989**, *45*, 6721. (c) Lohmann, F.; Trendel, J. M.; Hetru, C.; Albrecht, P. *J. Labelled Compd. Radiopharm.* **1990**, *28*, 377. (d) Stout, S. *Org. Geochem.* **1992**, *18*, 51.

(2) Bonnamour, L. *Les ponts romains de Chalon. Archéologie de la Saône, 150 ans de recherche: le fleuve gardien de la mémoire*; Errance: Paris, 2000; pp 83−88.

10.1021/ol100144t
Published on Web 03/10/2010

Scheme 1. Microbially Induced Degradation and Aromatization Pathways of C-3 Oxygenated Triterpenoids (e.g., β-Amyrin) during Early Diagenesis (Adapted from Wolf et al.)[1b]



($C_{24}-C_{28}$), C_{29} sterols, along with two pairs of predominant unknown compounds (**1**−**4**).

Compounds **1**−**4** could be isolated from the extract of one of the wood pieces from the third century (220−230 AD) using a sequence comprising fractionation by silica gel column and reversed-phase HPLC and yielded ca. 3 mg of each compound **1**−**4** with a purity >90% (GC), allowing their identification by high-resolution mass spectrometry and extensive NMR studies.

Identification of 1 and 2. The molecular formulas, $C_{25}H_{24}O$ and $C_{24}H_{22}O$, respectively, obtained for **1** and **2** by high resolution mass spectrometry analysis are compatible with those of tetraaromatic pentacyclic phenols, as could be proposed from the interpretation of the EI spectra of **1** and **2** which display predominant molecular ions ([M^+] 340 and 326, respectively) and present very similar mass fragmentation patterns characterized by an important [$M^+ - 56$] fragment possibly explained by a retro-Diels−Alder-type fragmentation.

Interpretation of 1D ^{13}C and DEPT spectra of **1** reveals the presence of eighteen aromatic and seven nonaromatic carbon atoms. The ^1H NMR spectrum shows the presence of two methyl group signals, one integrating for six protons and corresponding to two superimposed singlet methyl signals and the other corresponding to a benzylic methyl group (2.75 ppm). The basic skeleton could be established from the long-range ($^{2,3}J_{CH} = 7$ Hz) $^1H-^{13}C$ correlation network. The two superimposed methyl groups (C-29 and C-30) display connections to three other common carbon atoms, a quaternary carbon atom (C-20) and two methylenes (C-21 and C-19), respectively, the latter being located at a benzylic position (δ_H: 3.03 ppm). Connections of H-21 to C-29/C-30, C-20, C-19, a benzylic methylene C-22 (δ_H: 3.04 ppm), and a quaternary aromatic carbon atom C-17 could

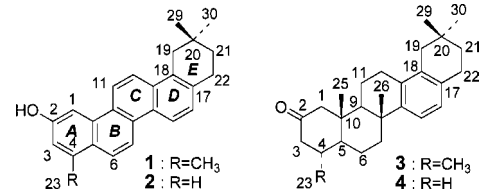


Figure 1. Novel 2-oxygenated triterpenoid structures identified from oak wood buried in freshwater sediments. Absolute stereochemistry of **3** and **4** is based on that known for triterpenes of the oleanane series occurring in higher plants.[4]

also unambiguously be observed. The carbon sequence deduced from the HMBC experiment shows therefore that the E ring corresponds to that of the triterpenoids from the oleanane series and is directly connected to an aromatic ring (ring D).

These data are, thus, compatible with a pentacyclic tetraaromatic triterpenoid phenol of the oleanane series. The complete skeleton was further established from the $^1H-^{13}C$ long-range ($^{2,3}J$) correlations. One particular structural feature deduced from this experiment is the localization of the phenol hydroxyl group at C-2 which is rather unexpected considering that triterpenoids of the oleanane series are generally oxygenated at position 3. In this respect, the overall structure of ring A deduced from the HMBC experiment was further supported by detailed investigation of the NOESY experiment which allowed the connectivities between ^1H-coupled sectors to be established. Thus, the detection of nuclear Overhauser effects (NOE) between H-11/H-1, H-1/OH, OH/H-3, H-3/CH_3-23, and CH_3-23/H-6 unambiguously allowed the assignment of the relative positions of CH_3-23, H-1 and H-3, and the phenolic hydroxyl group on ring A. The ^1H NMR spectrum of **2** is very similar to that of compound **1**. The main differences observed are, respectively, the absence of a benzylic methyl group and a modification of the spectrum in the aromatic region compatible with those expected to be induced by the removal of CH_3-23 (e.g., replacement of one singlet aromatic signal by two doublets). Detailed NMR investigation by homonuclear ($^1H-^1H$: COSY and NOESY) and heteronuclear ($^1H-^{13}C$: HSQC and HMBC) correlation experiments unambiguously confirmed the identification of **2** as being the 23-nor analogue of **1**.

Identification of 3 and 4. High-resolution mass spectrometry analysis of compounds **3** and **4** gave the molecular formulas $C_{26}H_{37}O$ and $C_{27}H_{39}O$, respectively, for protonated molecules $[M + H]^+$, indicating that the isolated compounds contain one oxygen atom. Compounds **3** and **4** are obviously homologues differing by one methylene or methyl group and share a similar mass fragmentation pattern (fragments at *m*/*z* 199, 213, and 225). The latter presents strong similarities with that of ring B monoaromatic fernane derivatives.[5] They were thus postulated to be monoaromatic triterpenoid ketones possibly related to the phenols **1** and **2**. The structures of

(3) See, e.g.: (a) Mahato, S. B.; Nandy, A. K.; Roy, G. *Phytochemistry* **1993**, *31*, 2199. (b) De Rosa, S.; Mitova, M.; Handjieva, N. A.; Popov, S.; Anchev, M. *J. Nat. Prod.* **2000**, *63*, 1012. (c) Calis, I.; Kirmizibekmez, H.; Tasdemir, D.; Rueedi, P. *Helv. Chim. Acta* **2004**, *87*, 611. (d) Allouche, Y.; Jimenez, A.; Uceda, M.; Aguilera, M. P.; Gaforio, J. J.; Beltran, G. *J. Agric. Food Chem.* **2009**, *57*, 3604.

(4) See, e.g.: (a) Corey, E. J.; Lee, J. *J. Am. Chem. Soc.* **1993**, *115*, 8873. (b) Wendt, K. U.; Schulz, G. E.; Corey, E. J.; Liu, D. R. *Angew. Chem., Int. Ed.* **2000**, *39*, 2812.

(5) Hauke, V.; Graff, R.; Wehrung, P.; Trendel, J. M.; Albrecht, P.; Schwark, L.; Keely, B. J.; Peakman, T. M. *Tetrahedron* **1992**, *48*, 3915.

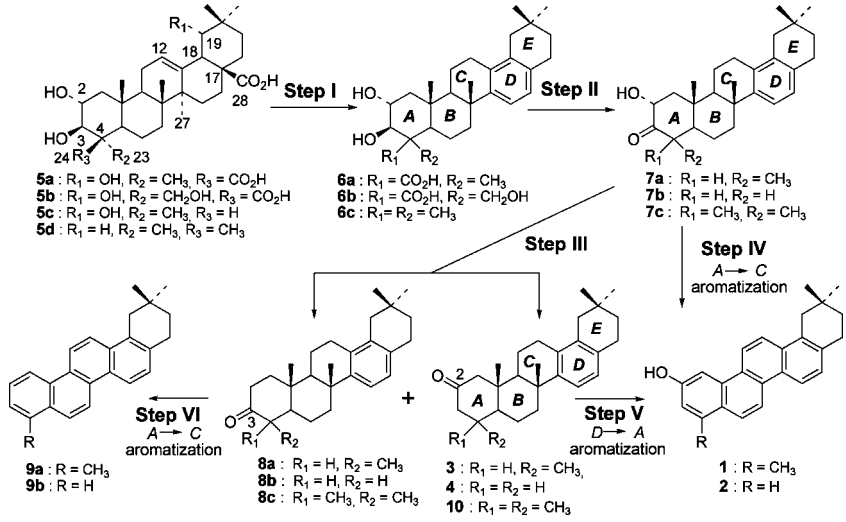


Figure 2. Proposed early diagenetic transformation pathway of 2,3-dioxygenated higher plant triterpenoids.

compounds **3** and **4** were unambiguously determined by means of extensive NMR studies similar to those used for the identification of **1** and **2**.

The ^{13}C NMR spectrum of compound **3** displays the signal of six aromatic carbon atoms (121.9−147.4 ppm range), thus confirming the presence of an aromatic ring. In addition, the signal at 211.3 ppm typically shows the presence of a carbonyl group. The ^{1}H NMR spectrum of **3** shows two protons at, respectively, 6.91 and 7.02 ppm, confirming the presence of one aromatic ring as well as the deshielded protons in the 1.9−2.8 ppm range likely corresponding to protons located at benzylic positions or α to a carbonyl. Four singlet methyl groups as well as one doublet corresponding to a methyl group on a methine could also be detected. The basic skeleton of **3** could be established from the long-range $^{1}H-^{13}C$ correlation network mainly starting from the methyl groups. In particular, a $^{1}H-^{13}C$ correlation network similar to that observed for rings D and E of compounds **1** and **2** could be evidenced in the case of compound **3**, showing that this compound has a similar ring D/E moiety typical of aromatized oleanane-related triterpenoids. The structure of rings A and B could be determined from the $^{1}H-^{13}C$ correlations starting from methyls CH_3-23, CH_3-25, and CH_3-26. A key structural feature was the localization of the carbonyl at C-2 similar to what has been observed for **1** and **2**. This could be established mainly based on the long-range $^{1}H-^{13}C$ correlation network since C-1, which has a typical chemical shift of a carbon located in the vicinity of a carbonyl group (55.7 ppm), can be "seen" by H-5 and H-25. Conversely, the presence of remote $^{2,3}J$ connections between protons at C-1 and carbon atoms C-5, C-10, C-25, and the C-2 carbonyl further supports this structural assignment. In addition, the H-1 protons appear as doublets, clearly indicating that there are no hydrogen atoms on the carbon atoms vicinal to C-1.

The NMR data of compound **4** show many analogies with those of **3**. The most remarkable difference resides in the absence of the doublet signal corresponding to the methyl located at position 4 on ring A. Further investigation of the 2D correlation patterns fully supports the identification of **4** as the 23-nor analogue of **3**.

From their structure, it appears that the four newly identified compounds are oleanane-related triterpenoids deriving from triterpenes present in the oak wood. The fact that the newly identified compounds all bear an oxygenated function at C-2 instead of the usual C-3 position strongly suggests that they might derive from 2,3-dioxygenated triterpenoid precursors widely distributed in higher plants.[3] Indeed, to the best of our knowledge, higher plant triterpenoids exclusively functionalized at position 2 have not been reported from oak, nor from any other plant species, whereas 2,3-dioxygenated triterpenoids have been identified from various oak species and comprise, notably, compounds **5a** and **5b** (Figure 2) occurring mainly as their related glycosides.[6] The latter can be regarded as potential precursor molecules of compounds **1**−**4**. Indeed, formation of **1**−**4** might be explained starting from **5a** and **5b** by a process involving, as a first step (step **I**; Figure 2), the aromatization of ring D. This step might also comprise the elimination of the hydroxyl group at C-19 followed by the decarboxylation at C-17 triggered by the presence of the Δ^{18} double bond and, finally, the removal of the C-27 methyl group, allowing aromatization of ring D. The further loss of the carboxyl and hydroxymethylene groups at C-23 and C-24 on intermediates **6a** and **6b** (Figure 2) might be explained by, successively, the oxidation of the C-3 alcohol functionality to a ketone followed by a decarboxylation and/or a retroaldol reaction leading to **7a** and **7b** (step **II**; Figure 2). Similar processes have been previously proposed to explain the

(6) (a) Romussi, G.; Parodi, B.; Falsone, G. *Pharmazie* **1983**, *38*, 787. (b) Arramon, G.; Saucier, C.; Colombani, D.; Glories, Y. *Phytochem. Anal.* **2002**, *13*, 305. (c) Fontana, N.; Bisio, A.; Romussi, G. *Pharmazie* **1998**, *53*, 653. (d) Chen, H. D.; Yang, S. P.; Zhang, C. R.; Yue, J. M. *Helv. Chim. Acta* **2006**, *89*, 1971.

formation of sedimentary 24- or 28-nor-triterpenoid hydrocarbons from unsaturated and ketotriterpenoids.[7] It should also be mentioned, in this respect, that 24-nor triterpenoids (e.g., **5c** and the related C-3 ketone), which might be considered as possible precursors of **2** and **4**, have been identified in *Quercus aliena*.[6d]

Formation of **3** and **4** from **7a** and **7b** would involve further removal of the functionality at C-3 (step **III**; Figure 2) via unknown reductive processes, possibly involving elimination of H_2O on intermediate diols resulting from the reduction of the C-3 ketone.

Two alternative pathways may be proposed to account for the formation of the phenols **1** and **2**. The first pathway involves the aromatization of 2,3-oxygenated intermediates like **7a** or **7b** or of related 2,3-diols (step **IV**; Figure 2) according to the microbially induced process prevailing for 3-oxygenated triterpenoids (Scheme 1).[1]

This aromatization would start from the oxygenated functionality located at C-3. The functionnality at C-2 would be preserved as a phenolic hydroxyl. However, it cannot be excluded that the oxygenated functionality at C-2 prevents such a microbially triggered aromatization process. Consequently, in the case of 2,3-dioxygenated triterpenoid precursors, aromatization would not start in ring A, but at other sites where functionalities (e.g., double bonds or oxygenated functionalities) likely to trigger the process are present as is the case here in ring D. Such an aromatization would then progress from ring D to ring A following a process different from that known to affect higher plant triterpenoids at the earliest stages of diagenesis. According to such a process, compounds **1** and **2** might result from the stepwise aromatization of **3** and **4** progressing from ring D to ring A (step **V**; Figure 2). Since this aromatization occurs under mild temperature and pH conditions and within a rather short time frame (<2000 years), this process is likely microbially induced, as opposed to a thermal process occurring at later stages of sedimentary burial.

It is worth mentioning that isomers of compounds **1**−**4** bearing the oxygenated functionality at C-3 (e.g., **8a** and **8b**) could not be detected. This reflects a high selectivity of the removal of the functionality at C-3 (step **III**; Figure 2), further suggesting that this reductive step is biologically triggered. Alternatively, the 3-keto analogues of **3** and **4** (e.g., **8a** and **8b**), if indeed formed, could be quantitatively transformed via the "classical" aromatization process described in Scheme 1 into aromatic hydrocarbons like **9a** and **9b**, indeed present in the investigated samples.[1]

The diagenetic pathway presented in Figure 2 for 2,3-dioxygenated triterpenoids bearing additional functionalities at C-23, C-24, and C-19 might be more general and could also apply for the more widely distributed 2,3-dioxygenated triterpenoids bearing nonfunctionalized methyl groups at C-4 (e.g., maslinic acid **5d**).[3] Phenol **1** might thus be formed following step **IV** (Figure 2) according to the process described in Scheme 1 via the intermediate **7c** derived from **5d**. Compound **10**, a C-4 dimethylated analogue of **3**, might also be formed from **7c** via step **III** (Figure 2). In this respect, we have indeed tentatively identified such a dimethylated analogue **10** in freshwater sediments by GC−MS analysis. Finally, it cannot be excluded that intermediates with a functionalized methyl at C-4 (such as **6a**) can be formed by microbial transformation of **5d**, thus opening another pathway toward the formation **1** and **3**. The diagenetic pathway followed by 2,3-oxygenated triterpenoids, a series of terpenoids largely distributed among higher plants,[3] is thus different from that known to affect 3-oxygenated triterpenoids at the earliest stages of diagenesis. This pathway would have similarities with that reported for hopanoid pentacyclic triterpenoids[8] and would start, notably, by the aromatization of the ring D of the hydrocarbon skeleton (Figure 2).

Acknowledgment. We thank E. Motsch and P. Wehrung (Université de Strasbourg) for mass spectral analysis, L. Allouche, R. Graff, J. D. Sauer, and M. Schmitt (Université de Strasbourg) for NMR measurements, and L. Bonnamour (Musée Denon, Chalon-sur-Saône, France) for providing the archaeological samples.

Supporting Information Available: Experimental procedures and spectroscopic data (MS, HRMS, NMR) for all newly identified compounds. This material is available free of charge via the Internet at http://pubs.acs.org.

OL100144T

(7) (a) Peakman, T. M.; ten Haven, H. L.; Rullkötter, J.; Curiale, J. A. *Tetrahedron* **1991**, *47*, 3779. (b) Trendel, J. M.; Graff, R.; Albrecht, P. *Tetrahedron Lett.* **1991**, *32*, 2959. (c) Armanios, C.; Alexander, R.; Kagi, R. I.; Skelton, B. W.; White, A. H. *Org. Geochem.* **1995**, *23*, 21. (d) Rouquette, N.; Schaeffer, P.; Schweigert, M. C.; Trendel, J. M.; Kowalewski, I.; Levaché, D.; Albrecht, P. *Org. Geochem.* **2005**, *36*, 1227.

(8) Greiner, A. C.; Spyckerelle, C.; Albrecht, P.; Ourisson, G. *J. Chem. Res., Synop.* **1977**, 334.