Phytochemistry 1971 Vol. 10, pp. 3161 to 3166. Pergamon Press. Printed in England.

THE STRUCTURES OF SALICORTIN AND TREMULACIN*†

IRWIN A. PEARL and STEPHEN F. DARLING
The Institute of Paper Chemistry, Appleton, Wisconsin 54911, U.S.A.

(*Received* 10 *March* 1971)

Abstract—The structure of the labile glucoside, salicortin, an important component of the bark and leaves of all *Populus* and *Salix* species, was proved to be the ω-(1-hydroxy-6-oxo-2-cyclohexene-1-carboxylic acid ester) of salicin by means of hydrogenolysis, acid and alkaline hydrolysis, and IR, NMR and mass spectra. In similar manner, the structure of tremulacin, an important glucoside component of several *Populus* species barks and leaves, was demonstrated to be the 2-*O*-benzoyl ester of salicortin.

INTRODUCTION

IN 1964, Thieme[1] isolated a new glucoside, salicortin, from the bark of *Salix purpurea*, and in a subsequent paper,[2] demonstrated by means of paper chromatography that salicortin was an important component of all the *Populus* and *Salix* species barks he had investigated. Thieme[1] reported that salicortin was not hydrolysed by emulsin and that it was converted by hydrochloric acid into ω-salicyloylsalicin. Cleavage with sodium methylate gave salicin and a colorless oil which, upon alkaline hydrolysis, yielded salicylic acid and an unidentified aliphatic hydroxy acid. These results led Thieme to conclude that salicortin is a derivative of ω-salicyloylsalicin in which the primary hydroxyl of the glucose moiety and the hydroxyl of the salicyloyl moiety are bound by an unidentified aliphatic hydroxy acid.

In our continuing studies on the barks and leaves of the family Salicaceae, we have isolated salicortin in quantity from the barks and leaves of a number of *Populus* and *Salix* species[3–6] and have confirmed Thieme's earlier conclusion[7] that this labile glucoside is probably the precursor of much of the salicin found in these materials. The possible structure suggested by Thieme[1] became suspect when his salicortin acetate‡ was submitted to mass spectrometry and exhibited major peaks at *m/e* 331, 169 and 109. These peaks, corresponding with the tetraacetylglucoseoxonium ion and two of its fragmentation ions, demonstrated that the glucose moiety in salicortin was unsubstituted,[8] and therefore, could not be bound by the unidentified aliphatic hydroxy acid. In addition, β-glucosidase was found to cleave salicortin to yield glucose as one of the products, thus confirming the

* Part XXVIII in the series "Studies on the Barks of the Family Salicaceae" and Part XV in the series "Studies on the Leaves of the Family Salicaceae".

† A preliminary communication of these results appeared in *Tetrahedron Letters* **44**, 3827 (1970).

‡ Kindly supplied by DR. H. THIEME, Karl-Marx University, Leipzig, Germany.

[1] H. THIEME, *Die Pharmazie* **19**, 725 (1964).
[2] H. THIEME, *Planta Medica* **15**, 35 (1967).
[3] I. A. PEARL and S. F. DARLING, *Phytochem.* **8**, 2393 (1969).
[4] I. A. PEARL and S. F. DARLING, *Phytochem.* **9**, 1277 (1970).
[5] I. A. PEARL and S. F. DARLING, *Phytochem.* **10**, 483 (1971).
[6] I. A. PEARL and S. F. DARLING, *Can. J. Chem.* **49**, 49 (1971).
[7] H. THIEME, *Planta Medica* **13**, 431 (1965).
[8] I. A. PEARL and S. F. DARLING, *Phytochem.* **7**, 831 (1968).

unsubstituted glucose moiety in the glucoside. Subsequent experiments proved that Thieme's conclusions were invalid in several other respects and led to our investigation of the structure of salicortin (I), m.p. 135–137°, and that of the related glucoside, tremulacin (II), m.p. 122–123°, originally isolated from *P. tremula* bark by Thieme and Richter,[9] and which we isolated from *P. tremuloides* bark.[5] The present paper reports the results of our studies.

RESULTS

The structure of I has been established as the ω-(1-hydroxy-6-oxo-2-cyclohexene-1 carboxylic acid ester) of salicin and that of II as the 2-*O*-benzoyl ester of I in the following manner. A solution of I in dilute NaOH, after standing at 22° for 15 min, was acidified with dilute HCl to yield salicin (III), carbon dioxide, and an oil which on further alkaline treatment, yielded pyrocatechol (IV). No trace of salicylic acid (V) could be found, and thus, I could not be a derivative of ω-salicyloylsalicin (VI). Nevertheless, treatment with acid converted I into VI in high yield, and the assigned structure for I had to resolve this apparent anomaly. The obtainment of carbon dioxide upon alkaline hydrolysis of I suggested a carbonate ester or a β-keto acid ester of III. The possibility of a carbonate ester was eliminated when the IR spectrum of I demonstrated the presence of two carbonyl groups, and when I absorbed only 2 mole of H_2 upon catalytic hydrogenolysis. Such catalytic hydrogenolysis at 25° and atmospheric pressure converted I into *o*-cresol-β-D-glucoside (VII), adipoin (VIII) and CO_2.

In an attempt to obtain Thieme's unidentified aliphatic hydroxy acid and to learn more concerning the nonsalicin portion of the unknown glucoside, I was treated with sodium methylate in absolute methanol The reaction mixture yielded III and a colourless oil which was identified as dimethyl 2-hydroxy-3-heptenedioate [IX, ($C_9H_{14}O_5$)] by mass, IR and NMR spectra. The mass spectrum gave the molecular ion, *m/e* 202, and a prominent peak at *m/e* 143 corresponding with the stable carbonium ion $CH_3OCOCH_2CH_2CH{=}CH^+(OH)$ formed by the loss of —CO—OCH_3 from the molecular ion. The NMR spectrum in $CDCl_3$ with TMS as the internal standard showed the following signals: 7·5 τ (multiplet, 4H, methylenes), 6·2 τ (singlet, 3H methyl) and 6·3 τ (singlet, 3H methyl), 5 τ (doublet, 1H, —CHOH—) and 4·2–4·6 τ (multiplet, 2H, vinyl). The latter multiplet in the vinyl region was simplified by irradiating the methylene region at 7·47 τ to yield a typical AB system influenced by the adjacent C—H from the —CHOH— at 5 τ. The IR spectrum confirmed the structure IX.

The dimethyl ester IX obtained by transesterification in methanol of I apparently is the dimethyl ester of the unidentified hydroxy acid obtained by Thieme[1] upon alkaline hydrolysis of the product he obtained by sodium methylate cleavage of I. (Found by Thieme: C, 48·3; H, 5·18. The free acid corresponding with IX, $C_7H_{10}O_5$ requires: C, 48·27; H, 5·79%.)

NMR spectra of I were determined in deuterated acetone with TMS as internal standard and in deuterium oxide with DSS as internal standard. The spectrum in deuterated acetone indicated 24 protons, in agreement with the analysis for $C_{20}H_{24}O_{10}$, assigned as follows: 4H in the aromatic region downfield between 2·6 τ and 3·0 τ are the protons of the salicyl alcohol moiety; 2 vinyl protons between 3·8 τ and 4·2 τ appear as a typical AB pair of doublets with distant coupling; 2 benzyl protons are a distinct AB pair of doublets at 4·6 τ and 4·7 τ; the single proton appearing as a doublet at 4·9 τ with a $J = 7{\cdot}4$ is the anomeric proton of glucose, consistent with the β form of the glucoside; the single alicyclic

[9] H. THIEME and R. RICHTER, *Die Pharmazie* **21**, 251 (1966).

hydroxyl proton appears as a broad triplet at 5·2 τ; the two C6 glucose protons are th strong singlet at 6·1 τ; the other 4 glucose C-H protons have a strong singlet at 6·5 τ; the 4 glucose hydroxyl protons are the strong singlet at 7·05 τ; and 4 methylene protons appear as a complicated multiplet between 7·2 τ and 7·5 τ. The 5 hydroxyl protons were not present in the spectrum determined in deuterium oxide. As in the case of the NMR spectrum of IX, irradiation of the methylene region at 7·47 τ caused decoupling in the vinyl region to yield a simple AB doublet and no change in the aromatic region, the benzyl proton responses, nor the anomeric proton response. Therefore, it appears that the vinyl protons of the original NMR spectrum of I were coupled only with methylene protons since the response due to the —CHOH— in the spectrum of IX is missing in the spectrum of I, and the alicyclic ring of I must contain the arrangement —CH=CHCH$_2$CH$_2$—. Attempts to decouple the methylene region by irradiation at 3·8 and 4·2 in the vinyl region produced almost no change. The response due to the —CHOH— proton noted in the spectrum of IX is missing entirely from the spectrum of I since this proton appears only when the alicyclic ring is cleaved by sodium methylate in methanol.

All the chemical reactions of salicortin are accounted for by the structure I as pictured in Fig. 1. The possible alternative structure in which the methylene groups and vinyl group in the alicyclic ring of I are interchanged was dismissed because such a conjugation of double bond and carbonyl should exhibit strong UV absorption at 227 nm. However, subtraction of the UV molecular absorption spectrum of VII from that of I indicated no such conjugation. In addition, the sodium methylate cleavage of the alternative structure would require an unlikely shift of a double bond in a stable conjugated system to produce the system of IX.

In their communication on II, Thieme and Richter[10] noted that treatment with dilute acid yielded salicyloyltremuloidin (X), and, therefore, concluded that II was a derivative of X. Because this reaction was so analogous to the formation of VI from I under essentially identical reaction conditions, we suspected that I and II possessed similar structures. Accordingly, II was submitted to hydrogenolysis under conditions identical with those used for I and absorbed 2 mole of H_2 to yield CO_2, VIII, and 2-*O*-benzoyl-*o*-cresol-β-D-glucoside (XI), identified by analysis and mass spectrum of its acetate. The hydrogenolysis results together with the analysis of II and the fact that dilute acid treatment of II yields X, the 2-*O*-benzoyl derivative of VI, prove that II is the 2-*O*-benzoyl derivative of I.

In Fig. 1 the reaction of I with boiling acid to yield VI is indicated. Thieme[1] also noted that dilute acid at room temp. yielded VI. Repetition of Thieme's experiment under controlled conditions demonstrated that in addition to VI, I also yielded CO_2 and IV. Thus, besides the dehydration reaction (Fig. 1), acid also causes hydrolysis of I to form III, and 1-hydroxy-6-oxo-2-cyclohexene-1-carboxylic acid, which in turn, eliminates CO_2 and dehydrogenates in the presence of air to give IV. Under conditions of boiling with stronger acid, both dehydration and hydrolysis reactions are accelerated, but under all conditions, the hydrolysis is much slower than the competing dehydration reaction of Fig. 1 to form VI. Furthermore, under strong acid conditions, the dihydropyrocatechol intermediate does not dehydrogenate to form IV.

During the course of these studies, many attempts were made to prepare the acetate of I, a compound reported by Thieme[1] to melt at 90–92°, and whose mass spectrum demonstrated the unsubstituted glucose moiety. Upon purification, all acetylated products proved

[10] I. A. PEARL and S. F. DARLING, *Arch. Biochem. Biophys.* **102,** 33 (1963).

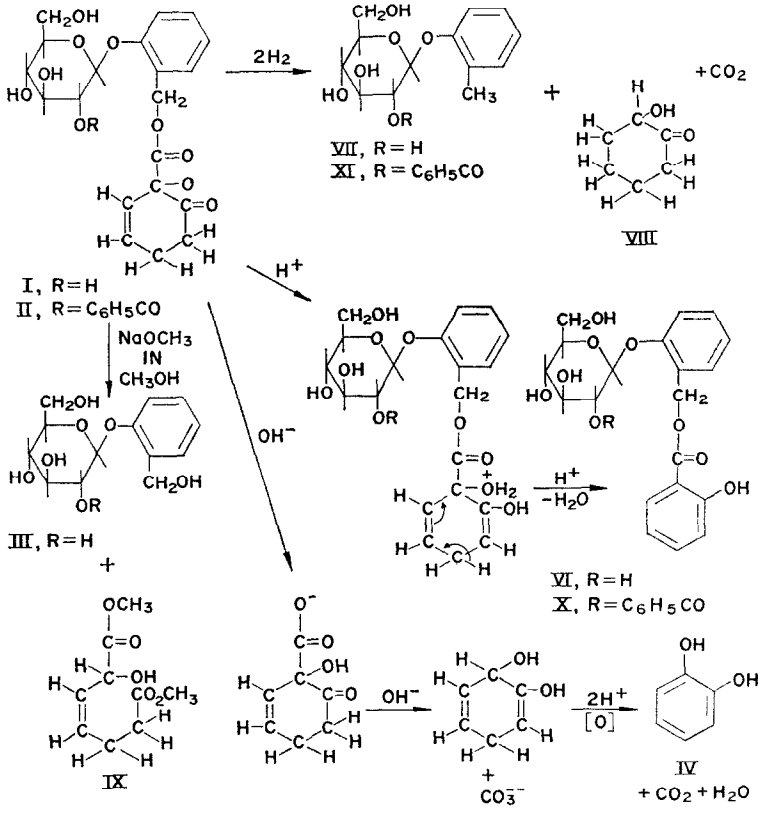

FIG. 1. REACTIONS OF SALICORTIN (I) AND TREMULACIN (II).

to be completely acetylated VI by mixed melting point and identity of IR and mass spectra with authentic material.[8,10]

In addition to being the precursors of III and tremuloidin (and/or populin), respectively, in organs of the Salicaceae, it is suggested that the alicyclic ring of these two labile glucosides, I and II, are also the precursors of IV and/or V, found many times in the past during the processing of these organs. Furthermore, it is possible that this type or similar type of labile alicyclic ring might be responsible for the isolation of IV, V, or related compounds from other plant materials.

EXPERIMENTAL

Salicortin (I)

Salicortin, isolated in quantity from the barks of *P. balsamifera*[3] and *S. purpurea*,[4] was recrystallized from EtOH to give colourless crystals, m.p. 135–137°, λ_{max}^{EtOH} 270, 214 (log ϵ 2·93, 3·75). ν_{max} (KBr pellet) 3380, 3310, 2920, 1740, 1604, 1590, 1492, 1458, 1439, 1412, 1381, 1352, 1321, 1290, 1260, 1240, 1219, 1180, 1132, 1074, 1046, 1011, 954, 919, 895, 866, 845, 806, 754, 684, 629, 590, 509, 471, 440, 415, 371, 348, 325, 291 and 288 cm^{-1}, identical with that of authentic I.[8] (Found: C, 56·61, 56·48; H, 5·70, 5·74. $C_{20}H_{24}O_{10}$ required: C, 56·60; H, 5·70%)

Acid Hydrolysis of I

(a) *With weak acid at room temp.* A solution of 0·5 g I in 10 ml 0·1 N HCl was allowed to stand 10 days at 22°. Filtration yielded 58·1 mg VI, and ether extraction gave a residue containing VI, IV, and a little salicyl alcohol by TLC. Extraction of the residue with petroleum ether gave crystals of IV, and GLC[11] indicated 19·5 mg IV.

(b) *With hot stronger acid.* A solution of 1·0 g I in 24 ml N HCl was boiled under reflux. After 1 hr, the solution was cooled and extracted with ether. The extract was fractionated with $NaHCO_3$ solution to give 143 mg V and 322 mg saliretin, the dimer of salicyl alcohol. In a similar experiment conducted in a quantitative CO_2 determination apparatus under a current of air, 35·2 mg of CO_2 were obtained. No trace of IV was found by TLC in either of the hot acid reactions.

Alkaline Hydrolysis of I

To a solution of 6·0 g I in 120 ml H_2O were added 24 ml N NaOH, and the solution was allowed to stand at 22° for 15 min during which time all I had disappeared as shown by TLC. The reaction mixture was just acidified with dil. HCl causing CO_2 evolution, and extracted with ether to give 0·66 g of ether extractives as a colorless oil. The oil was boiled 10 min under reflux with N NaOH, acidified, and extracted with ether. Removal of the ether left a dark semicrystalline residue of crude IV.

The aq. raffinate was exactly neutralized with $NaHCO_3$ solution and evaporated to a thick syrup. The syrup was covered with abs. EtOH and filtered to remove NaCl. The EtOH was concentrated and allowed to stand to yield 4·7 g of crude III, m.p. 195–198°. In a similar experiment on 1·0 g I in the CO_2 apparatus, 70·1 mg CO_2 was evolved. No trace of V was found by TLC in the products of either reaction.

Enzymic Hydrolysis of I

A clear solution of 0·25 g I in 10 ml 0·1 N NaOAc buffered to pH 5 with AcOH was treated with 50 mg β-glucosidase and allowed to stand overnight at 35°. The mixture, containing a heavy flocculent precipitate, was extracted with ether. Upon evaporation, the ether left a colorless oil. Glucose was identified in the aq. raffinate by paper chromatography.

Hydrogenolysis of I *to form* o-*Cresol-β-*D-*glucoside* (VII) *and Adipoin* (VIII)

A solution of 1·0 mmole (0·424 g) I in 10 ml abs. MeOH was treated with 40 mg 5% Pd on $BaSO_4$ catalyst and submitted to hydrogenolysis at 22° in an atmospheric pressure quantitative hydrogenation apparatus equipped with a tube of Ascarite between the hydrogenolysis flask and the hydrogen reservoir. Hydrogen uptake (2 mmole) was complete after 10 min. The reaction mixture was filtered, and the catalyst was washed with a little MeOH. The filtrate was evaporated to dryness and the residue was covered with anhydrous ether. The crystals were filtered, washed with anhydrous ether, and recrystallized from EtOAc to give 0·26 g VII, m.p. 161–162°, identical with the product prepared by hydrogenolysis of III by Richtmeyer[12] and containing the following bands in its IR spectrum: 3380, 2921, 2890, 1602, 1590, 1494, 1461, 1392, 1330, 1295, 1283, 1238, 1192, 1103, 1073, 1040, 993, 930, 895, 840, 772, 746, 712, 634, 565, 530, 515, 489, 472, 440, 427, 411 and 388 cm^{-1}. Its UV spectrum in EtOH contained maxima at 275, 269 and 214 nm (log ε 2·96, 3·02, 3·89).

Acetylation with pyridine and Ac_2O and recrystallization from EtOH gave o-*cresol-β-*D-*glucoside tetraacetate* as shiny needles m.p. 141–142°. Its IR spectrum contained bands at 3475, 2990, 2965, 2920, 2905, 2880, 1752, 1743, 1605, 1590, 1495, 1465, 1441, 1382, 1373, 1366, 1315, 1238, 1212, 1190, 1139, 1118, 1106, 1095, 1069, 1044, 990, 958, 929, 909, 894, 865, 840, 779, 759, 719, 701, 673, 639, 600, 560, 530, 510, 479, 445, 430 and 380 cm^{-1}, and its mass spectrum contained the following major and important *m/e* peaks with relative *m/e* intensity for each peak noted in parentheses: 43 (100), 81 (6·0), 97 (6·0), 108 (10·3), 109 (38·0), 127 (17·0), 139 (4·7), 145 (4·7), 169 (38·0), 187 (1·6), 203 (0·6), 211 (1·6), metastable 222 (corresponding with loss of CH_3CO as AcOH from tetraacetylglucoseoxonium ion), 229 (1·4), 271 (1·6), 287 (0·09), 288 (0·11), 311 (10·7), 345 (0·06), 365 (0·08), 379 (0·05) and M^+ 438 (0·007). (Found: mol. wt. by mass spectrometry, 438. $C_{21}H_{26}O_{10}$ required: mol. wt. 438.)

The ether filtrate from the VII crystals was evaporated to dryness and purified by preparative GLC to give VIII as a colorless oil identified by identity of its IR spectrum with that of authentic VIII purified by GLC.

An identical experiment was performed except that the tube of Ascarite was not employed. The same results were obtained, but the apparent uptake of H_2 was just 1 mmole because of the evolution of 1 mmole of CO_2 during the hydrogenolysis.

[11] I. A. Pearl and S. F. Darling, *Phytochem.* **7**, 1855 (1968).
[12] N. K. Richtmeyer, *J. Am. Chem. Soc.* **56**, 1633 (1934).

Reaction of I *with Sodium Methylate to form* III *and Dimethyl 2-hydroxy-3-heptenedioate* (IX)

To a solution of 1 0 g I in 10 ml abs. MeOH was added 4 ml of a solution of 0·2 g Na in 25 ml abs. MeOH. After standing at 22° for 10 min, the yellow solution was acidified slightly with anhydrous AcOH and evporated under reduced pressure The residue was diluted with 20 ml H_2O and extracted with ether. The aq. raffinate was evaporated to dryness, triturated with abs. EtOH, and filtered to give 0·48 g pure III, m.p. 197–198°. GLC of the ether extract indicated only one component The ether was evaporated to leave 0·4 g oil which was distilled at 15 mm to give IX as a colorless oil, λ_{max}^{EtOH} 300 shoulder, 277, 213 (log ϵ 0·98, 1·61, 3·08). ν_{max} (film) 3470, 3020, 2960, 2850, 2070, 2000, 1995, 1735, 1439, 1364, 1255, 1204, 1169, 1080, 1017, 906, 879, 855, 834 and 800 cm^{-1}. (Found: mol. wt. by mass spectrometry, 202. $C_9H_{14}O_5$ required: mol. wt. 202.)

Tremulacin (II)

Tremulacin was isolated in quantity from the leaf stem twigs of diploid *P. tremuloides*,[5] recrystallized from EtOAc containing a little petroleum (b.p. 30–60°), and obtained as colorless crystals, m.p. 122–123°, ν_{max} (KBr pellet) 3425, 3090, 3020, 2965, 2920, 2890, 1725, 1602, 1494, 1452, 1395, 1360, 1318, 1272, 1180, 1072, 1028, 1010, 981, 930, 900, 850, 804, 760, 704, 689, 634, 599, 535 and 529 cm^{-1}. (Found: C, 61·16, 61·09; H, 5·47, 5·45. $C_{27}H_{28}O_{11}$ required: C, 61·36; H, 5·34%.)

Hydrogenolysis of II *to form* VIII *and* 2-O-*Benzoyl-*o*-cresol-*β*-*D*-glucoside* (XI)

II (0·54 g, 1·0 mmole) was subjected to hydrogenolysis and processed exactly as described for I. Again, 2 mmole H_2 were absorbed, and CO_2 was liberated. Upon triturating the evaporated reaction mixture with anhydrous ether and filtering, XI was obtained as 0·31 g of colorless crystals which were recrystallized from EtOAc to give crystals, m.p. 190–193°.

Acetylation with Ac_2O and pyridine gave 2-O-*benzoyl-*o*-cresol-*β*-*D*-glucoside triacetate* as colorless needles (dil. EtOH), m.p. 138–140°. Its mass spectrum contained the following major and important *m/e* peaks: 43 (24·9), 77 (8·5), 105 (100), 109 (8·5), 169 (9·8), 231 (1·0), 273 (0 4), metastable 282·5 (corresponding with loss of CH_3CO as AcOH from benzoyltriacetylglucoseoxonium ion), 291 (0·2), 333 (0·6), 351 (0·06), 393 (6·6), 427 (< 0·1) and 441 (< 0·1). The molecular ion at 500 was missing. (Found: C, 62 39, 62·17; H, 5·59, 5·53. $C_{26}H_{28}O_{10}$ required: C, 62·39; H, 5·64%.)

Acknowledgements—We are indebted to Dr. S. D. Darling of Southern Illinois University who suggested the quantitative hydrogenolysis experiments and also performed the decoupling NMR measurements on I and IX. We also thank the following of The Institute of Paper Chemistry staff: Mrs. C. Robbins for the preparation in quantity of I and II, Mr. L. Sell for UV, IR and NMR spectra, and Dr. D. Johnson for help in the interpretation of NMR spectra. Mass spectra were determined by Morgan-Schaffer Corp., Montreal, Canada, and analyses were made by Micro-Tech Laboratories, Skokie, Illinois.