Tetrahedron Letters No.11, pp. 1193–1204, 1966. Pergamon Press Ltd.
Printed in Great Britain.

THE CONSTITUTION OF THEAFLAVIN

A. G. Brown, C. P. Falshaw, E. Haslam, A. Holmes, and W. D. Ollis

Department of Chemistry, The University, Sheffield

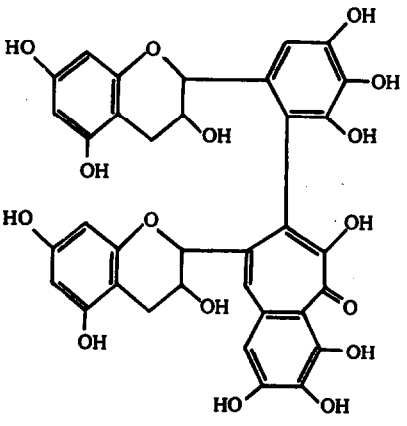
(Received 20 January 1966)

The recent publication[1] describing (a) the determination of the molecular weight of theaflavin acetate, (b) the deuterium analysis of theaflavin deuteroacetate, and (c) the interpretation of the NMR spectrum of the theaflavin–methanol solvate, makes it desirable for us to report our evidence regarding the constitution of theaflavin. Our results are in certain respects more extensive and more detailed than those reported by Gianturco et al.[1] and in particular we have prepared a number of new derivatives* of theaflavin which have been useful in defining its novel constitution. The configurational formula which we now propose for theaflavin is different from that given by Gianturco et al.[1]

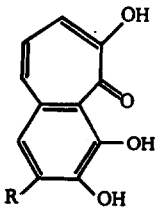
Crystalline theaflavin was first obtained by Roberts and Myers[2] and later

---

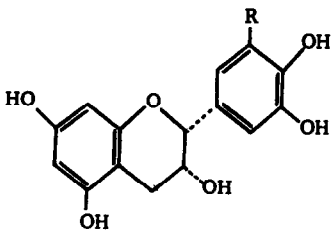
* See Table 1



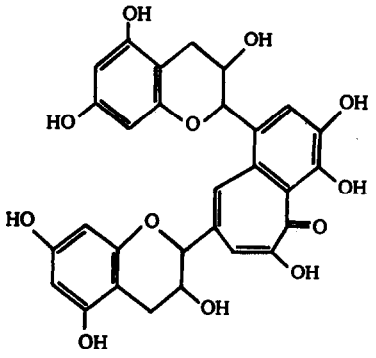
I



II; R = OH
III; R = H



IV; R = OH
V; R = H



VI

TABLE 1

	Molecular formula	Melting point		Infrared Spectra				
Theaflavin (VII)	$C_{29}H_{24}O_{12}\cdot 2\cdot 5H_2O$	237° (decomp.)	(Nujol)	3300 – 3100		1630		1605
Theaflavin nonacetate (VIII)	$C_{29}H_{15}O_3(OAc)_9$	167-168°	($CHCl_3$)	1775	1755	1675	1627	1595
Theaflavin heptamethyl ether (IX)	$C_{29}H_{15}O_3(OH)_2(OMe)_7$	-	($CHCl_3$)	3575		1675	1625	1598
Theaflavin heptamethyl ether diacetate (X)	$C_{29}H_{15}O_3(OAc)_2(OMe)_7$	186-187°	($CHCl_3$)		1740	1680	1620	1598
Theaflavin heptamethyl ether dimesylate (XI)	$C_{27}H_{15}O_3(OMes)_2(OMe)_7$	155-157°	($CHCl_3$)			1675	1630	1603

These compounds have been fully characterised by analysis and mass spectrometry and their UV, IR, and NMR spectra are in full accord with the indicated constitutional formulae. The details will be reported in our full publication.

Roberts[3] proposed for theaflavin the constitution (I). This suggestion was based mainly on the supposition that the enzyme catalysed (tea oxidase) formation of theaflavin gallate involved the oxidative coupling[4] of catechin type precursors and the opinion that the UV spectrum of theaflavin indicated the presence of a benzotropolone nucleus of the purpurogallin (II) type. Our initial structural studies on theaflavin were guided by these hypotheses, but it soon became clear that modification of the constitution (I) was necessary. Meanwhile, Takino and Imagawa[5] made the important observation that oxidation of mixtures of (-)epigallocatechin (IV) and (-)epicatechin (V) either with tea oxidase or with potassium ferricyanide yielded a reddish-orange compound, $C_{29}H_{24}O_{12}$. The identity of this compound, $C_{29}H_{24}O_{12}$, with theaflavin was suspected and on the basis of a reasonable mechanism for its formation the constitution (VI) was considered.[5,6] Later the identity of theaflavin with the oxidation product, $C_{29}H_{24}O_{12}$, was established.[7]

While these studies by Takino et al.[5,6,7] were in progress, we were examining natural theaflavin isolated from black tea by a modification of Roberts' procedure.[2] Theaflavin was characterised by the derivatives listed in Table 1 and these results established that theaflavin had the molecular formula $C_{29}H_{15}O_3(OH)_9$, in which seven hydroxyl groups were phenolic or enolic and the other two hydroxyl groups were alcoholic. The infrared spectral data (see Table 1) indicated that theaflavin contained a highly conjugated carbonyl group (ν_{CO} 1630 cm^{-1}.) and the shift of the carbonyl band (ν_{CO} 1675-1680 cm^{-1}.) in appropriate derivatives suggested that this carbonyl group was internally hydrogen

bonded in theaflavin. This observation was obviously compatible with Roberts' suggestion[3] that theaflavin was a benzotropolone derivative. However, comparison of the UV spectra of theaflavin and its derivatives with appropriate models (see Table 2) clearly favoured the presence of a 1', 2'-dihydroxy-3, 4-benzotropolone (III)[8, 9] grouping rather than a 1', 2', 3'-trihydroxy-3, 4-benzotropolone (II)[10] residue.

The presence of a 1', 2'-dihydroxy-3, 4-benzotropolone grouping in theaflavin and its molecular formula, $C_{29}H_{24}O_{12}$, was clearly compatible with its formation by an oxidative process

$$C_{15}H_{14}O_6 + C_{15}H_{14}O_7 \xrightarrow{(-CH_4O)} C_{29}H_{24}O_{12},$$

involving precursors of the catechin, $C_{15}H_{14}O_6$, and gallocatechin, $C_{15}H_{14}O_7$ types. This process was clearly analogous to the oxidative coupling of catechol and pyrogallol to yield 1', 2'-dihydroxy-3, 4-benzotropolone,[8, 9]

$$C_6H_6O_2 + C_6H_6O_3 \xrightarrow{(-CH_4O)} C_{11}H_8O_4$$

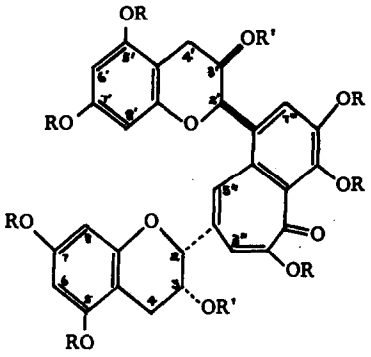
and as both (-)epicatechin (V) and (-)epigallocatechin (IV) occur in green tea,[3] their oxidative coupling to yield theaflavin (I)[5, 6, 7] has ample precedent.[8, 9] We now propose for theaflavin the configurational formula (VII).

TABLE 2

Ultraviolet Spectra, λ_{max} (ε_{max}) in ethanol

Theaflavin (VII)	216	(35, 500)	229*	(25, 100)	271	(19, 500)	290*	(17, 400)	384	(8, 700)	470 (3, 600)
Theaflavin + EtOH−$AlCl_3$	218	(40, 700)	230*	(33, 100)	278	(26, 900)	318	(26, 900)	406	(13, 500)	544 (5, 750)
Theaflavin nonacetate (VIII)	211	(25, 100)	221*	(20, 000)	250	(13, 200)	314	(7, 100)	353	(4, 700)	
Theaflavin heptamethyl ether (IX)	213	(52, 500)	222*	(31, 600)	269	(18, 200)	315	(8, 900)	384	(2, 950)	
Theaflavin heptamethyl ether diacetate (X)	211	(67, 600)	225	(33, 900)	268	(20, 000)	315	(11, 000)	375-385	(3, 250)	
Theaflavin heptamethyl ether dimesylate (XI)	207	(93, 000)	240*	(28, 800)	268	(17, 000)	318	(8, 900)	380	(4, 900)	
1', 2'-Dihydroxy-3, 4-benzotropolone (V)	230-235	(12, 000)	265	(21, 400)	276	(21, 900)			376	(10, 700)	460 (3, 700)
1', 2'-Dihydroxy-3, 4-benzotropolone + EtOH−$AlCl_3$			235	(24, 600)	273	(27, 500)	300	(27, 500)	398	(12, 000)	532 (3, 250)
1', 2'-Dihydroxy-3, 4-benzotropolone triacetate			233	(24, 000)	248*	(19, 100)	316	(9, 100)	340*	(6, 300)	
1', 2'-Dihydroxy-3, 4-benzotropolone trimethyl ether	210	(8, 300)			267	(12, 600)	328	(1, 700)	387	(2, 500)	

* = shoulder



(VII) R = R' = H
(VIII) R = R' = Ac
(IX) R = Me; R' = H
(X) R = Me; R' = Ac
(XI) R = Me; R' = Mes
(XII) R = R' = $SiMe_3$

In the derivation of the configurational formula (VII) the assumption is made that the formation of theaflavin takes place with the preservation of absolute configuration at the chiral centres in the precursors. This assumption is reasonable, but in view of the known epimerisation at C_2 in the catechin series the possible consequences of this during the formation of theaflavin (VII) from (-)epicatechin (V) and (-)epigallocatechin (IV) must be considered. In this connection one facet of the chemistry of theaflavin which has not been previously examined is whether it is optically active. We have now established that theaflavin and its derivatives all give optical rotatory dispersion curves. These will be discussed in our full publication.

The formula (VII) for theaflavin may also be deduced by an independent agreement based on the NMR and mass spectra of theaflavin and its derivatives. The NMR spectra are summarised in Table 3. The presence of two pairs of meta-related protons (positions 6, 8, 6', and 8') are recognisable as two AB systems ($J \sim 2$ c/s) and singlets may be assigned to three protons (positions 3",

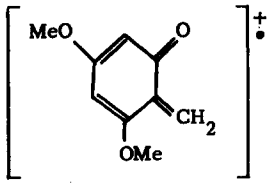
TABLE 3

	Positions of Protons												Groups
	2	2'	3	3'	CH_2 groups at 4 and 4'	6	8	6'	8'	7"	3"	5"	
Theaflavin (VII) in CD_3COCD_3	4.95	4.25	5.35-5.65 m		—7.08— m	Two coincident AB systems 3.92 and 3.97 (J = 2 c/sec)				2.45	1.95	2.05	-5.0 (OH)* 1.6—2.2 $(OH)_6$* 5.7—6.3 $(OH)_2$*
Theaflavin nonacetate (VIII) in $CDCl_3$	5.22	4.45	——4.65 m		—7.04— m	3.10	——	——	3.45 m	2.12	2.82		7.66-7.80 $(OAc)_7$ 8.05 (OAc) 8.13 (OAc)
Theaflavin heptamethyl ether (IX) in $CDCl_3$	5.35	4.49	5.45-5.8 m		—7.12— m	3.73	——	——	4.05 m	2.30	2.79		8.09 $(OH)_2$* 6.20-6.27 $(OMe)_5$ 6.05 (OMe) 6.01 (OMe)
Theaflavin heptamethyl ether diacetate (X) in $CDCl_3$	5.24	4.35	——4.62 m		—7.04— m	3.70	——	——	4.0 m	2.40	2.90		6.0—6.18 $(OMe)_7$ 8.06 (OAc) 8.11 (OAc)
Theaflavin heptamethyl ether dimesylate (XI) in $CDCl_3$	5.23	4.40	——4.83 m		—6 88— m	3.74	——	——	3.90 m	2.33	2.85		6.16-6.26 $(OMe)_5$ 6.05 (OMe) 6.00 (OMe) 7.55 (OMes) 7.23 (OMes)
Theaflavin trimethylsilyl ether (XII) in CCl_4	5.45	4.59	5.50-5.95 m		—7.22— m	3.85 m	——	——	4.25	3.79	2.44	2.78	

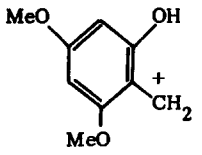
The chemical shifts are given on the τ scale.
m = multiplet. Either the range of the multiplet or the centre of broad multiplets is quoted.
Unless otherwise indicated, the signals described in this Table are singlets.
* These assignments to hydroxyl groups were confirmed by deuteration studies.

5", and 7") on the benzotropolone residue. Broad signals associated with four protons may be assigned to two benzylic methylene groups (positions 4 and 4'). The benzylic protons (positions 2 and 2'), when they are observable, appear as singlets. This is good evidence[11] for the relative stereochemistry shown in the configurational formula (VII) because the singlet nature of the signals due to the protons (positions 2 and 2') requires axial–equatorial relations, that is cis-stereochemistry between C_2–H and C_3–H and between $C_{2'}$–H and $C_{3'}$–H. The signals due to the protons (positions 3 and 3') are broad, but they show the expected downfield shift on O-acetylation and O-mesylation.

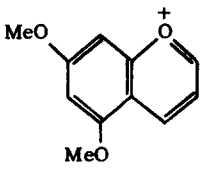
The mass spectra of theaflavin nonacetate (VIII), theaflavin heptamethyl ether (IX), theaflavin heptamethyl ether diacetate (X), and theaflavin heptamethyl ether dimesylate (XII) have been determined. Detailed discussion will be presented shortly, but it may be noted that the results provide excellent support for the proposed structure (VII) for theaflavin. In particular, generation of ions of the type (XIII), (XIV), and (XV) clearly indicates the presence of the grouping (XVI) and the fragmentation pattern associated with retro-Diels-Alder reactions[12] (see arrows, XVI) requires the presence of two such groupings (XVI). The mass spectrum of theaflavin heptamethyl ether exemplifies this scheme in that it shows a molecular ion peak at m/e 662, fragment ions at m/e 496, 495, 330, 329, 191, 167, and 166, and metastable peaks due to the processes $662^+ \rightarrow 496^+ + 166$ and $496^+ \rightarrow 330^+ + 166$.



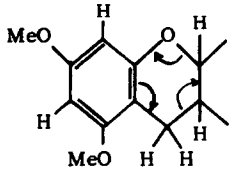
XIII m/e 166



XIV m/e 167



XV m/e 191



XVI

Thus the mass spectral evidence requires the presence of two groupings (XVI) and when these are linked to the benzotropolone (III) in positions defined by the NMR spectral data then this leads directly to the constitution (VI) and the configuration (VII) for theaflavin.[13]

The description of theaflavin as a natural product is debatable since its formation involved an enzyme catalysed process which operates during the fermentation of green tea to give black tea. However, its novel C_{29}-structure is certainly derived by the oxidative coupling of C_{15}-flavonoid precursors and as such its structure may be regarded as a new variant upon the C_{30}-structures characteristic of the biflavonyls[14] and the dimeric proanthocyanidins.[15]

This work is a continuation of the studies of the late Dr. E.A.H. Roberts. We thank the Indian Tea Association (London) for financial support and Professor R.D. Haworth, F.R.S., for his close interest.

REFERENCES

(1) Y. Takino, A. Ferretti, V. Flanagan, M. Gianturco, and M. Vogel, Tetrahedron Letters, No. 45, 4019 (1965).

(2) E.A.H. Roberts and M. Myers, J.Sci.Fd.Agric., 10, 172, 176 (1959).

(3) E.A.H. Roberts, "The Chemistry of Flavonoid Compounds", ed. T.A. Geissman, Macmillan, New York, 1962, p. 468.

(4) A.I. Scott, Quart.Revs., 19, 1 (1965) and references there cited.

(5) Y. Takino and H. Imagawa, Agric.Biol.Chem.(Japan), 27, 319 (1963).

(6) Y. Takino, H. Imagawa, H. Horikawa, and A. Tanaka, Agric.Biol.Chem.(Japan), 28, 64 (1964).

(7) Y. Takino and H. Imagawa, Agric.Biol.Chem.(Japan), 28, 255 (1964).

(8) M. Murakami, K. Suzuki, and E. Mishima, J.Chem.Soc.Japan, 75, 620 (1954).

(9) L. Horner, W. Dürckheimer, K.H. Weber, and K. Dölling, Chem.Ber., 97, 312 (1964).

(10) P.L. Pauson, Chem.Revs., 55, 9 (1955).

(11) J.W. Clark-Lewis, L.M. Jackman, and T.M. Spotswood, Aust.J.Chem., 17, 632 (1964).

(12) R.I. Reed and J.M. Wilson, J.Chem.Soc., 5949 (1963).

B. Willhalm, A.F. Thomas, and F. Gautschi, Tetrahedron, 20, 1185 (1964).

A.P. Johnson, A. Pelter, and M. Barber, Tetrahedron Letters, No. 20, 1267 (1964).

(13) This constitutional proposal for theaflavin has been reported by Dr. D.S. Bendall in a lecture on the "Biochemistry of Tea Fermentation", British Association Meeting, Cambridge, September 1965.

(14) W. Baker and W.D. Ollis in "Recent Developments in the Chemistry of Natural Phenolic Compounds", ed. W.D. Ollis, Pergamon Press, London, 1961, p. 152.

W. Baker, A.C.M. Finch, W.D. Ollis, and K.W. Robinson, J.Chem.Soc., 1477 (1963).

(15) W.G.C. Forsyth and J.B. Roberts, Biochem.J., 74, 374 (1960).

K. Freudenberg and K. Weinges, Tetrahedron Letters, No. 8, 267 (1961); Angew.Chem., 74, 182 (1962); "The Chemistry of Flavonoid Compounds", ed. T.A. Geissman, Macmillan, New York, 1962, p. 197.

K. Weinges, Chem.Ber., 94, 3032 (1961).

K. Weinges and K. Freudenberg, Chem.Comm., 220 (1965).

T.A. Geissman and H.F.K. Dittmar, Phytochemistry, 4, 359 (1965).