Pergamon

PII: S0031-9422(97)00508-6

Phytochemistry, Vol. 46, No. 8, pp. 1397–1402, 1997

Printed in Great Britain
0031-9422/97 $17.00+0.00

A POLYPHENOLIC PIGMENT FROM BLACK TEA

ADRIENNE L. DAVIS, JOHN R. LEWIS, YA CAI,* CHRIS POWELL, ALAN P. DAVIS, JOHN P. G. WILKINS, PAUL PUDNEY and MIKE N. CLIFFORD†

Unilever Research, Colworth Laboratory, Sharnbrook, Bedford MK44 1LQ, U.K.; † University of Surrey, School of Biological Sciences, Guildford, Surrey, GU2 5XH, U.K.

(*Received in revised form* 25 *April* 1997)

Key Word Index—*Camellia sinensis*; Theaceae; black tea; tea polyphenols; flavanols; catechins; theaflavins; thearubigins; theacitrins; NMR.

Abstract—A yellow polyphenolic compound has been isolated from black tea (fermented leaves of *Camellia sinensis*) and its structure characterized using a number of spectroscopic techniques. This compound, theacitrin A, represents a new class of polyphenolic pigments in black tea. © 1997 Elsevier Science Ltd

INTRODUCTION

The characteristic colour of black tea is generated during its manufacturing process. During this process, the colourless catechins which are abundant in fresh leaves of *Camellia sinensis* ($\leqslant$35% dry weight, [1]) are oxidized both enzymatically and chemically to give two major groups of pigments, theaflavins and thearubigins [2]. The theaflavins are formed by oxidative coupling of the dihydroxybenzene and trihydroxybenzene rings of an appropriate pair of flavan-3-ols, (e.g. epicatechin and epigallocatechin, respectively), resulting in a benzotropolone ring, which gives a yellow colour. The major theaflavin compounds have been identified [3]. By contrast, the thearubigins are an extremely complex, heterogenous mixture of pigments, and their structures are largely unknown [4]. Thearubigins are predominant in black tea leaf (15–20% dry weight) and infusions (20–35% of soluble solids) and are believed to make the greatest contribution to the colour of black tea liquor.

Recently, great effort has been made to identify polyphenolic compounds present in teas. A large number of polyphenolic monomers (catechins) and dimers (theaflavins) have been isolated and characterized [5, 6], including several novel theaflavin compounds [7]. From the thearubigin fractions of an Assam black tea, three yellow compounds have been isolated [8] and preliminary structural data reported [9, 10]. However, these compounds were found to be highly unstable and attempts to elucidate their structures were not successful. In the current study, these yellow compounds have been re-isolated from the same black tea. One of them, termed theacitrin A, has been characterized unequivocally as a novel polyphenolic compound. Spectroscopic evidence suggests that the other two (theacitrins B and C) have a similar structural skeleton. Details of their separation, purification and characterization are described below.

RESULTS AND DISCUSSION

Theacitrin A (**1**) was isolated from decaffeinated black tea leaves using a combination of column chromatography and semi-preparative HPLC methods. This yellow compound was found to be highly soluble in water and methanol, but less soluble in acetone. It is not stable in solution, particularly in the presence of acid, where it undergoes rapid structural changes. Consequently, it is very difficult to prepare this compound to the purity at which good quality structural information can be generated.

The negative ion electrospray mass spectrum indicated that the M_r of theacitrin A is 760 Daltons and the ^{13}C-^1H NMR spectrum showed the presence of 37 carbon atoms. Therefore, the empirical formula was provisionally deduced as $C_{37}H_{28}O_{18}$.

The ^1H NMR spectrum (Fig. 1 and Table 1) clearly indicated the presence of two flavan-3-ol units, part of which has been transformed. Two sets of resonances from the heterocyclic C-rings are readily identifiable (H-2, H-3, H-4α, H-4β and H-2′, H-3′, H-4′α, H-4′β) and a group of aromatic resonances is observed between 6.0 and 6.4 ppm, attributable to the two A-rings in the molecule (two pairs of doublets each exhibiting a 2.3 Hz coupling consistent with *meta*-substitution). The coupling pattern of H-2 and H-2′ indicates that both transformed flavan-3-ol units retain epicatechin-like stereochemistry [11]. The pres-

* Author to whom correspondence should be addressed.

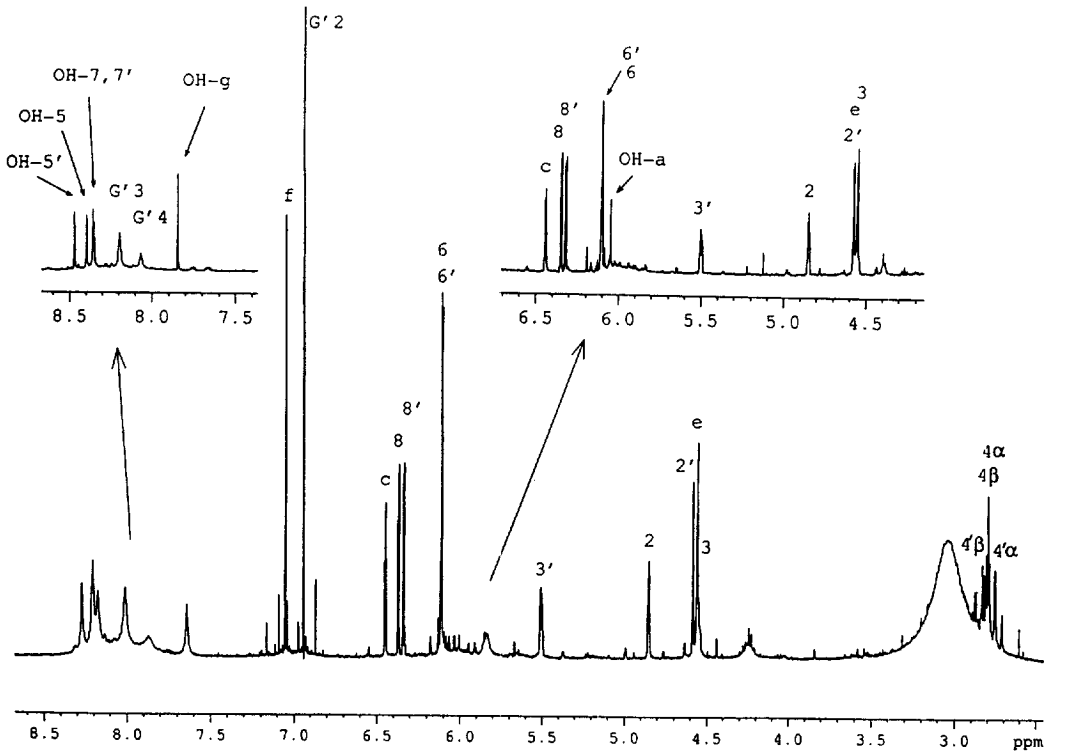
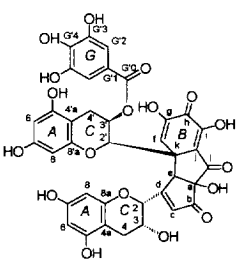


Fig. 1. ^{1}H NMR spectrum of theacitrin A in acetone-d_6 at 303 K (main spectrum). Expansions show the positions of the OH peaks (run at 283 K).

ence of a two-proton singlet at 6.94 ppm indicates that one of the flavan-3-ol units possesses a galloyl ester group (attached at C-3′ as deduced from HMBC connectivities, see below). The resonances for the structural fragments listed above are observed in the ^{13}C-^{1}H NMR spectrum (Table 2), confirming these assignments.



Theacitrin A (**1**)

Subtracting the identified fragments from the proposed empirical formula of theacitrin A leaves a fragment of $C_{12}H_6O_6$, which is presumably derived from the two B rings of the catechin precursors. Two dimensional NMR spectroscopy, specifically DQF-COSY [12], HMQC [13] and HMBC [14] experiments, was conducted in order to characterize this fragment. The structurally diagnostic long-range proton carbon correlations extracted from the HMBC spectrum are listed in Table 3.

Three non-labile resonances are so far unidentified in the ^{1}H NMR spectrum H-f (δ 7.05), H-c (δ 6.45) and H-e (δ 4.56). Examination of the chemical shifts of these protons and their attached carbons shows that H-f and H-c are olefinic (or aromatic) but H-e is not. H-e is coupled to H-2, whilst H-c is coupled to H-2 and to H-e. The magnitude of the coupling constants is *ca* 1.4 Hz in each case, strongly suggesting that the coupling from H-c to H-2 and H-e is allylic in nature and that C-2 and C-e are attached to a quaternary olefinic carbon that participates with C-c in a double bond. These conclusions are confirmed, and the quaternary olefinic carbon is identified as C-d, by the long-range correlations observed in the HMBC spectrum (Table 3).

The chemical shifts for C-c (δ 131.56) and C-d (δ 174.84) demonstrate that the C-c to C-d bond is polarised. Examination of the long-range correlations from H-c show that this is due to the presence of a ketone substituent, since H-c (and H-2, very weakly) has a correlation to C-b (δ 200.3 ppm), a ketonic carbonyl. Correlations from both H-e and H-c to C-a (δ 87.57) indicate that a ring system must be present in this part of the molecule, though at this stage in the analysis it is unknown if it is a five- or six-membered. The chemical shift of C-a suggests the presence of at least one oxy-containing substituent on C-a (low temperature experiments confirm the presence of an OH substituent, see below). The above assignments and the data in Tables 1–3 clearly demonstrate that the unknown molecular fragment contains three carbonyl groups and three double bonds; therefore, three rings must be present in this part of the structure.

Table 1. ^{1}H NMR data for theacitrin A*

Proton	Chemical shift (δ)	Coupling constant (J)
2	4.852 *q*	$J(2, 3) = 1.2$; $J(2, e) = 1.4$; $J(2, c) = 1.4$
3	4.552 *m*	$J(3, 4\beta) = 4.4$; $J(3, 4\alpha) = 4.4$
4β†	2.828 *dd*	$J(4\alpha, 4\beta) = 16.9$
4α†	2.774 *dd*	
6	6.110 *d*	$J(6, 8) = 2.3$
8	6.367 *d*	
2′	4.588 *s br*	
3′	5.504 *dd*	$J(3', 4'\beta) = 4.4$; $J(3', 4'\alpha) = 2.0$
4′β	2.852 *dd*	$J(4'\alpha, 4'\beta) = 17.4$
4′α	2.735 *dt*	$J(2', 4'\alpha) = 1.4$
6′	6.110 *d*	$J(6', 8') = 2.3$
8′	6.336 *d*	
e	4.562 *t*	$J(c, e) = 1.4$
c	6.446 *t*	
f	7.051 *s*	
G′2	6.940 *s*	
OH-5′‡	8.47 *s*	
OH-5‡	8.40 *s*	
OH-7,7′‡	8.36 *s br*	
OH-G′3‡	8.20 *s br*	
OH-G′5‡	8.10 *s br*	
OH-a‡	6.05 *s*	
OH-g‡	7.85 *s*	
OH-i§	10.89 *s*	

* Measured at 303 K in acetone-d_6 solvent. Chemical shifts are quoted in ppm relative to the solvent peak = 2.052 ppm. Coupling constants are in Hz.

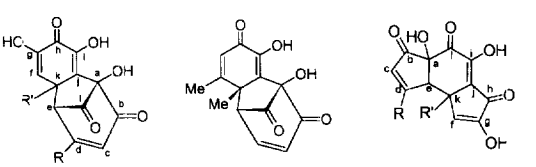
† Assignment interchangeable.

‡ Measured in 283 K.

§ Measured at 183 K.

Another part of this fragment is identified from the observation that one of the broad, labile proton resonances, for OH-g (δ 7.85), has long-range correlations to C-f (δ 120.42), C-g (149.47) and a carbonyl carbon, C-h (δ 178.05). This is clearly an $\alpha\beta$-unsaturated carbonyl group with an hydroxyl substituent at the α-position. The chemical shifts of C-f and C-g are consistent with this conclusion, since the hydroxyl group polarizes the double bond more strongly than the carbonyl group and in the opposite direction. It is now possible to link together all parts of the unknown fragment so far identified by further examination of the data in Table 3: H-2′ correlates to C-k, C-e and C-f, whilst H-f and H-e also correlate to C-k. Therefore, the quaternary carbon C-k must link C-e, C-f and C-2′ as shown (**1**).

This leaves three carbons unaccounted for, C-j (δ 120.42), C-i (δ 150.47) and C-l (δ 196.28), and three rings to construct. Both H-e and H-f correlate strongly to C-j and, consequently, a bond must be present between C-k (δ 51.29) and C-j. C-j is an olefinic carbon and must therefore be attached to C-i, the only remaining olefinic carbon. The chemical shift of the carbonyl carbon C-h can be satisfied by attaching it to C-i, thereby making it doubly $\alpha\beta$-unsaturated, closing one ring. In order to make the chemical shifts of C-j and C-i correct, it is necessary to attach an hydroxyl substituent to C-i. There now remains only an OH group and a ketonic carbonyl group to account for. Both groups must be attached to C-a in order to satisfy its valency requirements. Therefore, in order to make an unequivocal structure with three rings, one question remains—does the carbonyl group bridge between C-a and C-j [with a bond between C-a and C-e, (**1**)] or between C-e and C-a [with a bond present between C-a and C-j, (**2**)]?



(2) (3) (4)

Structure 2.

Structure (**2**) has a carbon skeleton identical to a compound (**3**) which was synthesized chemically by Dürckheimer *et al.* from cyclocondensation of 4,5 dimethyl-*O*-benzoquinone and 3-hydroxy-*O*-benzo-

Table 2. ^{13}C NMR data for theacitrin A*

Carbon	Chemical shift (δ)
2	76.71
3	64.45
4	28.48
4a	100.29
5	157.67†
6	97.25‡
7	157.78†
8	96.35
8a	155.80
2′	82.42
3′	64.14
4′	28.83
4′a	99.44
5′	157.81†
6′	97.16‡
7′	157.98†
8′	96.30
8′a	157.29
a	87.57
b	200.26
c	131.56
d	174.84
e	58.22
f	120.18
g	149.47
h	178.05
i	150.47
j	120.42
k	51.29
l	196.28
G′0	165.87
G′1	121.48
G′2	110.36
G′3	145.68
G′4	139.21

* Measured at 303 K in acetone-d_6. Chemical shifts are quoted in ppm relative to the solvent peak = 29.80 ppm.
†,‡ Assignments interchangeable.

Table 3. Diagnostic long-range proton-carbon correlations of theacitrin A measured using HMBC

Proton	Carbon
OH-a	a†‡, e†‡, l§
OH-g	f, g, h
f	e, 2′†, J, g, h, l†
c	e, 2, a, d, b
e	k, 2′, a, j, c, d, l
3′	4′a, G′O†
2′	4′, k, e, 3′, f
2	4, c, 8a, d, b†

* Measured at 303 K in acetone-d_6 unless otherwise stated.
† Weak correlations at 303 K.
‡ Confirmed by measurement at 283 K.
§ Observable only at 283 K from summation of several columns of the spectrum.

quinone and characterized by X-ray crystallography [15]. Consequently, despite the fact that its ring systems are rather strained, structure (**2**) is plausible. However, (**3**) has a CO stretch at 1780 cm^{-1} [15] assigned to the bridging carbonyl group [ν(CO) = 1780, 1690 and 1650 cm^{-1} (KBr)], whilst theacitrin A has no analogous CO stretch [ν(CO) = 1707, 1657(sh), 1629 and 1609 cm^{-1} (KBr)]. This suggests that structure (**2**) is not correct.

This conclusion is confirmed by the data in Table 3. A D_2O-exchangeable proton resonance is observed to correlate to C-a and (more strongly) to C-e, demonstrating that the resonance must be OH-a. An extremely weak correlation from OH-a to C-l was also observed. It is known that for any given proton, the intensity of its correlations in the HMBC experiment are proportional to the magnitude of ^nJ(CH) [14]. Therefore, (**2**) cannot be the correct structure of theacitrin A, since it would require the four bond C-e to OH-a coupling constant to be larger in magnitude than both the three bond C-l to OH-a coupling and the two bond C-a to OH-a coupling, which is very unlikely [16].

The final structural variation that has not been considered, but also fits all the data discussed so far, is shown as (**4**). Here, C-j rather than C-i is adjacent to C-h. This structure is ruled out by the observation of a weak long-range correlation from H-f to C-l and the absence of any correlation from H-f to C-i. Consequently, we conclude that theacitrin A must have structure (**1**). The observation of the four-bond correlation from H-f to C-l is consistent with this structure, since it is believed that long-range couplings transmitted over four bonds are favoured by a 'W'-type geometry [16].

All of the hydroxyl proton resonances of theacitrin A are visible in a ^1H NMR spectrum acquired at 283 K (see Fig. 1), with the exception of OH-i, which is not observable. In order to locate this resonance, low temperature spectra were acquired. At a temperature of 193 K, an additional broad resonance was observed at *ca* 10.8 ppm and when the temperature was further decreased to 183 K, the line-width of this peak was observed to reduce significantly. We assign this resonance as OH-i (δ 10.89). Based upon the spectroscopic evidence, the structure for theacitrin A is unequivocally established as (**1**).

Two compounds with UV-Vis spectra very similar to theacitrin A have also been isolated from the same black tea, but in much smaller quantities. One of these compounds (termed theacitrin C) elutes at 18.27 min under the HPLC analytical conditions as described below, and has a M_r of 912 Daltons as determined by mass spectrometry, suggesting one galloyl ester group more than theacitrin A. Treatment of theacitrin C with tannase yielded theacitrin A amongst the reaction products. Therefore, theacitrin C is believed to have the same structure as theacitrin A, with an additional galloyl ester group at C-3. The other compound (theacitrin B) elutes at 11.93 min, and has the same UV/Vis

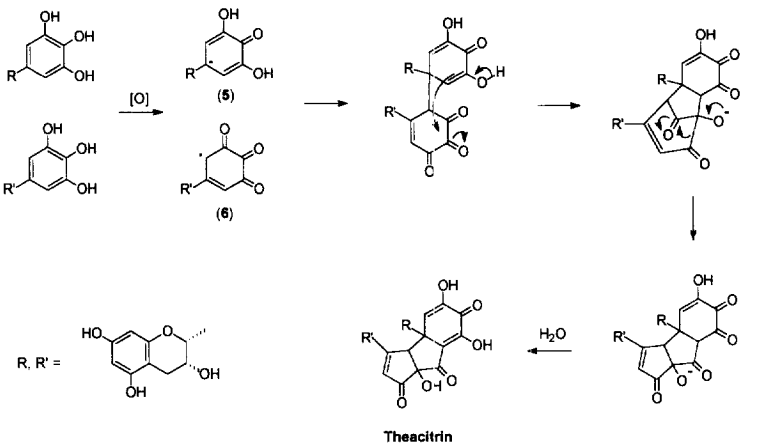


Fig. 2. Proposed mechanism for formation of theacitrins.

absorption as theacitrin A. It is thus postulated to be an isomer of theacitrin A, with the galloyl ester group linked to C-3 rather than C-3′. Confirmation of the structures of these two compounds is in progress.

The identification of theacitrin A is of great interest, since it represents a new class of coloured, dimeric, flavanol-derived polyphenolic compound present in black tea. It has some structural features in common with oolongtheanin [17], which also has a flavan-3-ol and a flavan-3-*O*-gallate joined together by a three-fused ring system. However, the major difference between oolongtheanin and theacitrin A lies in that the former has a M_r 28 Daltons (equivalent to a carbonyl group) less than theacitrin A. A mechanism for the formation of theacitrin A has been proposed and is shown in Fig. 2. The B-rings of EGC and EGCG are oxidized to give rise to two free radicals (**5**, **6**) and they are coupled *via* a C—C bond. This intermediate then undergoes intramolecular cyclization, rearrangement and hydration, finally leading to the formation of theacitrin A.

It is also of interest to note that oxidation of catechins during tea fermentation takes place principally on their B-rings, leading to formation of a variety of dimeric compounds, such as theaflavins, theasinesins [18], theaflavates [19], oolongtheanin and theacitrins. Some of these dimers are of particular importance to the colour of black tea infusions. It is expected that further oxidation of these dimers will certainly give rise to more diverse trimers, tetramers, etc, known as thearubigins. Thus, it is not difficult to speculate how complex, heterogenous the structures of thearubigins should be.

EXPERIMENTAL

Extraction. An Indian black tea blend (Lattakari Assam, Importers Ltd) was purchased from commercial sources. The tea leaf (50 g) was decaffeinated by repeated washing with $CHCl_3$. Decaffeination was completed by infusing the leaf into hot H_2O, removing the leaf debris by filtration and then repeatedly extracting the resulting soln with $CHCl_3$. Removal of $CHCl_3$, followed by freeze-drying, yielded a brown solid (7 g).

Fractionation. The procedures used are similar to those reported previously [9, 20]. Dried tea extract (7 g) was dissolved in 50% aq. MeOH and applied to a solka floc cellulose column (i.d. 80 mm). The column was eluted with MeOH and then Me_2CO until the eluent from the column was clear and colourless. The required fr. was then eluted with 50% aq. Me_2CO. Removal of solvent yielded a dark brown powder (3.4 g). This material was dissolved in 20% aq. Me_2CO and applied to a Sephadex LH-20 column (200 mm × 80 mm i.d.). Elution with 20, 40 and, finally, 80% aq. Me_2CO, yielded 5 frs. Freeze-drying of these frs produced 5 light brown samples [217, 82, 162, 140 and 15 mg, respectively], of which the third contained theacitrin A.

Isolation and purification of theacitrin A. Fr. 3 (30 mg) was dissolved in 10% aq. MeCN and 250 μl of this soln was applied to a 5 μm Hypersil C_{18} column (300 mm × 10.0 mm i.d.) operating with UV detection at 280 nm. The column was eluted with MeCN–HOAc–H_2O (10:0.5:89.5) at a solvent flow rate of 4.75 ml min^{-1}. Multiple injections were made, with the column being subjected to a wash procedure [MeCN–HOAc–H_2O (60:0.5:39.5)] and re-equilibration (5 min of the initial solvent composition) between injections. The peak eluting at 7 min was identified as theacitrin A by its UV-Vis spectrum (theacitrin B and C elute at 11.93 and 18.27 min, respectively). The frs eluting between 6.5 and 8.0 min were collected, combined and then concd. Further purification of theacitrin A was achieved by the above method but with a modified mobile phase composition: MeCN–HOAc–H_2O (8.0:0.5:91.5). Under these conditions, theacitrin A eluted at 9.35 min. Appropriate frs were collected from multiple injections and then combined and concd. The concentrate was loaded onto a Sephadex LH-20 column (40 mm × 10 mm i.d.) and the column

repeatedly eluted with H_2O in order to remove HOAc. Finally, the desired product was eluted with 50% aq. Me_2CO. Removal of solvent followed by freeze-drying yielded pure theacitrin A as a yellow solid (4 mg).

UV-Vis spectroscopy. The UV-Vis spectrum was obtained by the diode-array detector of the HPLC system. Two absorption maxima at 277 and 379 nm were observed.

Mass spectrometry. Negative ion electrospray MS were obtained at cone potentials of 30 and 70 V using a triple quadrupole instrument: $[M—H]^-$ detected at *m/z* 759, M_r 760.

NMR. Spectra were measured on a Bruker AMX400 spectrometer operating at a probe temp. of 303, 283 or 183 K using either a dual ^1H-^{13}C 5 mm probe or a multinuclear 5 mm inverse probe as appropriate. The solvent used was acetone-d_6 and spectra were referenced relative to the solvent peak (δ = 2.052 and 29.80 ppm, ^1H and ^{13}C, respectively). Sample concns were typically 2–3 mg per 0.5 ml. HMQC and HMBC expts were acquired with the parameters described previously [6].

IR. ν(CO) = 1707, 1657(sh), 1629 and 1609 cm^{-1} (KBr disc).

Acknowledgements—The authors wish to thank Prof. E Haslam, University of Sheffield, for advancing the probable mechanism for formation of theacitrins.

REFERENCES

1. Roberts, E. A. H., *The Chemistry of Flavonoid Compounds*, ed. T. A. Geissman. Pergamon, Oxford, 1962, p. 468.
2. Roberts, E. A. H., Cartright, R. A. and Oldschool, M., *Journal of Science and Food Agriculture*, 1957, **8**, 720.
3. Takino, Y., Imagawa, H., Horikawa, H. and Tanaka, A., *Agricultural and Biological Chemistry*, 1964, **28**, 64.
4. Opie, S. C., Robertson, A. and Clifford, M. N., *Journal of Science and Food Agriculture*, 1990, **50**, 547.
5. Davis, A. L., Cai, Y., Davies, A. P. and Lewis, J. R., *Magnetic Resonance Chemistry*, 1996, **34**, 887.
6. Davis, A. L., Cai, Y. and Davies, A. P., *Magnetic Resonance Chemistry*, 1995, **33**, 549.
7. Lewis, J. R., Davis, A. L., Cai, Y. and Davies, A. P., Unpublished results, 1996.
8. Powell, C., Ph.D. thesis, University of Surrey, U.K.
9. Powell, C., Clifford, M. N., Opie, S. C. and Gibson, C. L., *SCI Lecture Papers*, Series No. 0027, 1994.
10. Powell, C., Clifford, M. N., Opie, S. C. and Gibson, C. L., in *Polyphenols* 94. *XVIIe Journeés Internationales Groupe Polyphénols*, eds R. Brouillard, M. Jay and A. Scalbert. INRA Editions, Paris, France, 1995, p. 279.
11. Hemingway, R. W., Tobiason, F. L., McGraw, W. G. and Steynberg, J. P., *Magnetic Resonance Chemistry*, 1996, **34**, 424.
12. Derome, A. E. and Williamson, M. P., *Journal of Magnetic Resonance*, 1990, **88**, 177.
13. Bax, A. and Subramanian, S., *Journal of Magnetic Resonance*, 1986, **67**, 565.
14. Bax, A. and Summers, M. F., *Journal of the American Chemical Society*, 1986, **108**, 2093.
15. Dürckheimer, W. and Paulus, E. F., *Angewandt Chemisch*, 1985, **97**, 219.
16. Hansen, P. E., *Progress in NMR Spectroscopy*, 1981, **14**, 175.
17. Hashimoto, F., Nonaka, G. I. and Nishioka, I., *Chemical and Pharmaceutical Bulletin*, 1988, **36**, 1676.
18. Nonaka, G., Kawahara, O. and Nishioka, I., *Chemical and Pharmaceutical Bulletin*, 1983, **31**, 3906.
19. Wan, X., Nursten, H., Cai, Y., Davis, A. L., Wilkins, J. P. G. and Davies, A. P., *Journal of Science and Food Agriculture*, 1997, **74**, 401.
20. Bailey, R. G., Nursten, H. E. and McDowell, I., *Journal of Science and Food Agriculture*, 1992, **59**, 365.