Tetrahedron. Vol 28. pp 3259 to 3269 Pergamon Press 1972 Printed in Great Britain

STRUCTURE DETERMINATION OF FLAVONOID DISACCHARIDES BY MASS SPECTROMETRY

R. D. SCHMID
Institut de Chimie des Substances Naturelles-Gif-sur-Yvette, Essonne, France

(*Received in the UK* 24 *January* 1972; *Accepted for publication* 6 *March* 1972)

Abstract—The mass spectra of 13 perdeuteriomethylated flavonoid disaccharides comprising flavones, flavanones, flavonols and an isoflavone are reported. It is shown that their fragmentation pattern allows unequivocal differentiation according to the structural type of the aglycone, the sugar sequence and the position of the interglycosidic linkage. Moreover, information regarding substitution on the aglycone can be obtained. All compounds except one exhibited molecular ion peaks thus allowing direct molecular weight determination.

TRADITIONALLY, the structure of a flavonoid glycoside is elaborated after partial or complete hydrolysis leading finally to the constituent aglycone and the sugar residues.[1,2] A considerable amount of work employing analytical, physicochemical and synthetic methods is involved in order to establish:
(a) the type of aglycone (e.g., flavanone or flavone); (b) its substitution pattern; (c) the linkage position of the sugar moiety; and in flavonoid oligosaccharides, (d) the sequence of sugars; (e) the interglycosidic linkage.
Moreover, a good stock of reference compounds for chromatographic comparison is necessary. In an attempt to find a complementary method which circumvents some of these problems we recorded Electron Impact mass spectra of some per-deuteriomethylated flavonoid disaccharides.

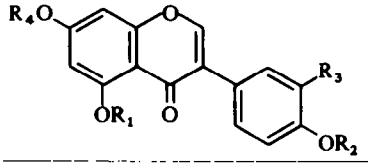
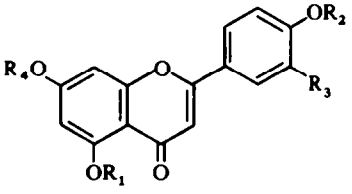
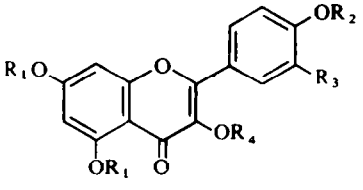
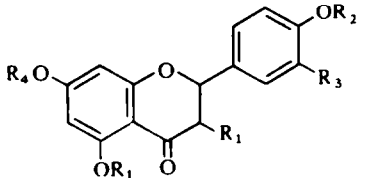
Mass spectrometry is an established tool in the analysis of flavonoid aglycones.[3] In many cases it yields sufficient information to elucidate their structural type and substitution pattern. On the other hand, mass spectrometric sequencing of oligosaccharide derivatives[4] has been much refined during the last years[5] to furnish not only the sequence of the sugars involved but also in many cases the position of their interglycosidic linkage and even information about the stereochemistry at the anomeric center.

In spite of this progress, no systematic study seems to have been made on the application of mass spectrometry to the structure determination of flavonoid oligoglycosides. While comprehensive papers exist on the fragmentation behaviour of underivatized[6] and peracetylated[10] flavonoid-C-monoglycosides including one underivatized flavone-O-monoglucoside,[6] the use of mass spectra of peracetylated flavonoid-O-monoglycosides encountered occasionally in the literature seems to have been limited to the verification of their molecular weight and, to the best of our knowledge, no spectra of flavonoid oligosaccharides have been published at all. We here report the mass spectrometric analysis of 13 perdeuteriomethylated flavonoid disaccharides obtained by a clean and rapid derivatization procedure.

RESULTS AND DISCUSSION

Table 1 gives the formulas of the isoflavone bioside (Ia), 3 flavone biosides (IIa, IIIa, VIa) 5 flavonol biosides (IVa, Va, VII–IXa), and 3 flavanone biosides (XI–XIIIa). The formula of the flavonol diglycoside (Xb) is given on the mass spectrum (Fig 1). III was available only as a mixture with II; IV and V were obtained as a mixture.

TABLE 1

	R_2	R_3	R_4	a- and b-series*	Compound name (R_1 = H)
	CH_3	H	1 → 2 api-glu	I	Lanceolarin
	R_1	H	1 → 2 api-glu	II	Apiin
	R_1	OCH_3	1 → 2 api-glu	III	Graveobioside B
	CH_3	H	1 → 6 rha-glu	VI	Linarin
	R_1	H	1 → 2 api-glu	IV	Kaempferol-apiosylglucosid
	CH_3?	OR_1?	1 → 2 api-glu	V	Methylquercetin-apiosylglucosid
	R_1	OR_1	1 → 6 rha-glu	VII	Rutin
	R_1	H	1 → 6 glu-glu	VIII	Kaempferol-gentiobioside
	R_1	OR_1	1 → 6 glu-glu	IX	Quercetin-gentiobioside
	R_1	H	3,7-rha$_2$	X	Kaempferitrin
	CH_3	OR_1	1 → 6 rha-glu	XI	Hesperidin
	R_1	H	1 → 2 rha-glu	XII	Naringin
	CH_3	H	1 → 2 rha-glu	XIII	Citrifolin

* R_1 = −H: a-series, natural products
R_1 = $-CD_3$: b-series, sugars perdeuteriomethylated; derivatives used for MS
abbreviations used: api = apiose, rha = rhamnose, glu = glucose

TABLE 2

IIb + IIIb	730(0·5), 700(0·6), 576(0·4), 546(0·5), 519(0·4), 518(1·5), 489(0·5), 488(2), 397(1·5), 363(3), 362(16), 335(9), 334(8), 306(5), 305(17), 304(15), 275(4), 207(4), 198(5), 195(4), 184(4), 149(100), 114(30)
IVb + Vb	763(0·1), 762(0·1), 733(0·1), 729(0·1), 699(0·15), 685(0·1), 645(0·1), 615(0·3), 579(0·3), 568(0·3), 563(0·4), 552(3), 551(9), 522(3), 521(15), 397(1), 396(1·5), 369(7), 368(37), 367(84), 362(3), 339(7), 338(68), 337(115), 336(3), 321(4), 320(3), 308(6), 290(3), 184(3), 149(100), 114(18)
VIb	711(0·5), 689(0·4), 678(1), 677(2·5), 676(0·3), 675(0·4), 659(0·3), 645(0·8), 640(0·6), 599(1·5), 560(1), 542(0·7), 531(0·6), 503(1), 499(4), 498(9), 431(0·7), 430(2), 411(0·3), 410(2·5), 361(4), 335(4), 334(9), 331(6), 330(4), 317(5), 316(15), 303(8), 302(37), 301(65), 300(5), 299(5), 272(8), 253(2), 198(100), 163(25), 128(8)
VIIb	780(0·4), 757(0·2), 745(0·2), 710(0·3), 629(0·3), 567(0·4), 566(0·8), 532(0·5), 464(1·5), 458(0·3), 430(0·8), 416(0·4), 411(0·5), 410(2), 387(10), 372(18), 371(100), 370(680), 369(19), 368(3), 367(7), 353(4), 352(7), 342(4), 341(55), 338(3), 337(10), 336(3), 324(5), 323(7), 198(11), 163(6), 128(3)
IXb	813(0·2), 790(0·2), 779(0·3), 765(0·3), 744(0·2), 629(0·3), 600(0·4), 566(1), 531(0·3), 482(0·5), 458(0·7), 443(0·9), 388(4), 387(15), 372(19), 371(100), 370(800), 369(18), 368(3), 367(6), 353(3), 352(4), 342(3), 341(13), 337(6), 324(4), 323(7), 231(2), 230(6), 196(17), 161(10)
XIIIb	732(0·4), 731(0·9), 730(2), 713(0·6), 532(0·4), 519(0·4), 518(1·4), 441(0·3), 424(0·4), 412(5), 411(25), 377(9), 376(44), 322(4), 321(20), 320(6), 292(10), 291(9), 198(100), 163(37), 128(11)

* Below *m/e* 250, only the values for the T-series are given.

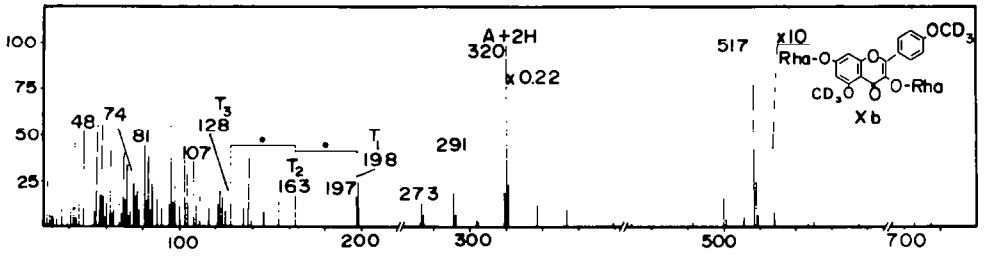


FIG 1

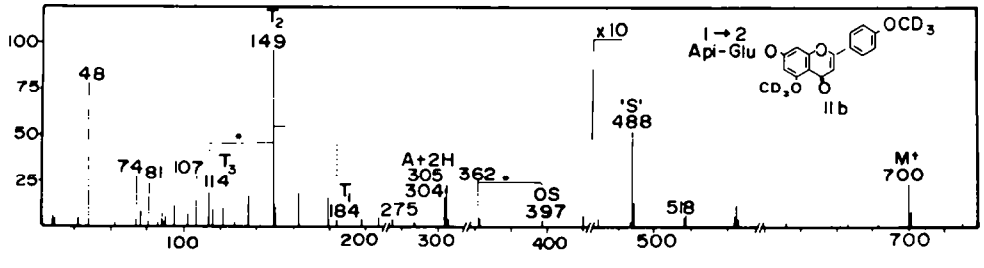


FIG 2

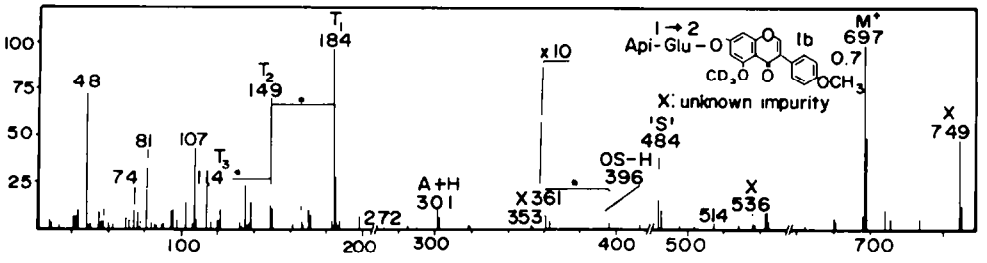


FIG 3

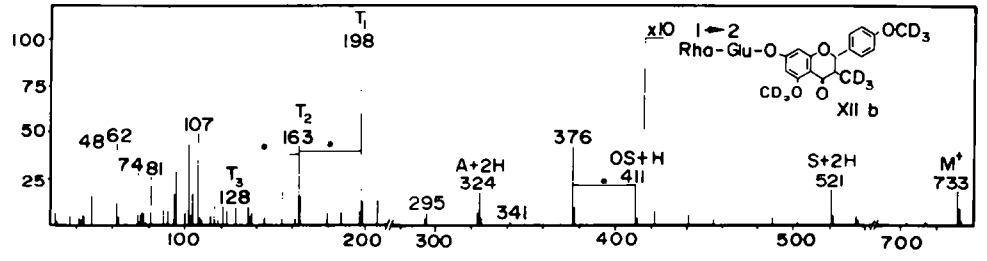


FIG 4

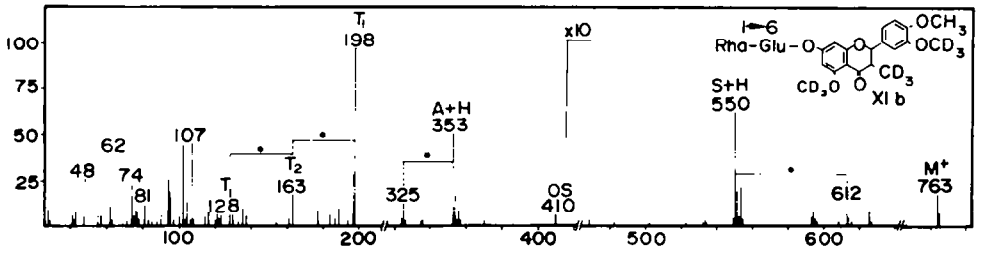


FIG 5

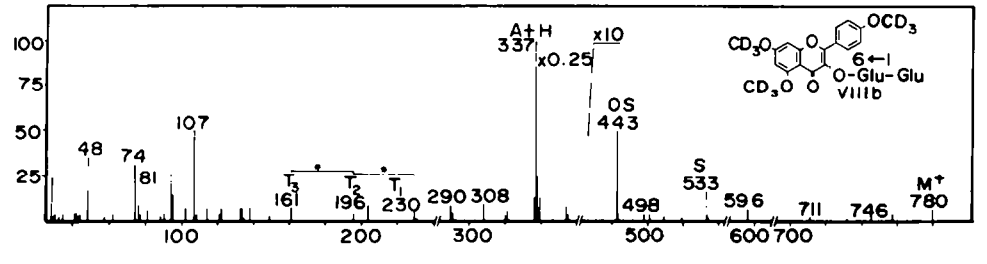


FIG 6

The spectra of compounds Ib, IIb, VIIIb, Xb, XIb and XIIb are reproduced in Figs 1–6. The other spectra are in Table 2. For most compounds spectra of the permethylated counterparts have been utilized to help interpretation of the fragmentation pathways.

Except Xb, all compounds exhibit molecular ion (M^+) peaks. The principal fragmentation pattern is illustrated in Fig 7.

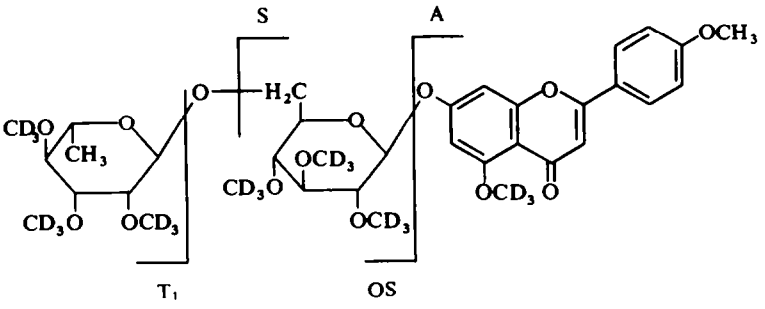


FIG 7

Fission of the glycosidic carbon oxygen bond of the terminal sugar leads to the ion T_1 which successively loses CD_3-methanol to give T_2 and T_3.

Rupture of the etheral carbon–oxygen bond between the terminal and the second sugar gives rise to the sequence ion S. Fission of the glycosidic carbon–oxygen bond between the sugar and the aglycone is indicated by the aglycone peak A invariably formed by transfer of hydrogen and followed by loss of CO.

Retention of charge on the disaccharide residue after this type of rupture leads to the oligosaccharide ion OS. In the spectra of apiosyl–glucose containing flavonoids, the S-peak is very weak. Instead, a strong peak 13 m.u. lower (isoflavone) or 12 m.u. lower (flavones, flavonols) is observed the origin of which is not clearly understood. The lower mass region of all spectra is dominated by peaks of the T-series and by fragment ions appearing at *m/e* 107, 94, 81, 74 and 48 which are usually observed in the spectra of (deuterio-)methylated sugars.[4] Peaks due to retro-Diels–Alder cleavage of the flavonoid aglycone are small or absent as it is often observed in highly substituted compounds of this type.[3]

From a more detailed inspection of the fragmentation pattern of compounds Ib–XIIIb, some facts emerge which allow their differentiation according to the mass and sequence of sugars, the nature and substitution of the aglycone and the position of the interglycosidic linkage.

Sequence of sugars

Information about the terminal sugar is obtained from the difference ($M^+ - S$) and from peaks due to ions of the T-series. As can be seen from Table 3, the T-series is the more reliable indication since the ($M^+ - S$) value can be changed by H-transfer or, in the case of the apiosyl-containing compounds, even more complicated reactions.

The diglycoside Xb exhibits a peak at *m/e* 197 in addition to the T_1 peak. This species, not observed in the biosides, is assumed to be rhamnosyl ion formed by loss of its anomeric hydrogen. Since it has been found in another non-linear flavonol tri-

TABLE 3

Compound	(M^+-S)	T_1	terminal sugar	(S-A)	$(OS-T_1)$	second sugar
Ib	(213)	184	api	(183)	212	glu
IIb	(212)	184	api	(184)	213	glu
IVb + Vb	(212)	184	api	(183)	213	glu
VIb	213	198	rha	197	213	glu
VIIb	214	198	rha	196	212	glu
VIIIb	247	230	glu	196	213	glu
IXb	247	230	glu	196	213	glu
XIb	213	198	rha	197	213	glu
XIIb	212	198	rha	197	213	glu
XIIIb	212	198	rha	197	213	glu

abbreviations used: api = apiose, rha = rhamnose, glu = glucose

saccharide,[7] it is tempting to suggest that its presence can be used as an indication of a nonlinear flavonoid oligosaccharide.

Only compounds containing glucose as the second sugar moiety have been available. For these cases studied, the $(OS-T_1)$ rather than the (S-A) difference seems to be the more useful for mass identification (Table 3).

Aglycone type and substitution

Some peaks of the isoflavone, flavones, flavonol and flavanone derivatives studied exhibit features which allow their distinction. In Table 4, intensities of these peaks are given in relation to the peak at *m/e* 107 of the same spectrum. The latter has been chosen because its formation from C-2, C-3 and C-4 of per(deuterio)methylated

TABLE 4

Compound	Aglycone	M^+	S or S + H or S − H	A + H	A + 2H
Ib	isoflavone	31	8	2·5	2
IIb	flavone	8	15	70	80
VIb		1	12	88	50
VIIb	flavonol	1	1	1830	400
VIIIb		0·5	1	820	200
IXb		1	3	800	210
Xb		—	280	40	1800
XIb	flavanone	4	13	110	35
XIIb		4	2	11	36
XIIIb		4	2	12	38

Intensities are expressed relative to *m/e* 107 = 100

hexopyranosides[4] justifies the assumption that its occurrence is of similar probability, in compounds VIb–XIIIb. This need not be true, however, for the apiose containing disaccharides Ib–Vb.

Table 4 shows that all flavonol derivatives exhibit very intense A + H (A + 2H for Xb) peaks which allow their unequivocal identification. The aglycon peaks of the flavone, flavanone and isoflavone derivatives are much smaller. The latter, however, can be recognized by its strong M^+ signal.

Peak intensity ratios do not seem to allow distinction between flavone and flavanone glycosides. The former, however, showed two peaks of similar intensity at A + H and A + 2H, while the latter are the only compounds of this study which do not give loss of methanol-(d_3) from the molecular ion.

The substitution pattern of the aglycone cannot be deduced from retro-Diels–Alder fragments because these are very weak or absent. By use of perdeuteriomethylation it is possible, however, to determine the number of hydroxyl- and methoxyl-groups present before derivatization, an odd mass number for the A + H peak representing an odd number of hydroxyl-groups; methoxyl-groups originally present can be detected by mass calculation for the limited number of probable structures.

Ambiguous cases usually can be resolved by running the spectra of the permethylated compounds, which, by comparison with their perdeuteriomethylated counterparts, yield the exact number of methyl groups introduced by the derivatization procedure.

Perdeuteriomethylation of the flavanones XIa–XIIIa resulted in the introduction of an extra deuteriomethyl group. By mass spectrometric evidence (loss of CO from the aglycone) and IR data (CO frequency at 1650 cm^{-1}) it is assumed that C-methylation α to the carbonyl group, not O-methylation of the carbonyl group after enolization, took place.

An application of the methods and reasonings described above is given by interpreting the structure of "impurity X" in the spectrum of lanceolarin (Ib, Fig. 3). Comparison of the spectra after perdeuteriomethylation and permethylation indicated that "impurity X" contains two extra methyl groups which should be located on the aglycone moiety since no disaccharide peak other than that corresponding to apiosyl-glucose can be observed. The intensities of the molecular ion M^+ and of the aglycone peak A + H suggest an aglycone similar to an isoflavone. The occurrence of peaks $M^+ + 52$ m.u., (S − 13) + 52 m.u. and (A + H) + 52 m.u. can be explained by an aglycone containing an extra ($OCD_3 + CD_3 + H$). Therefore "impurity X" is postulated to be a apiosyl-glucosyl isoflavanone perdeuteriomethylether which before derivatization contained one hydroxyl group more than lanceolarin (Ib); the C-deuteriomethyl group is introduced during derivatization.

Position of the interglycosidic linkage

The sequence peak S is formed without hydrogen transfer (flavonols) or with single hydrogen transfer (flavone, flavanone) in the case of 1 → 6 linked flavonoid disaccharide derivatives, while transfer of 2 hydrogens takes place in the 1 → 2 linked compounds. The aglycone and oligosaccharide fragments A and OS behave similarly: for 1 → 6 linked derivatives, the A + H and OS peaks are prominent, while A + 2H and OS + H peaks are observed for 1 → 2 linked compounds. (Table 6). Further differentiation is possible by a strong (OS − CD_3-methanol) peak only

TABLE 5

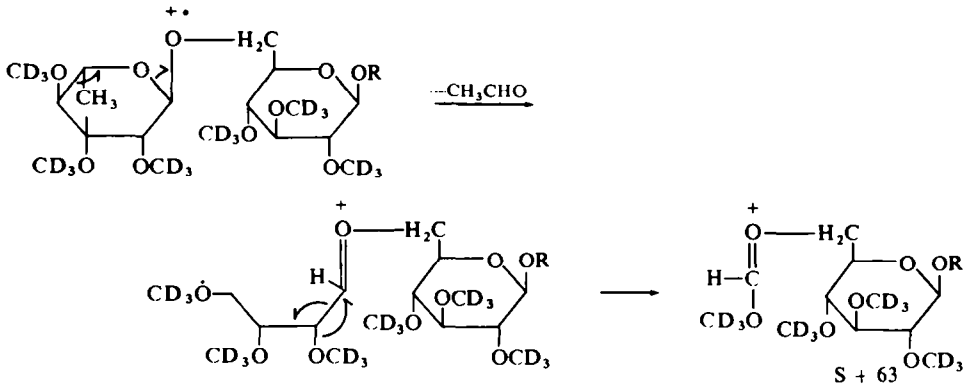
Aglycone	Compound	Sugar linkage	M^+		S		A			OS	
			M^+	M^+ -MeOH	S-CH	S-C	A + H	A + 2H	OS-H	OS	OS -MeOH
isoflavone	Ib	1 → 2	697	663	484		301	302	396		361
flavone	IIb	1 → 2	700	666 631		488	304	305		397	362
	IIIb	1 → 2	730	696 661		518	334	335		397	362
flavonol	IVb	1 → 2	763	729		551	367	368	396	397	362
	Vb	1 → 2	733	699		521	337	338	396	397	362

TABLE 6

Aglycone	Compound	Sugar linkage	M^+		S					A		OS		
			M^+	M^+ -MeOH	S	S + H	S + 2H	S + 62	S + 63	A + H	A + 2H	OS	OS + H	OS -MeOH
flavone	VIb	1 → 6	711	677		498		560		301	302	410		
flavonol	VIIb	1 → 6	780	745 710	566				629	370		413		
	VIIIb	1 → 6	780	746 711	533				596	337		443		
	IXb	1 → 6	813	779 744	566				629	370		443		
flavanone	XIb	1 → 6	763			550		612		353		410		
	XIIb	1 → 2	733				521				324		411	376
	XIIIb	1 → 2	730				518				321		411	376

present in the 1 → 2 linked glycosides, and by the S + 62 (flavone, flavanone) or S + 63 (flavonols) peak observed only in 1 → 6 linked compounds and explained in Scheme 1.

SCHEME 1



The 1 → 2 linked apiosyl-glucosyl flavonoid derivatives (Ib–Vb) show a slightly different fragmentation behaviour reproduced in Table 5. The main feature distinguishing them from the other compounds is the occurrence of the S-12 or S-13 peak mentioned above. In all five compounds loss of CD_3-methanol from the OS species (flavones, flavonols) or (OS–H) species (isoflavone) is observed in agreement with the 1 → 2 linked disaccharide derivatives not containing apiose.

CONCLUSION

A considerable amount of structural information can be obtained from the mass spectra of perdeuteriomethylated flavonoid disaccharides. For the 13 compounds described in this paper it has been possible to identify the aglycone type by an intense aglycone peak (flavonols), an intense molecular ion peak (isoflavone) or by C-methylation and lack of peaks due to loss of CD_3-methanol from the molecular ion (flavanones). The number of hydroxyl- and methoxyl-functions originally present in the aglycone can be determined by perdeuteriomethylation. The mass and the sequence of the sugars was established unequivocally. The diglycoside Xb could be distinguished from the biosides by the presence of an A + 2H peak, a peak at *m/e* 197 and a very strong peak due to the loss of the first sugar residue.

Finally, differences in hydrogen transfer to some peaks allowed prediction of the position of the interglycosidic linkage. Taking into account the ease of preparation of the derivatives, the minute amount of substance required and the ability of mass spectrometry to obtain valid information even on mixtures of compounds as shown for IIb + IIIb and IVb + Vb, it is concluded that mass spectrometry is a valuable complementary tool for the structure elucidation of flavonoid disaccharides. Preliminary evidence indicates that this is true for flavonoid trisaccharides as well.[7]

EXPERIMENTAL

For derivatization 0·5 to 2 μmoles of the flavonoid glycoside are perdeuteriomethylated by use of

CD_3I and NaH in DMF.[9] In many cases, the course of the reaction can be followed by the fading of the colour of the suspension. After 1 hr at room temp, workup is accomplished by partition between $CHCl_3$–H_2O. Yield is essentially quantitative. Purity of the products can be checked by TLC (silicagel) using EtOAc as solvent.

Mass spectra were recorded on an AEI MS 9 at 100 μA and 70 eV using a direct inlet system. Temperature varied between 210° and 240°.

Acknowledgments—Gifts of samples VIIIa and IXa from Prof. Dr. L. Hoerhammer, Institut fuer Arzneimittellehre, University of Munich, Ia–Va from R. Ortmann and Dr. K. Hahlbrock, Institut fuer Biologie II, University of Freiburg, and of the other compounds from Prof. Dr. R. Paris and Dr. A. Cave, Faculty of Pharmacy, Paris, are gratefully acknowledged.

The author expresses his gratitude to Dr. B. C. Das for his steady encouragement. He wants to record his gratitude to Prof. Dr. E. Lederer and Dr. Das for the possibility to work in this institute and for their help in manuscript preparation.

A scholarship of Deutsche Forschungsgemeinschaft is gratefully acknowledged.

REFERENCES

[1] J. B. Harborne, *Comparative Biochemistry of the Flavonoids*, Academic Press, London (1967)
[2] T. J. Mabry, K. R. Markham and M. B. Thomas, *The Systematic Identification of Flavonoids*. Springer, New York (1970)
[3] D. G. I. Kingston, *Tetrahedron* **27**, 2691 (1971)
[4] N. K. Kochetkov and O. S. Chizov, *Advan. Carbohydrate Chem.* **21**, 39 (1966)
[5] O. S. Chizov, N. K. Kochetkov, N. N. Malysheva, A. K. Shiyonok and V. L. Chashchin, *Org. Mass Spectrom.* **5**, 1157 (1971)
[6] A. Prox, *Tetrahedron* **24**, 3697 (1968)
[7] R. D. Schmid, unpublished results
[8] W. Hoesel and W. Barz, *Phytochem.* **9**, 2053 (1970)
[9] J. S. Brimacombe, B. D. Jones, M. Stacey and J. J. Willard, *Carbohyd. Res.* **2**, 167 (1966)
[10] M. Aritomi, T. Komori and T. Kawasaki *Liebigs Ann. Chem.* **734**, 91 (1970)