Phytochemistry, 1968, Vol. 7, pp. 831 to 837. Pergamon Press. Printed in England

# MASS SPECTROMETRY AS AN AID FOR DETERMINING STRUCTURES OF NATURAL GLUCOSIDES[1]

IRWIN A. PEARL and STEPHEN F. DARLING

The Institute of Paper Chemistry, Appleton, Wisconsin, 54911

(*Received* 11 *October* 1967)

**Abstract**—Mass spectral fragmentation patterns were determined for acetates of a number of glucosides of unequivocally known structure which have been isolated from the barks and leaves of trees of the family Salicaceae. All glucoside acetates, irrespective of size, indicated primary fragmentation at the C-1 bond, essentially similar to that found for completely acetylated glucose and methyl glucoside. In addition, the spectra of the acetates of glucosides with free phenolic groups demonstrated loss of acetyl as ketene either concurrently with or before the rupture of the C-1 bond. Presence or absence of ester substitution in the glucose moiety of the parent glucosides could thus be determined with certainty, and possible location of substitution is indicated by relative intensities of fragment peaks.

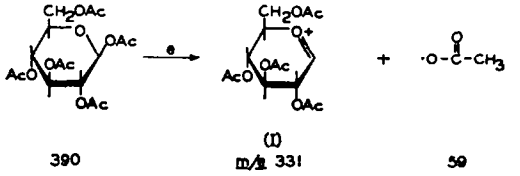
IN OUR continuing studies on the isolation and characterization of glucosides from the barks and leaves of *Populus* and *Salix* species, we have isolated during the past few years a number of phenolic glucosides of unknown composition. In many instances, the small yields of pure products precluded the use of classical methods comprising specific chemical reactions employed so effectively for the earlier characterization of tremuloidin,[2] grandidentatin,[3] and the 2-*O*-benzoate and 6-*O*-benzoate of salicyloylsalicin.[4,5] In other cases, the nature of the aglucone made such reactions as methylation impractical. Thus, it became necessary to employ alternate methods for structure determination of the newly isolated glucosides.

Because of the success attained in other laboratories on the determination of structures of different types of naturally occurring compounds isolated in extremely small quantities by means of mass spectrometry, we decided to employ this method in our studies. As a first step in this direction, it became necessary to ascertain the mass spectral fragmentation patterns for volatile derivatives of already isolated glucosides of unequivocally known structure. Biemann *et al.*,[6] in a study of the mass spectrometry of sugars and some of their derivatives, found that sugars gave no useful mass spectra, but sugar acetates were well suited for mass spectrometric examination. Accordingly, we determined the mass spectra of the acetates of a number of glucosides isolated from various *Populus* and *Salix* species barks and leaves.

Biemann and co-workers,[6] in their earlier study employing deuterated compounds, demonstrated that primary mass spectral fragmentation of glucose pentaacetate took place at the C-1 carbon in accordance with Scheme I to give a prominent peak in the mass spectrum

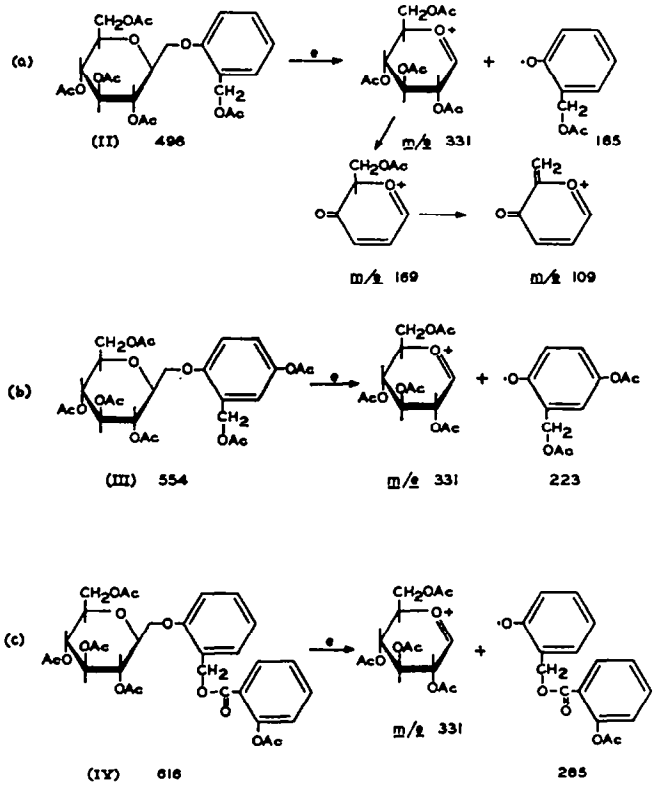
[1] Paper No. XV in the series "Studies on the Barks of the Family Salicaceae" and No. X in the series "Studies on the Leaves of the Family Salicaceae".
[2] I. A. PEARL and S. F. DARLING, *J. Org. Chem.* **24**, 731 (1959).
[3] I. A. PEARL and S. F. DARLING, *J. Org. Chem.* **27**, 1806 (1962).
[4] I. A. PEARL and S. F. DARLING, *Arch. Biochem. Biophys.* **102**, 33 (1963).
[5] I. A. PEARL and S. F. DARLING, *Tappi* **48**, 506 (1965).
[6] K. BIEMANN, D. C. DEJONGH and H. K. SCHNOES, *J. Am. Chem. Soc.* **85**, 1763 (1963).

at *m/e* 331 corresponding with the tetraacetylglucose oxonium ion (I). The mass spectrum contained several other peaks at *m/e* 169 and 109 due to secondary fragmentation of the primary ion by loss of acetic acid or ketene and at *m/e* 43 due to the acetyl ion. These authors also established the fact that acetates of methylglucoside and phenylglucoside follow essentially this same fragmentation pattern, the main difference being that the spectra of these two glucoside acetates contained even more prominent peaks at *m/e* 331 than did that of glucose pentaacetate, due apparently to the greater instability of the C-1 bond in these glucoside acetates.




SCHEME I.

Mass spectra were determined for the acetates of three *Populus*-derived glucosides, salicin (II), salirepin[7] (III), and salicyloylsalicin (IV), all of which contain the tetraacetyl-

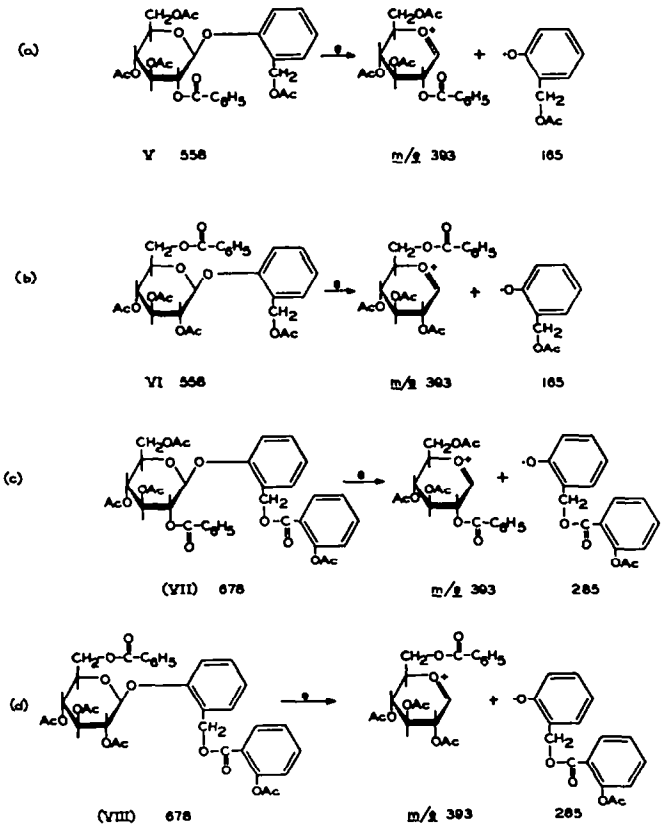



SCHEME II.

[7] Debenzoylated salireposide.

glucose moiety. These spectra all showed prominent peaks at *m/e* 331, 169, 109, and 43, the same prominent peaks found by Biemann and co-workers for the acetates of glucose, methyl glucoside, and phenyl glucoside. Thus, it appears from the mass spectra of the three glucosides that they also undergo primary fragmentation at the glucosidic linkage resulting in the formation of the tetraacetylglucose oxonium ion (I). These possible fragmentation patterns are pictured in Scheme II. It should be noted that Scheme IIa includes the recent postulations of the structural formulas for the secondary fragments *m/e* 169 and *m/e* 109 as proposed by Smale and Waight.[8]

The mass spectra of the acetates of four glucosides containing benzoyl substitution on the glucose moiety were determined next. The spectra of acetates of tremuloidin (2-*O*-benzoylsalicin) (V), populin (6-*O*-benzoylsalicin) (VI), salicyloylsalicin-2-*O*-benzoate (VII), and salicyloylsalicin-6-*O*-benzoate (VIII) are all characterized by the complete absence of a peak at *m/e* 331. Instead, all four exhibit a prominent peak at *m/e* 393 corresponding with a primary fragmentation at the C-1 bond leaving this time a triacetylbenzoylglucose oxonium ion of *m/e* 393 instead of the tetraacetylglucose oxonium ion of *m/e* 331 as in the case of the nonbenzoylated glucosides. These fragmentation patterns are pictured in Scheme III.

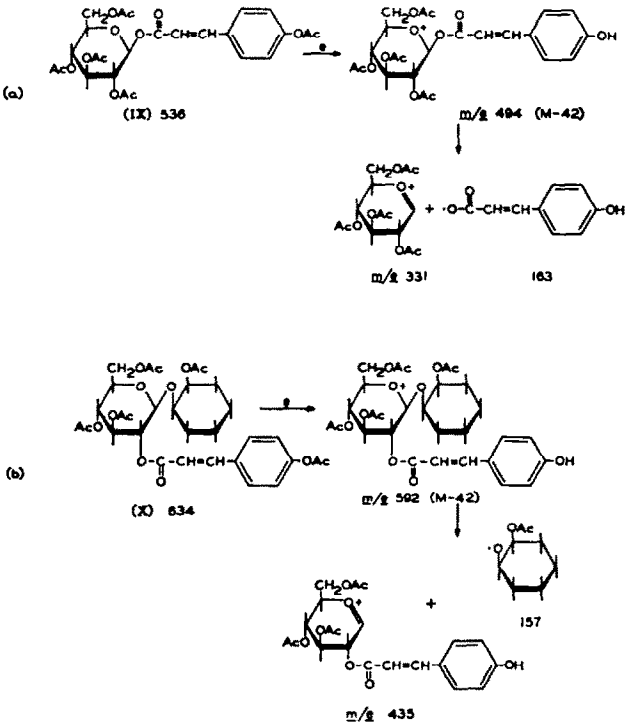



SCHEME III.

[8] T. C. SMALE and E. S. WAIGHT, *Chem. Comm.* 680 (1966).

The mass spectra of these four benzoylated glucosides demonstrate the effect of position of the benzoyl ester substitution on the glucose moiety of glucosides. While the spectra of the two isomeric pairs of benzoylated glucosides are essentially the same, they differ in that the peaks at *m/e* 231 in the spectra of the two 6-*O*-benzoylated glucoside acetates (VI) and (VIII) are quite prominent, but the same peaks are very weak in the spectra of the two 2-*O*-benzoylated glucoside acetates (V) and (VII). We attribute this peak to the loss of two molecules of acetic acid (mol.wt. 60) and one molecule of ketene (mol.wt. 42) from the primary ion at *m/e* 393. Apparently, these fragmentations take place much more readily when all three acetyl groups are contiguous as they are in the 6-*O*-benzoylated glucoside acetates than they do when they are separated as they are in the 2-*O*-benzoylated glucoside acetates. Thus, there is a difference in the mass spectra of isomeric glucoside ester acetates which appears to be due to the difference in position of the ester group on the glucose unit. This difference in spectra should prove useful in locating the ester groups in similar compounds of unknown structure.

The acetates of two *Populus*-derived glucosides containing the *p*-coumaroyl group were submitted next to mass spectrometry. These were 1-*O*-*p*-coumaroyl-β-D-glucose pentaacetate (IX) and grandidentatin pentaacetate (X). The spectrum of IX exhibits a prominent peak at *m/e* 331. This peak, corresponding with the ion, I, was expected from the structural relation of IX to II, III, and IV. In addition, the spectrum of IX contains a peak at *m/e* 494 (M-42) attributed to the loss of acetyl as ketene (mol.wt. 42) from the molecular ion. This loss of acetyl from the *p*-hydroxyl group of the *p*-coumaroyl ester residue takes place before or concomitant with the fragmentation at the C-1 bond of the glucose. Thus, the acetyl

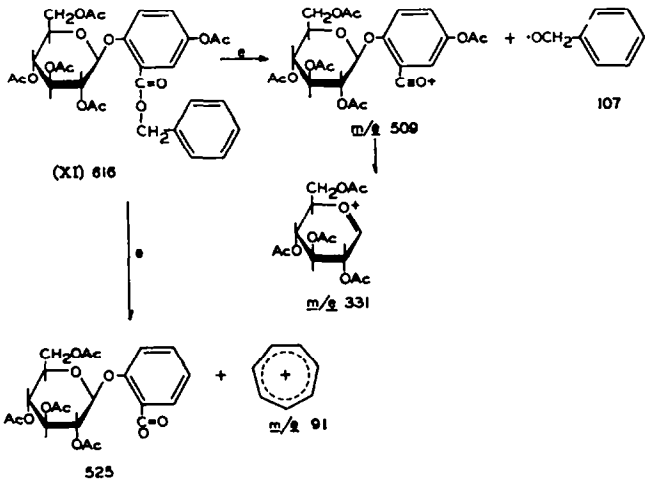



SCHEME IV.

group on the *p*-hydroxyl appears to be less stable than the C-1 linkage. The fragmentation pattern of IX is shown in Scheme IVa.

Similarly, the mass spectrum of the *p*-coumaroylated glucoside acetate X contains a peak at M-42, and the formed intermediate ion of *m/e* 592 reaches the target in the mass spectrometer before all of the C-1 bond is severed. The fragmentation pattern of X is pictured in Scheme IVb.

We next examined the mass spectral fragmentation of trichocarpin pentaacetate (XI). The structure of trichocarpin has been established by Loeschcke and Francksen[9] as the benzyl ester of gentisic acid β-D-glucoside. The spectrum contains a prominent peak at *m/e* 331 due to the I ion from the unsubstituted glucose unit in the glucoside acetate XI. The spectrum also contains a prominent peak at *m/e* 509 corresponding with M-107. The loss of a mass of 109 is attributed to the loss of the benzyloxy group from the molecular ion. The strong peak at *m/e* 91 is accounted for by the tropylium ion from the fragmentation of the molecular ion. This latter fragmentation step is apparently the preferred one as judged by the prominence of the *m/e* 91 peak. The suggested fragmentation patterns of XI are pictured in Scheme V.




SCHEME V.

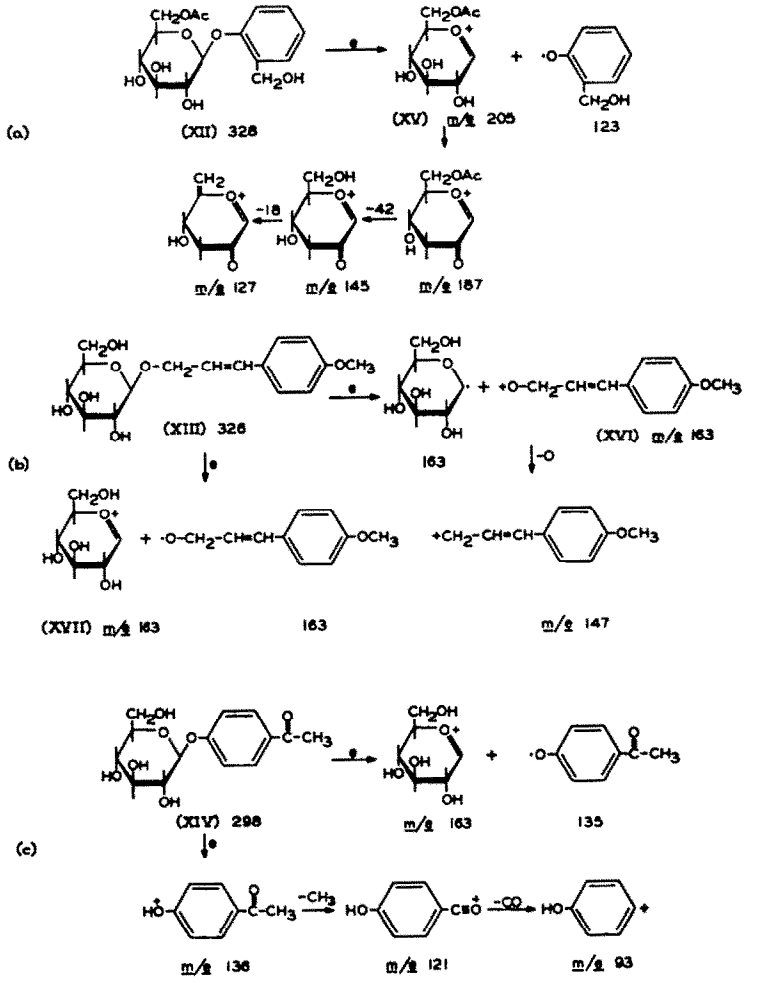
Finally, we determined the mass spectra of three low-melting *Salix* species-derived glucosides directly without prior acetylation. These glucosides were fragilin (6-*O*-acetyl salicin) (XII), vimalin (1-*O*-*p*-methoxycinnamyl-β-D-glucose) (XIII), and picein (1-*O*-*p*-acetylphenyl-β-D-glucose) (XIV). The spectra of these glucosides have fewer peaks than those of the glucoside acetates. However, the peaks that are present appear to be related to a systematic fragmentation rather than to a general pyrolysis.

The fragmentation of XII is characterized by primary rupture of the C-1 bond to give the 6-acetylglucose oxonium ion (XV) with *m/e* 205 as noted in Scheme VIa. Scheme VIa also pictures other fragmentations of this intermediate ion (XV) to account for other major peaks in the mass spectrum of XII by loss of water (mol.wt. 18), and acetyl as ketene (mol.wt. 42) or acetic acid (mol.wt. 60). The peak at *m/e* 43 is due to the acetyl ion.

[9] V. LOESCHCKE and H. FRANCKSEN, *Naturwissenschaften* **51**, 140 (1964).

The fragmentation of XIII is an interesting case because primary fragmentation at the C-1 bond would give rise to two possible ions, both with a unit mass of *m/e* 163 as depicted in Scheme VIb. The mass spectrum shows a prominent peak at *m/e* 163, but whether this is due to the *p*-methoxycinnamyl ether ion (XVI) or to the glucose oxonium ion (XVII) or both, we are unable to determine. A mass spectrum of this compound obtained by a high resolution instrument should be able to ascertain whether this peak at *m/e* 163 is a singlet or a doublet due to two ions of nearly the same *m/e* or not. It seems probable, however, that part of the *m/e* 163 peak is due to the XVI ion because this would account for the strong peak at *m/e* 147 by loss of oxygen from XVI.

The spectrum of XIV indicates that the glucosyl ion, *m/e* 179, and the glucose oxonium ion (XVIII) both reach the target in the mass spectrometer. Other major peaks at *m/e* 136, *m/e* 121, and *m/e* 93 can be attributed to the aglucone ion and further fragmentation products as noted in Scheme VIc. Again, the peak at *m/e* 43 is due to the acetyl ion.




SCHEME VI.

## EXPERIMENTAL[10]

### *Materials*

All glucoside acetates were prepared from glucosides isolated from *Populus* species barks and leaves in our laboratories, and most have been described in previous papers as follows: salicin pentaacetate (II),[11] salicyloylsalicin pentaacetate (IV),[4] tremuloidin tetraacetate (V),[2] salicyloylsalicin-2-*O*-benzoate tetraacetate (VII),[4] salicyloylsalicin-6-*O*-benzoate tetraacetate (VIII),[4] 1-*O*-*p*-coumaroyl-β-D-glucose pentaacetate (IX),[12] and grandidentatin pentaacetate (X).[3]

*Salirepin hexaacetate (III).* Salireposide, isolated pure in quantity from the bark of triploid *Populus tremuloides*,[12] was debenzoylated with $Ba(OH)_2$ according to Rabaté.[13] The resulting salirepin (0·19 g) was acetylated with 3 ml of pyridine and 2 ml of acetic anhydride, and the resulting product was recrystallized from dilute ethanol to yield 0·133 g of salirepin hexaacetate melting at 103–105° and having a specific rotation $[\alpha]_D^{20}$ −18·0° (*c*, 2·7 in chloroform). Its i.r. absorption contained bands at 2·93, 3·41, 5·70, 6·68, 7·00, 7·27, 8·15, 8·40, 8·98, 9·32, 9·64, 10·55, 11·05, 11·49, 11·92, 12·12, 12·45, and 12·91 μ. (Found: C, 54·30; H, 5·34. Calc. for $C_{25}H_{30}O_{14}$: C, 54·11; H, 5·41 per cent.)

*Populin tetraacetate (VI).* Populin[14] (3·5 g) was acetylated in the same manner, and the product was recrystallized from ethanol to give 3·4 g of white needles of populin tetraacetate melting at 126–126·5° and having a specific rotation $[\alpha]_D^{24}$ −6·0° (*c*, 3·74 in chloroform). Its i.r. absorption contained bands at 2·95, 3·41, 5·70, 5·77, 6·23, 6·68, 6·88, 6·98, 7·26, 7·60, 7·85, 8·05–8·10, 8·26, 8·64, 8·90, 9·30, 9·37, 9·50, 9·68, 10·31, 10·85, 11·03, 11·31, 11·79, 12·58, 13·16, 13·70, and 14·05 μ. (Found: C, 60·19; H, 5·42. Calc. for $C_{28}H_{30}O_{12}$: C, 60·21; H, 5·41 per cent.)

*Trichocarpin pentaacetate (XI).* Trichocarpin (0·4 g) isolated from the bark of *P. balsamifera*,[15] was acetylated as above. The product was recrystallized from ethanol to yield 0·4 g of trichocarpin pentaacetate melting at 117–118° and having a specific rotation $[\alpha]_{546}^{25}$ mμ −31·2° (*c*, 3·33 in chloroform). Its i.r. absorption spectrum contained bands at 2·94, 3·41, 5·70, 6·18, 6·27, 6·68, 7·02, 7·30, 7·68, 8·07, 8·27, 8·45, 8·95, 9·26, 9·40, 9·60, 10·18, 11·01, 11·34, 11·68, 11·94, 12·45, 12·78, 13·29, and 14·38 μ. (Found: C, 58·51, 58·65; H, 5·31, 5·27. Calc. for $C_{30}H_{32}O_{14}$: C, 58·44; H, 5·23 per cent.)

*Nonacetylated glucosides.* The *Salix*-derived nonacetylated glucosides fragilin (XII), vimalin (XIII), and picein (XIV) were kindly supplied by Dr. H. Thieme, Pharmaceutical Dept., Karl Marx University, Leipzig, Germany.

### *Mass Spectra*

All mass spectra were made on a Model RMU-6D Hitachi Perkin–Elmer Spectrometer. The spectrum of IX was kindly prepared by Mr. Will Major, Perkin–Elmer Corp., Norwalk, Conn. All other mass spectra were prepared by Morgan-Schaffer Corp., Montreal, Quebec, Canada.

The major and important *m/e* peaks taken from the actual mass spectrograms run at 70 eV are given in the following list which includes relative *m/e* intensity for each peak in parentheses.

II: 496 (<0·1), 331 (16·8), 169 (68·5), 109 (54·5), 43 (100).
III: 554 (<0·1), 331 (8·9), 169 (46·0), 109 (34·0), 43 (100).
IV: 616 (<0·1), 331 (8·3), 169 (77·2), 109 (50·0), 43 (100).
V: 558 (0·1), 393 (7·7), 231 (1·0), 169 (9·0), 109 (7·7), 105 (100), 43 (31·2).
VI: 558 (0·1), 393 (11·7), 231 (19·2), 169 (4·2), 109 (8·5), 105 (100), 77 (7·7), 43 (47·0).
VII: 678 (0·1), 393 (16·2), 231 (1·2), 169 (18·1), 121 (14·3), 109 (13·7), 105 (100), 77 (8·8), 43 (65·0).
VIII: 678 (0·1), 393 (23·0), 231 (38·1), 169 (10·5), 121 (17·8), 109 (13·8), 105 (100), 77 (5·9), 43 (67·2).
IX: 536 (0·1), 494 (31·8), 331 (45·5), 189 (50·0), 169 (63·5), 147 (81·8), 109 (70·5), 43 (100).
X: 634 (0·1), 592 (2·4), 435 (9·5), 225 (4·8), 189 (16·7), 147 (100), 99 (33·4), 82 (28·6), 43 (73·7).
XI: 616 (0·1), 509 (3·1), 331 (62·6), 169 (78·2), 109 (72·0), 91 (75·0), 43 (100).
XII: 328 (0·1), 205 (25·0), 187 (12·0), 166 (11·2), 145 (25·0), 127 (12·0), 124 (16·5), 106 (50·0), 78 (40·0), 43 (100).
XIII: 326 (5·0), 164 (12·5), 163 (12·5), 147 (100), 91 (7·5).
XIV: 298 (4·0), 179 (4·5), 163 (6·5), 145 (7·5), 136 (77·5), 121 (100), 93 (20·0), 43 (75·0).

### *Infrared Spectra*

I.r. absorption spectra were obtained with a Perkin–Elmer model 21 recording spectrophotometer using a NaCl prism and KBr pellets prepared by hand grinding with sample before pressing.

[10] All melting points are uncorrected. Analyses were performed by Micro-Tech Laboratories, Skokie, Illinois. I.r. absorption spectra were determined by Mr. Lowell Sell of The Institute of Paper Chemistry Analytical Department.

[11] D. H. BRAUNS, *J. Am. Chem. Soc.* **47**, 1286 (1925).

[12] I. A. PEARL and S. F. DARLING, *Tappi* **50**, 193 (1967).

[13] J. RABATÉ, *Bull. Soc. Chim. Biol.* **17**, 328 (1935).

[14] N. K. RICHTMYER and E. H. YEAKEL, *J. Am. Chem. Soc.* **56**, 2495 (1934).

[15] I. A. PEARL and C. R. POTTENGER, *Tappi* **49**, 152 (1966).