Chemometric Elucidation of Methylation Analysis of Polysaccharides

[a]M. ĎURINDOVÁ, [b]P. JURÁŠEK, [a]I. KLBIK, and [a]V. KOVÁČIK

[a]*Institute of Chemistry, Slovak Academy of Sciences, CS-842 38 Bratislava*

[b]*Department of Wood, Pulp, and Paper, Faculty of Chemical Technology, Slovak Technical University, CS-812 37 Bratislava*

Received 27 October 1992

The suitability of a chemometric approach to the evaluation of GC—MS measurements and to the interpretation of mass spectra of partially methylated alditol acetates has been examined. The products of methylation analysis of polysaccharide, prepared from water-soluble pectin, isolated from the spruce callus culture (*Picea abies* (L.) Karst.) walls, were interpreted manually by the expert and by way of the computer-assisted library search method. Two large commercial mass spectral libraries and library of mass spectra of partially methylated alditol acetates (PMAA), formed for the purpose to test the usefulness and reliability of small, specialized libraries, were used to compare the results of manual and computer-assisted interpretation. The identification of mass spectra of components in the examined mixture by way of large libraries failed. On the contrary, use of library PMAA led to satisfactory results, comparable with those obtained manually by an expert in the field.

The function of carbohydrates as structural constituents of the cell wall of plants, fungi, and bacteria, and their role as storage polymers is well known. Because of their importance, the interest of chemists, biochemists, and technologists in the chemical structure of carbohydrates constantly increases. The variability of inter-sugar linkages and their linkages to other constituents of natural material results in chemical and structural diversity. A fundamental step in the elucidation of tertiary structure of polysaccharides is the determination of the mode of intersaccharidic linkages.

The most widely used method for the determination of the structure of polysaccharides and their conjugates (*e.g.* lignin—saccharide complexes) is the methylation analysis [1]. Here, the free hydroxyl groups of a polymer are methylated and the methylated material is hydrolyzed. The unmasked hydroxyls in the formed partially methylated mono- or oligomers indirectly reveal the positions of the linkages in the original material. Chemical reduction of the reducing sugars to the respective alcohols permits, after acetylation, the determination of carbons participating originally in the ring and in the linkages.

The technique of gas chromatography—mass spectrometry (GC—MS) allows to analyze the products of the methylation analysis, *e.g.* a mixture of partially methylated alditol acetates. The mass spectra of alditol acetates differ depending on the number and positions of methyl groups. Additional information about the stereochemistry of the positional isomers and/or the presence of individual monosaccharides in a polysaccharide can be obtained from retention times of alditol acetates in gas chromatograms. Interpretation of mass spectral data is difficult, time-consuming and its correctness depends upon the expertise of the interpreter.

Chemometric approaches seem to be generally useful for the purposes of interpretation of similar chemical problems [2]. At present, from the numerous computer-assisted interpretations [3, 4] the library search [5—7] is the most common approach used. The search algorithms, *e.g.* PBM [8] and SISCOM [9], have been thoroughly elaborated by producers of GC—MS systems and are incorporated in combined GC—MS—DS systems. Other computer-aided interpretation systems as principal-component analysis in combination with library search [10], chemometric detectors [11—13], pattern recognition [14, 15], spectra modelling [16], or expert systems [17, 18] are at the present time hardly used.

In order to apply the chemometric approach in the evaluation and interpretation of mass spectra of partially methylated alditol acetates, the efficiency and the accuracy of library search technique as compared with manual interpretation was investigated. The mass spectrometric interpretation of the products of methylation analysis was performed as follows:

1. Manually, by an expert, taking into account the fragmentation rules known for the encountered substances;

2. by the use of two commercial libraries of mass spectra;

3. by the use of specialized library of mass spectra of partially methylated alditol acetates (PMAA), created for this purpose.

EXPERIMENTAL

The studied sample of partially methylated alditol acetates (METANA5*) was prepared from water-soluble pectin, isolated from the spruce callus culture (*Picea abies* (L.) Karst.) walls [19].

The GC—MS analysis was performed using an instrument HP 5940 (Hewlett—Packard) under the following conditions: capillary column SP 2330 (30 m × 0.25 mm, i.d. = 0.2 μm); inlet pressure 125 kPa; carrier gas helium; flow-rate 20 $cm^3 min^{-1}$, split ratio 29; column temperature, programmed to 180 °C for 1 min up to 240 °C at the temperature gradient 3 °C min^{-1}; injector temperature 240 °C; ion-source temperature 250 °C; electron energy 70 eV; mass scan range 29—450.

The library search using the Hewlett—Packard software is based on the method of Probability Based Matching (PBM) [8]. Three types of libraries of electron impact (EI) mass spectra were used with the aim to mutually compare the results of computer interpretation:

– Wiley/NBS NBS43K library (43 005 mass spectra) using the HP-300 computer without a coprocessor (HP 57790C MS Chemstation); the search strategy parameters used: Delta U + A 2, the number of significant peaks 10, NO tilt and the scan range set at Full.

– Wiley/NBS NBS49K library (49 469 mass spectra) using PC AT 286 with a coprocessor (HP DOS Chemstation); the search strategy parameters: Delta U + A 2, Tilting ON, the minimum purity 50 %, the peaks significance higher than 3 %.

– the special library of partially methylated alditol acetates (PMAA; 47 mass spectra) using the HP-300 computer (HP 57790C MS Chemstation) with the search strategy parameters as above. The library PMAA has been created from the published EI mass spectra of compounds [20], without participation of the expert who would interpret the results of methylation analysis and to whose interpretation of mass spectra the library search outputs would be compared.

RESULTS AND DISCUSSION

The mixture of partially methylated alditol acetates of spruce callus culture cells was subjected to GC—MS analysis. The respective gas chromatogram is shown in Fig. 1. In the sample, 15 compounds were identified applying rules of fragmentation for methyl and acetyl derivatives of saccharides. These do not distinguish the stereochemistry of compounds under investigation. The manual interpretation by an expert was used as a base for the evaluation of the results obtained applying the computer library search method, using two commercially available mass spectra libraries and the one specially created for this purpose. Table 1 summarizes the results of interpretation.

Neither of the searches using commercial libraries (NBS43K and NBS49K) led to satisfactory results. The number and the position of substituents in partially methylated alditol acetates and/or the number of carbons in compounds could not be correctly established. Thus, the library search method provided erroneous identification of components yielding structures more or less different from the correct one. Regarding the use of a library search in commercial libraries the dissatisfactory results may

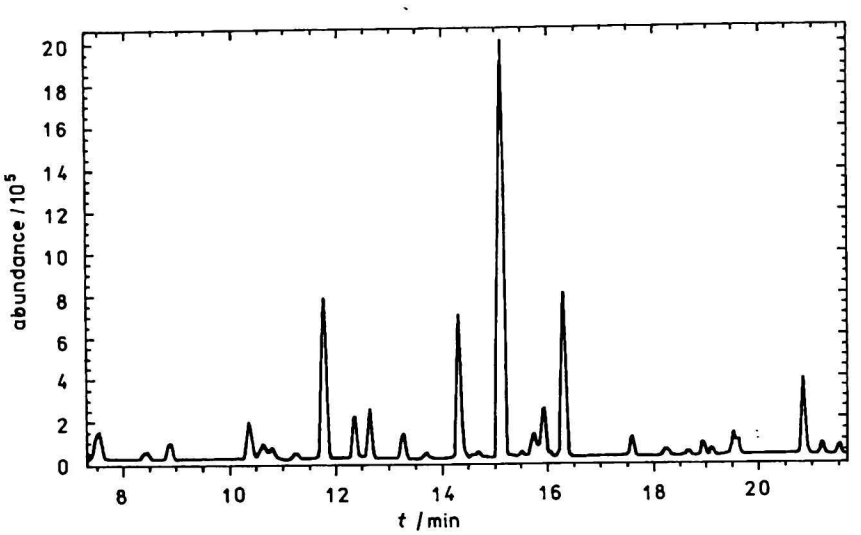

Fig. 1. Total ion current (TIC) in gas chromatogram of the sample METANA5*.

Table 1. Comparison of the Results of Library Search Using Libraries NBS43K, NBS49K, and PMAA for Sample METANA5*

Retention time	Manually interpreted	Similarity parameter*/%		
min	compound	NBS43K	NBS49K	PMAA
8.8	1,5-di-*O*-Ac 2,3,4-tri-*O*-Me pentitol	[20]	?	29
10.3	1,2,5-tri-*O*-Ac 6-deoxy-3,4-di-*O*-Me hexitol	?	[17]	37
10.6	1,5-di-*O*-Ac 2,3,4,6-tetra-*O*-Me hexitol	[41]	[59]	?
10.7	1,5-di-*O*-Ac 2,3,4,6-tetra-*O*-Me hexitol	[40]	[64]	?
11.8	1,5-di-*O*-Ac 2,3,4,6-tetra-*O*-Me hexitol	[86]	[90]	79
12.3	1,4,5-tri-*O*-Ac 2,3-di-*O*-Me pentitol	?	[50]	54
13.2	1,4,5-tri-*O*-Ac 2,3-di-*O*-Me pentitol	[29]	[43]	58**
14.3	1,2,4,5-tetra-*O*-Ac 6-deoxy-3-*O*-Me hexitol	?	[28]	76
15.1	1,4,5-tri-*O*-Ac 2,3,6-tri-*O*-Me hexitol	[52]	[53]	71
15.7	1,2,5-tri-*O*-Ac 3,4,6-tri-*O*-Me hexitol	45	47	25
15.9	1,4,5-tri-*O*-Ac 2,3,6-tri-*O*-Me hexitol	[25]	[47]	67
16.3	1,4,5-tri-*O*-Ac 2,3,6-tri-*O*-Me hexitol	[81]	[25]	54
17.6	1,4,5-tri-*O*-Ac 2,3,6-tri-*O*-Me hexitol	[40]	[43]	55
19.6	1,4,5,6-tetra-*O*-Ac 2,3-di-*O*-Me hexitol	?	[6]	29
20.8	1,4,5,6-tetra-*O*-Ac 2,3-di-*O*-Me hexitol	?	[23]	86
	time of search/s	20—40	4	4—6

* – calculated for compound searched in the 1st position of the list of reference spectra; **– correct determination in the 2nd position; ? - library search failed; [ ] – incorrect interpretation; determination of more or less different structure.

be caused by the following factors: errors in nomenclature, inexact intensities of peaks in the reference spectra, incorrect rounding-off of *m/z* values in the measured spectra, incompleteness of the reference spectra, insufficient purity of compounds the reference spectra of which are comprised in the libraries, or influence of background. The output of the library search for the component of the mixture investigated (retention time = 11.8 min) is shown in Fig. 2. The results obtained are far from satisfaction

PBM Search of Library file: NBS43K, Scan 128 (11.802 min) of V3: METANA5*, E1

	Name	M_r	Formula	Similarity parameter/%
1.	D-Galactitol, 1,3,4,5-tetra-*O*-methyl, d	322	$C_{14}H_{26}O_8$	86
2.	Galactitol, 1,3,4,5-tetra-*O*-methyl, dia	322	$C_{14}H_{26}O_8$	86
3.	Mannitol, 1,3,4,5-tetra-*O*-methyl, diace	322	$C_{14}H_{26}O_8$	75
4.	D-Glucose, 2,3,4,6-tetra-*O*-methyl (8CI9	236	$C_{10}H_{20}O_6$	67
5.	D-Glucose, 2,3,4-tri-*O*-methyl (8CI9CI)	222	$C_9H_{18}O_6$	24
6.	D-Glucose, 2,4,6-tri-*O*-methyl (8CI9CI)	222	$C_9H_{18}O_6$	11

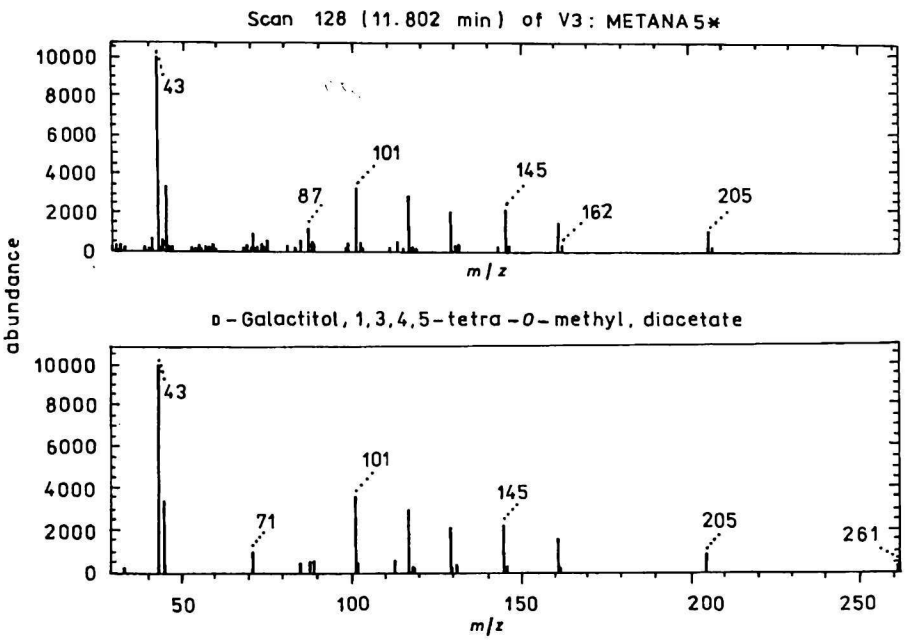

Fig. 2. Results of library search in the NBS43K library for the component having the retention time 11.8 min (correctly interpreted as 1,5-di-acetyl 2,3,4,6-tetra-*O*-methyl hexitol).

and correct interpretation cannot be achieved because of data errors in the library, *e.g.* incorrect numbering of substituent positions in the reference spectra, 1,3,4,5-tetra-*O*-Me instead of 2,3,4,6-tetra-*O*-Me.

The failure of the search by the use of commercially available libraries to distinguish and to interpret positional isomers prompted us to create a specialized library of mass spectra of partially methylated alditol acetates. The EI mass spectra [20] were reduced to 30 significant peaks, denoted by their *m/z* and intensity values. When necessary, the structurally not significant peaks (*e.g.* isotopic) in the lower part of the spectra have been omitted. As a rule, the most intensive peak in acetylated compounds (*m/z* = 43) is that of the CH_3CO^+ ions of acetyl (Ac) groups [1], and the intensities of fragments originating from the skeleton of the molecule are relatively low as compared with the peak at *m/z* = 43. Therefore, the peak intensities in mass spectra were assigned and calculated as follows: the intensity 100 per cent corresponds to the most intense fragment originating from the saccharide skeleton as well as the peak of Ac^+ ions with *m/z* = 43, if its measured intensity is at least 100 per cent of the base peak. The mass spectrum of 1,5-di-*O*-acetyl 2,3,4,6-tetra-*O*-methyl hexitol processed in this way is shown in Fig. 3. Such normalization of mass spectra emphasizes the most significant structural features typically found in the upper *m/z* range of a mass spectrum.

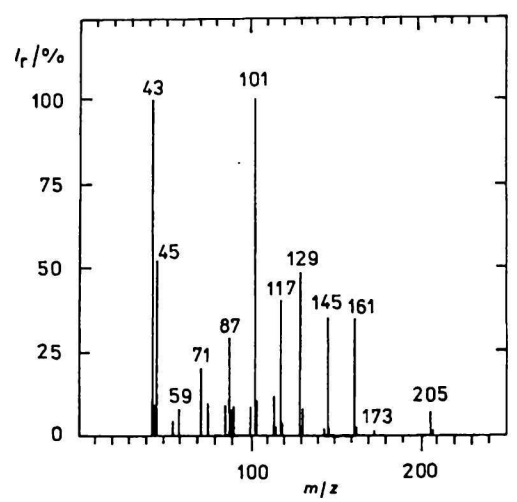

Fig. 3. EI mass spectrum of 1,5-di-acetyl 2,3,4,6-tetra-*O*-methyl hexitol, normalized in the way described in the text.

The results obtained by the use of this specialized library were satisfactory: of the 15 components in the mixture 13 compounds were correctly identified and matched in the 1st position of the list of reference spectra.

It is worth to mention that the use of library search method by way of the small, specialized library discussed here is significantly advantageous, because the shortcomings encountered with large data-bases are minimized.

The results of computer-assisted interpretation using the library PMAA were comparable with those obtained manually by the expert in the field. The conclusions drawn from the mutual comparison of the results suggest the usefulness and reasonability of development of specialized mass spectral data libraries, as for instance PMAA, which was used as successful tool for solving of structural problems of polysaccharides.

Acknowledgements. The authors thank Dr. P. Kováč, NIDDK, National Institutes of Health, Bethesda, for valuable comments.

REFERENCES

1. Biermann, Ch. J. and Ginnis, G. D., *Analysis of Carbohydrates by GLC and MS.* CRC Press, Boca Raton, Florida, 1989.
2. Chapman, J. R., *Computers in Mass Spectrometry.* Academic Press, London, 1978.
3. Martinsen, D. P. and Song, B. H., *Mass Spectrom. Rev. 4*, 461 (1985).
4. Small, G. W., *Anal. Chem. 59*, 535A (1987).
5. McLafferty, F. W. and Staufer, D. B., *Int. J. Mass Spectrom. Ion Process. 58*, 139 (1984).
6. Hertz, H. S., Hites, R. A., and Biemann, K., *Anal. Chem. 43*, 681 (1971).
7. Musil, S., Leško, J., Mocák, J., and Jurášek, P., *Chem. Listy 86*, 88 (1991).
8. Pesyna, G. M., Venkataraghavan, R., Dayringer, H. E., and McLafferty, F. W., *Anal. Chem. 48*, 1362 (1976).
9. Damen, H., Henneberg, D., and Weimann, B., *Anal. Chim. Acta 103*, 289 (1983).
10. Varmuza, K., Werther, W., Henneberg, D., and Weimann, B., *Rapid Commun. Mass Spectrom. 4*, 159 (1990).
11. Varmuza, K., *Trends Anal. Chem. 7*, 50 (1988).
12. Varmuza, K., Lohninger, H., and Werther, W., *Software Entwicklung in der Chemie,* Vol. 2. Springer-Verlag, Berlin, 1988.
13. Johnsen, S. and Kolset, K., *J. Chromatogr. 438*, 233 (1988).
14. Varmuza, K., *Lecture Notes in Chemistry (No. 21): Pattern Recognition in Chemistry.* Springer-Verlag, Berlin, 1980.
15. Fauset, D. W. and Weber, J. H., *Anal. Chem. 50*, 722 (1978).
16. Gray, N. A. B., Carhart, R. E., Lavanchy, A., Smith, D. H., Varkony, T., Buchanan, B. G., White, W. C., and Creary, L., *Anal. Chem. 52*, 1095 (1980).
17. Luinge, H. J., *Trends Anal. Chem. 9*, 66 (1990).
18. Wade, A. P., Crouch, S. R., and Betteridge, D., *Trends Anal. Chem. 7*, 358 (1988).
19. Hlavatý, J., Karacsonyi, Š., Kováčik, V., and Kakoniová, D., to be published.
20. Jansson, P. E., Kenne, L., Liedgren, H., Lindberg, B., and Lonngren, J., *Chem. Commun. Univ. Stockholm 8*, 1 (1976).

Translated by M. Ďurindová