

## Short Communication

---

# Application of topological indices to chromatographic data

## Calculation of the retention indices of anthocyanins

Dragan Amić and Dušanka Davidović-Amić

*Faculty of Agriculture, P.O. Box 117, 54001 Osijek (Croatia)*

Nenad Trinajstić\*

*Rugjer Bošković Institute, P.O. Box 1016, 41001 Zagreb (Croatia)*

(First received April 20th, 1993; revised manuscript received July 5th, 1993)

---

### ABSTRACT

The HPLC retention times ( $t_R$ ) of anthocyanins were calculated using a simple equation of the form  $t_R = a[W/(pn_{OH})] + b$ . The Wiener number,  $W$ , and polarity number,  $p$ , were used as structural descriptors. The other factor that controls the magnitude of the retention times is the number of OH groups ( $n_{OH}$ ) attached to the flavylium core. The calculated  $t_R$  values of several classes of anthocyanins are in excellent agreement with the experimental values.

---

### INTRODUCTION

The anthocyanins are part of the very large and widespread group of plant constituents known collectively as flavonoids. They are glycosylated polyhydroxy derivatives of 2-phenylbenzopyrylium or flavylium salts and are responsible for most of the red, pink, mauve and blue coloration of fruits, flowers and leaves [1]. Hundreds anthocyanins are known to exist in nature. Differences between individual anthocyanins occur in the pattern of hydroxylation

or methylation and the number and type of sugar(s) or acylated sugar(s) attached to the flavylium nuclei. The most common anthocyanidins (aglycones) are pelargonidin, cyanidin, peonidin, delphinidin, petunidin and malvidin. They differ only by the number and position of hydroxyl and/or methoxyl groups in the B-ring (see Fig. 1).

The importance of anthocyanins is reflected in their properties. They appear to be convenient food colouring materials [2–4], though their use is restricted to those foods and beverages whose pH is below 4. Additionally, they are substances with pharmaceutical properties [5–9].

Despite considerable efforts during the last

---

\* Corresponding author.

decade, qualitative and quantitative analysis of unknown mixtures of anthocyanins still remains difficult [10]. It is the structure and properties of the anthocyanins that make their separation difficult. The basic anthocyanidin skeletons of most of these pigments differ only little in molecular weight and substituents, making clearly defined separations of anthocyanin mixtures difficult. At present, the most satisfactory method for analysing mixtures of anthocyanins is the multistep method of quantification, separation and isolation by HPLC and peak identification by fast atom bombardment (FAB) MS and high-field NMR.

In this paper we investigate the relationship between HPLC retention indices and the structural properties of anthocyanins. Quantitative structure–chromatographic retention relationships (QSCRRs) with topological (graph-theoretical) indices represent a convenient model for studying the correlations between chromatographic and structural properties of molecules [11]. Topological indices in this approach act as molecular descriptors, which can be used to predict the HPLC retention indices via an appropriate statistical equation. Many graph–theoretical (topological) indices have been considered in the literature [12]. The Wiener number has been used extensively in structure–property–activity modelling. The Schultz index is another topological index that is successfully used in quantitative structure–property relationship (QSPR) modelling of properties of molecules.

It was shown by Hosoya [13] that the Wiener number [14],  $W = W(G)$ , of structure  $G$  can be defined as half of the sum of the elements of the distance matrix  $D(G) = D$ :

$$W(G) = \frac{1}{2} \sum_i \sum_j (D)_{ij} \quad (1)$$

where  $(D)_{ij}$  are off-diagonal elements of the distance matrix. The Wiener number appears to be a convenient measure of the compactness of the molecule [15].

The molecular topological index (MTI) or the Schultz index [16] is based on the adjacency matrix  $A$ , the distance matrix  $D$  and the valency matrix  $v$  of a molecule [17]:

$$MTI = \sum_i e_i \quad (2)$$

where  $e_i$  are the elements of the row matrix  $v [A + D]$ . Both Wiener number and Schultz index may be extended to heterosystems by replacing the elements of the adjacency matrix and of the distance matrix corresponding to heteroatoms and heterobonds with the values incorporating the corrections due to the changes induced when the carbon atom is replaced by the heteroatom [18].

Wiener also introduced the polarity number,  $p$ , which is equal to the number of pairs of atoms separated by three bonds [14,19]. The polarity number  $p = p(G)$  of a structure  $G$  is given by:

$$p = \frac{1}{2} \sum_i (p_3)_i \quad (3)$$

where  $p_3$  is the number of the paths of length 3 or the number of off-diagonal elements of  $D$  with the distance 3. Originally described as a polarity number,  $p$  is also related to steric aspects of a structure [20].

The topological indices used in this study were calculated using an Atari 1040ST computer and a modified computer program [21].

## RESULTS AND DISCUSSION

Structural characteristics of molecules are one of the most important factors responsible for molecular migration in the chromatographic process. The Wiener number is a molecular descriptor that reflects well the structure of molecule. The smaller the Wiener number, the larger the compactness of a structure (in terms of structural features such as branching and cyclicity). Hence, it can be reliably used for correlations with those physical and chemical properties that depend on the ratio of the volume to the surface of the molecule. Chromatographic retention data are typical molecular properties for which the OSPRs with the Wiener number yield correct predictions [22].

Experimental results [23,24] have shown that the polarity of the anthocyanins and anthocyanidins is most important factor affecting the HPLC retention times. Of the anthocyanidins, delphinidin was the first pigment

eluted, followed by cyanidin, petunidin, pelargonidin, peonidin and malvidin, showing the order of retention times to be in order of decreasing polarity of the compounds. Anthocyanins showed the same elution order. A topological index appropriate for modelling of molecular polarity is polarity number.

Of the individual substituents, OH groups have the greatest influence on retention times. As the number of hydroxyl groups on the B-ring increases, the retention time of the flavyliums decreases. Thus, delphinidin, with three hydroxy groups in the B-ring, is less retained than cyanidin with its two hydroxy groups. Methylation of the hydroxy groups increases the retention time: of the methylated flavyliums, petunidin is retained for a shorter time than peonidin. Malvidin shows the longest retention time. A reasonable explanation appears to be the formation of hydrogen bonds, which can greatly influence retention and separability. The number of hydrogen atoms bonded to oxygen atoms is a possible parameter describing hydrogen bonding. Thus, the more OH groups in the flavylium structure, the stronger the bonding with polar eluent and the smaller the retention time.

On the basis of the experimental results [23,24] we developed a very simple relationship between experimental HPLC retention times ( $t_R$ ) and Wiener number,  $W$ , polarity number,  $p$ , and number of OH groups,  $n_{OH}$ , in the flavylium structure:

$$t_R = a[W/(pn_{OH})] + b \quad (4)$$

where  $a$  and  $b$  are statistical parameters to be determined by least-squares regression.

Eqn. 4 indicates that  $t_R$  is proportional to the Wiener number. This is confirmed in the QSCRR studies between the  $t_R$  and the Wiener number for alkanes (*e.g.* ref. 25). However, the linear relationship between the  $t_R$  and  $W$  for anthocyanidins is rather poor ( $r = 0.4850$ ,  $s = 5.40$ ,  $F = 1.23$ ). Therefore, the Wiener number alone cannot be used to predict the  $t_R$  values of this class of compounds.

Experimental observation indicates that the  $t_R$  decreases with the polarity of the molecule. The simple rationale behind this observation is that

the polar molecule is better linked to a polar eluent and thus the retention time is shorter. We used the polarity number,  $p$ , as a measure of the polarity of the studied molecules. Thus, the  $t_R$  is inversely proportional to  $p$ . The linear relationship between  $t_R$  and  $W/p$  is again rather poor ( $r = 0.6003$ ,  $s = 4.94$ ,  $F = 2.25$ ).

It is also observed that the  $t_R$  decreases with an increase in the number of the OH groups on the flavylium core. This can be rationalized by considering the formation of the hydrogen bond between the flavylium salts and the polar eluent. The consequence of this is that more OH groups on the flavylium core result in stronger hydrogen bonding with the polar eluent and shorter retention time. This situation is introduced to the model through the number of the OH groups ( $n_{OH}$ ). Hence, the  $t_R$  is inversely proportional to  $n_{OH}$ . The linear relationship between  $t_R$  and  $W/n_{OH}$  is somewhat better than those above ( $r = 0.9247$ ,  $s = 2.35$ ,  $F = 23.59$ ), but still not particularly good. However, when we considered the linear relationship between the  $t_R$  and  $W/(pn_{OH})$ , the model improved considerably, as will be seen below. The point to note here is that  $W$  and  $p$  are not particularly intercorrelated quantities. For example, in the case of 159 undecanes, the correlation coefficient for the linear correlation between  $W$  and  $p$  is only 0.869.

We tested the validity of eqn. 4 for different types of flavylium salts: anthocyanidins, anthocyanin 3-glucosides and anthocyanin 3,5-diglucosides. The structures of the studied anthocyanidins, anthocyanin 3-glucosides and anthocyanin 3,5-diglucosides are shown in Figs. 1 and 2. Table I gives the Wiener numbers,  $W$ , Schultz indices, MTI, polarity numbers,  $p$ , and experimental ( $t_{R,exp}$ ) and calculated ( $t_{R,calc}$ ) retention indices. Experimental HPLC retention times for anthocyanidins were taken from a paper by Wilkinson *et al.* [23], and for anthocyanin 3-glucosides and for anthocyanin 3,5-diglucosides from a paper by Williams *et al.* [24].

The values of the statistical parameters are given in Table II. The statistical characteristics of all three correlations are of very good quality. Tables I and II show that the quality of fit between observed and calculated retention times is very good. This is confirmed by both high

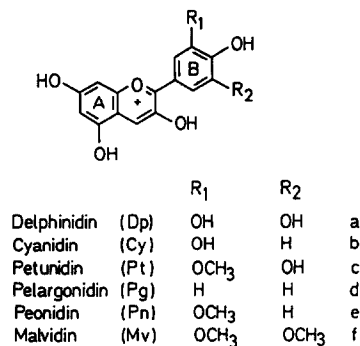


Fig. 1. Structures of the six most common anthocyanidins. Symbols in brackets represent the two-letter abbreviation of their names. The letters in the last column denote the B-ring substitution patterns.

correlation coefficients and the correctly predicted elution sequences. The high accuracy of the predicted models is also shown in Figs. 3–5,

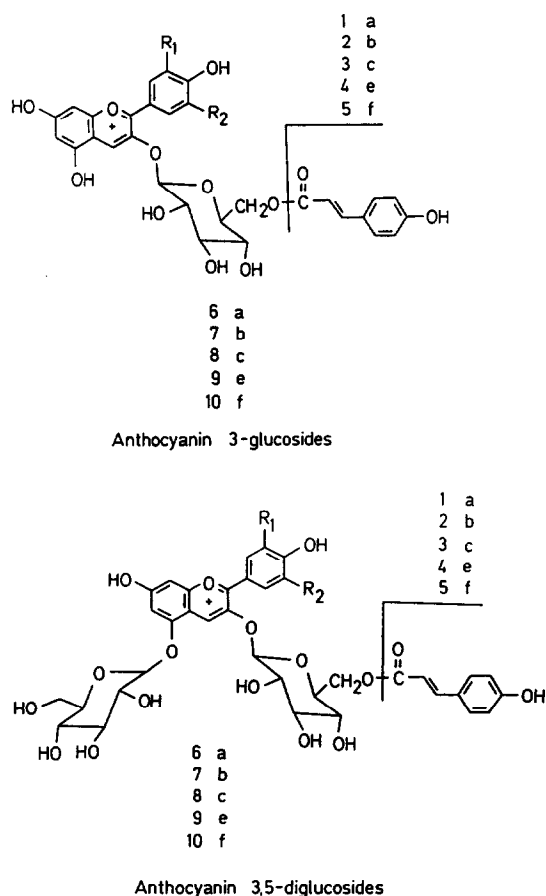


Fig. 2. Structural formulae of anthocyanins studied.

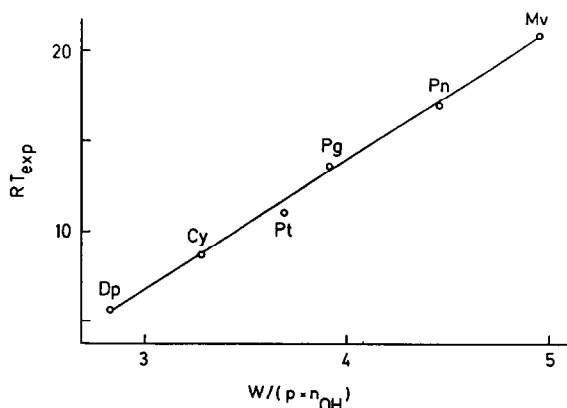


Fig. 3. Plot of  $t_{R,exp}$  ( $RT_{exp}$ ) vs.  $[W/(p-n_{OH})]$  for anthocyanidins.

in which the observed  $t_R$  vs.  $[W/(pn_{OH})]$  relationship are plotted.

The agreement between observed and calculated retention indices indicates that this model describes quite well the retention indices for a given class of flavylum salts: anthocyanidins, anthocyanin 3-glucosides and anthocyanin 3,5-diglucosides. There was a possibility of validating our model by comparison of measured retention times with calculated retention times. The retention times observed by Wulf and Nagel [26] for eleven acylated anthocyanin 3-glucosides correlate to a high degree with the retention times obtained using eqn. 4 ( $r = 0.9869$ ,  $s = 1.26$ ,  $F = 314.5$ ). Further, recently reported [27] retention times for six common anthocyanin 3-glucosides and their six corresponding malonyl esters are also quite well reproduced ( $r = 0.9848$ ,  $s = 1.07$ ,  $F = 320.8$ ).

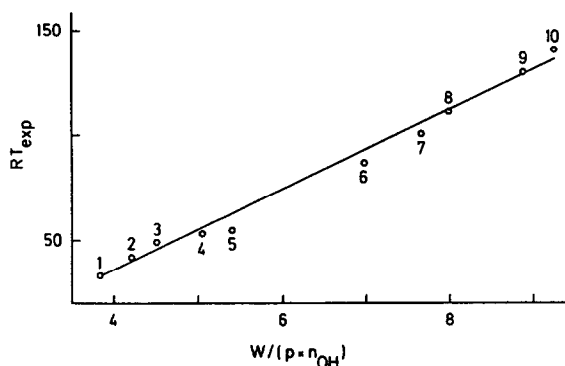


Fig. 4. Plot of  $t_{R,exp}$  ( $RT_{exp}$ ) vs.  $[W/(pn_{OH})]$  for anthocyanin 3-glucosides.

TABLE I

WIENER NUMBER, SCHULTZ INDEX POLARITY NUMBER,  $p$ , EXPERIMENTAL RETENTION INDICES AND CALCULATED RETENTION INDICES USING EQN. 4

Compound	$W$	MTI	$p$	$t_{R,exp}$	$t_{R,calc}$
<i>Anthocyanidins</i>					
1 Delphinidin (Dp)	663.02	2829.99	39	5.7	5.5
2 Cyanidin (Cy)	588.18	2531.71	36	8.7	8.7
3 Petunidin (Pt)	754.09	3195.28	41	11.1	11.6
4 Pelargonidin (Pg)	516.33	2243.92	33	13.6	13.3
5 Peonidin (Pn)	675.66	2884.17	38	17.0	17.1
6 Malvidin (Mv)	849.63	3577.04	43	20.8	20.7
<i>Anthocyanins</i>					
<i>Anthocyanin 3-glucosides</i>					
1 Dp 3-O-glucoside	2136.90	9025.11	62	33.2	31.5
2 Cy 3-O-glucoside	1986.87	8429.26	59	42.1	38.8
3 Pt 3-O-glucoside	2311.56	9723.13	64	49.3	44.6
4 Pn 3-O-glucoside	2157.94	9114.44	61	53.2	55.0
5 Mv 3-O-glucoside	2490.56	10436.98	66	55.1	61.4
6 Dp 3-O- <i>p</i> -coumarylglucoside	5369.47	22958.74	77	87.0	91.7
7 Cy 3-O- <i>p</i> -coumarylglucoside	5096.20	21860.23	74	100.9	104.7
8 Pt 3-O- <i>p</i> -coumarylglucoside	5675.61	24193.91	79	111.0	111.0
9 Pn 3-O- <i>p</i> -coumarylglucoside	5398.76	23082.57	76	129.9	128.2
10 Mv 3-O- <i>p</i> -coumarylglucoside	5986.09	25444.91	81	140.2	135.0
<i>Anthocyanin 3,5-diglucosides</i>					
1 Dp 3,5-diglucoside	4817.79	20300.56	85	22.0	23.0
2 Cy 3,5-diglucoside	4572.51	19323.26	82	26.0	27.8
3 Pt 3,5-diglucoside	5095.95	21414.52	87	30.1	31.3
4 Pn 3,5-diglucoside	4847.08	20424.39	84	37.8	37.4
5 Mv 3,5-diglucoside	5378.44	22544.31	89	42.1	41.1
6 Dp 3-O- <i>p</i> -coumarylglucoside-5-glucoside	9649.73	41069.65	100	64.5	60.0
7 Cy 3-O- <i>p</i> -coumarylglucoside-5-glucoside	9281.22	39589.70	97	70.0	67.5
8 Pt 3-O- <i>p</i> -coumarylglucoside-5-glucoside	10059.38	42720.76	102	72.0	70.9
9 Pn 3-O- <i>p</i> -coumarylglucoside-5-glucoside	9687.27	41227.98	99	79.1	80.1
10 Mv 3-O- <i>p</i> -coumarylglucoside-5-glucoside	10473.35	44387.71	104	79.1	83.7

In a recent report [28] on molecular descriptors derived from the distance matrix, it was found that the Wiener number,  $W$ , and the Schultz index, MTI, are strongly linearly correlated ( $r = 0.9999$ ) distance indices for alkanes. This result suggests the existence of a formal relation between  $W$  and MTI for alkanes. Such a relation was indeed found by Klein *et al.* [29]. We have also correlated the Wiener indices and the MTI for anthocyanins. The correlation coefficient obtained ( $r = 0.9999$ ) is indicative that in this case also there may exist a formal relation

between these two distance indices. The above result also suggests that the QSCRR with the MTI should be identical to those with the Wiener number. Some numerical experiments confirmed the above.

It was interesting to see how our model works in the case of some other classes of molecules. We tested eqn. 4 for calculating retention times of ten flavonoids [30]. If we consider that studied flavonoids possess different structures from anthocyanins, the correlation coefficient is satisfactory ( $r = 0.843$ ). Many anthocyanins contain

TABLE II

STATISTICAL CHARACTERISTICS OF THE RELATIONSHIP BETWEEN EXPERIMENTAL AND CALCULATED RETENTION TIMES OF FLAVYLIUMS USING EQN. 4

Parameter	Anthocyanidins	Anthocyanin 3-glucosides	Anthocyanin 3,5-diglucosides
<i>n</i>	6	10	10
<i>a</i>	7.1794 ± 0.1870	19.1354 ± 0.7039	13.6883 ± 0.5346
<i>b</i>	-14.7954 ± 0.7309	-41.7504 ± 4.6820	-41.6262 ± 3.7626
<i>r</i>	0.9986	0.9946	0.9940
<i>F</i>	1474.6	739.1	655.6
<i>s</i>	0.32	4.25	2.66

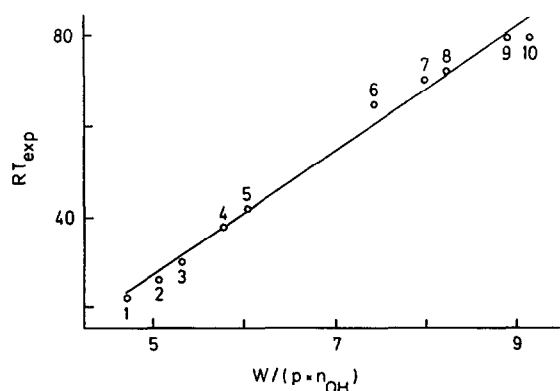


Fig. 5. Plot of  $t_{R,exp}$  ( $RT_{exp}$ ) vs.  $[W/(pn_{OH})]$  for anthocyanin 3,5-diglucosides.

acylated sugar moieties. The acyl groups are mostly derivatives of cinnamic acid. Our model works quite well in the case of eight cinnamic acid derivatives [31] ( $r = 0.972$ ).

Finally, an attempt was made to test the validity of our model in predicting the correct elution sequence for a mixture of twenty anthocyanin 3-glucosides and anthocyanin 3,5-diglucosides [24]. Unfortunately, it must be pointed out that in this case our model was less successful in reproducing the experimental elution sequence for all twenty anthocyanins ( $r = 0.825$ ).

## CONCLUSIONS

We have demonstrated that a very simple model, based on the structural properties of

molecules in terms of topological indices (Wiener number,  $W$ , and polarity number,  $p$ ) and number of OH groups in the flavylum structure (a possible parameter describing hydrogen bonding), can be used to predict successfully the HPLC retention indices of flavylum salts. With this model it is possible to reproduce quite accurately the experimental retention indices for different classes of flavylum salts such as anthocyanidins, anthocyanin 3-glucosides and anthocyanin 3,5-diglucosides.

## ACKNOWLEDGEMENT

This work was supported in part by Grant No. 1-07-159 awarded by the Ministry of Science of Croatia.

## REFERENCES

- 1 J.B. Harborne, *Comparative Biochemistry of the Flavonoids*, Academic Press, New York, 1967.
- 2 P. Markakis (Editor), *Anthocyanins as Food Colors*, Academic Press, New York, 1982.
- 3 G.A. Iacobucci and J.G. Sweeny, *Tetrahedron*, 39 (1983) 3005.
- 4 C.F. Timberlake and B.S. Henry, in V. Cody, E. Middleton, Jr., J.B. Harborne and A. Beretz (Editors), *Plant Flavonoids in Biology and Medicine*, Vol. II, Liss, New York, 1988, p. 107.
- 5 S.D. Varma and J.H. Kinoshita, *Biochem. Pharmacol.*, 25 (1976) 2505.
- 6 J.E. Ferrell, Jr., P.D.G. Chang Sing, G. Loew, R. King, J.M. Mansour and T.E. Mansour, *Mol. Pharmacol.*, 16 (1979) 556.
- 7 W.F. Hodnick, C.W. Bohmont, C. Capps and R.S. Pardini, *Biochem. Pharmacol.*, 36 (1987) 2873.

- 8 B. Gabetta, P. Morazzoni and G. Pifferi, *Planta Med.*, 56 (1990) 694.
- 9 G. Rastelli, L. Constatino and A. Albasini, *J. Mol. Struct. (Theochem)*, 276 (1993) 157.
- 10 G. Mazza and E. Miniati, *Anthocyanins in Fruits, Vegetables and Grains*, CRC Press, Boca Raton, FL, 1993.
- 11 R. Kaliszan, *Qualitative Structure Chromatographic Retention Relationships*, Wiley, New York, 1988.
- 12 N. Trinajstić, *Chemical Graph Theory*, CRC Press, Boca Raton, FL, 2nd ed., 1992, Ch. 10.
- 13 H. Hosoya, *Bull. Chem. Soc. Japan*, 44 (1971) 2332.
- 14 H. Wiener, *J. Am. Chem. Soc.*, 69 (1947) 17.
- 15 P.G. Seybold, M. May and U.A. Bagal, *J. Chem. Educ.*, 64 (1987) 575.
- 16 H.P. Schultz, *J. Chem. Inf. Comput. Sci.*, 29 (1989) 227.
- 17 W.R. Müller, K. Szymanski, J.V. Knop and N. Trinajstić, *J. Chem. Inf. Comput. Sci.*, 30 (1990) 160.
- 18 M. Barysz, G. Jashari, R.S. Lall, V.K. Srivastava and N. Trinajstić, in R.B. King (Editor), *Chemical Applications of Topology and Graph Theory*, Elsevier, Amsterdam, 1983, p. 222.
- 19 H. Wiener, *J. Am. Chem. Soc.*, 69 (1947) 2636.
- 20 J.R. Platt, *J. Phys. Chem.*, 56 (1952) 328.
- 21 J.K. Labanowski, I. Motoc and R.A. Dammkoehler, *Comput. Chem.*, 15 (1991) 47.
- 22 N. Bošnjak, Z. Mihalić and N. Trinajstić, *J. Chromatogr.*, 540 (1991) 430.
- 23 M. Wilkinson, J.G. Sweeny and G.A. Iacobucci, *J. Chromatogr.*, 132 (1977) 349.
- 24 M. Williams, G. Hrazdina, M.M. Wilkinson, J.G. Sweeny and G.A. Iacobucci, *J. Chromatogr.*, 155 (1978) 389.
- 25 N. Bošnjak, N. Adler, M. Perić and N. Trinajstić, in Z.B. Maksić (Editor), *Modelling of Structure and Properties of Molecules*, Wiley, New York, 1987, p. 103.
- 26 L.W. Wulf and C.W. Nagel, *Am. J. Enol. Vitic.*, 29 (1978) 42.
- 27 J.H. Kim, G.-I. Nonaka, K. Fujieda and S. Uemoto, *Phytochemistry*, 28 (1989) 1503.
- 28 Z. Mihalić, S. Nikolić and N. Trinajstić, *J. Chem. Inf. Comput. Sci.*, 32 (1992) 28.
- 29 D.J. Klein, Z. Mihalić, D. Plavšić and N. Trinajstić, *J. Chem. Inf. Comput. Sci.*, 32 (1992) 304.
- 30 K.H. Law and N.P. Das, *J. Chromatogr.*, 388 (1987) 225.
- 31 E. Idaka, T. Ogawa, T. Kondo and T. Goto, *Agric. Biol. Chem.*, 51 (1987) 2215.