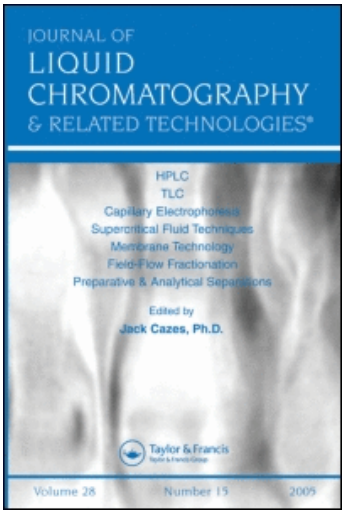

Journal of Liquid Chromatography & Related Technologies

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597273

QSRR of Flavones: Evaluation of Substituent Contributions to RP HPLC Retention

Marina Stefova[a]; Trajče Stafilov[a]; Svetlana Kulevanova[b]; Gjoshe Stefkov[b]; Vassya S. Bankova[c]
[a] Faculty of Science, Institute of Chemistry, Skopje, R. Macedonia [b] Faculty of Pharmacy, Institute of Pharmacognosy, Skopje, R. Macedonia [c] Institute of Organic Chemistry with Centre of Phytochemistry, Bulgarian Academy of Sciences, Sofia, Bulgaria

To cite this Article Stefova, Marina , Stafilov, Trajče , Kulevanova, Svetlana , Stefkov, Gjoshe and Bankova, Vassya S.(2007) 'QSRR of Flavones: Evaluation of Substituent Contributions to RP HPLC Retention', Journal of Liquid Chromatography & Related Technologies, 30: 8, 1035 — 1049
To link to this Article: DOI: 10.1080/10826070601128402
URL: http://dx.doi.org/10.1080/10826070601128402

Journal of Liquid Chromatography & Related Technologies®, 30: 1035–1049, 2007

ISSN 1082-6076 print/1520-572X online
DOI: 10.1080/10826070601128402

QSRR of Flavones: Evaluation of Substituent Contributions to RP HPLC Retention

Marina Stefova and Trajče Stafilov
Institute of Chemistry, Faculty of Science, Skopje, R. Macedonia

Svetlana Kulevanova and Gjoshe Stefkov
Institute of Pharmacognosy, Faculty of Pharmacy, Skopje, R. Macedonia

Vassya S. Bankova
Institute of Organic Chemistry with Centre of Phytochemistry, Bulgarian Academy of Sciences, Sofia, Bulgaria

Abstract: A non-parameter method based on the assumption of additive substituent increments on retention is employed for estimation of the effects of the substituents in the flavone ring on the retention in the reversed phase HPLC. The influence of OH- and OCH_3-groups in the positions: 3, 5, 6, 7, 8, 3′, 4′, and 5′ on reversed phase HPLC retention is studied in a group of 21 flavones.

The multiple linear regressions performed using 15 substituent codes as independent variables and the logarithm of the retention factor (log k) of each solute obtained with isocratic elution as dependent variables, gave a good correlation with R^2 from 0.9884 to 0.9984, and $\tilde{\mathbf{R}}^2$ from 0.9535 to 0.9937. The regression coefficients confirm that incorporation of an OH-group does not necessarily reduce retention, and OCH_3-groups, depending on the position, can either cause a decrease or increase in retention.

The proposed approach was used for prediction of the structure of an unknown flavone. The predicted substituent pattern of an unknown flavone found in an

Address correspondence to Marina Stefova, Institute of Chemistry, Faculty of Science, P. O. Box 162, 1001, Skopje, R. Macedonia. E-mail: marinaiv@iunona.pmf.ukim.edu.mk

extract of *Teucrium polium* is then experimentally confirmed by mass spectrometry and nuclear magnetic resonance of the isolated flavone–cirsiliol.

Keywords: Flavones, Quantitative structure-retention relationships, QSRR, Substituent contribution, Indicator variable, Multiple linear regression

INTRODUCTION

Quantitative structure chromatographic retention studies are directed towards two goals: obtaining more information about the chromatographic process by evaluating the dependence of retention on various structure descriptors of the solutes, and prediction of the retention of new solutes.

Free and Wilson[1] in 1964 introduced a general mathematical approach for quantitative evaluation of the substituents' additive effects in a series of analogue compounds on the biological activity, which is one of well known additive schemes. Accordingly, the molecules of the series were represented as products of substitution of one, parent, compound in different positions and the following series of linear equations was constructed:

$$\mathrm{BA}_i = \sum_j a_j X_{ij} + \mu$$

where BA_i is the parameter of biological activity of the compound; X_j is the indicator variable, which equals 1 if the j-substituent is present and 0 if not; a_j is the contribution of the j-substituent to BA, and μ is the mean overall activity of the whole series.

An analogous procedure for evaluation of the substituent effects on the chromatographic retention was used by Molnár and Horvát[2] and Chen and Horvát[3] on catecholamines and related compounds, and later by Gill et al.[4] on 2-phenylethylamines, assuming that the difference between the logarithm of the retention factor of the compound-i ($\log k_i$) and the one of the corresponding parent compound ($\log k_p$) can be written as follows:

$$\log r_{i,p} = \log k_i - \log k_p = \sum_{j=1}^{m} \tau_{i,j}$$

The same approach has also been used for a series of purine compounds[5] and on steroid hormones.[6] It is suitable for studying the effects of various substituents in different positions when a group of compounds can be derived from one "parent compound". The potential of the functional group contribution values for the understanding and for prediction of relative and absolute retention is thoroughly examined in the review of Smith.[7]

In this work, the non-parameter method, based on the assumption of additive substituent effects on retention, is employed for estimation of the effects of two different substituents (OH and OCH_3 groups) in eight possible positions in the flavone ring on the retention in the reversed phase

HPLC. Flavones are very convenient for studying the influence of substituents on retention because they form a group of compounds with the same "parent compound" flavone. The substituents are usually present in the positions 3, 5, 7, 3′, and 4′, and not so often 6, 8, and 5′ (Figure 1).[8] Several attempts have been made to correlate flavonoid structure to their chromatographic retention using different parameters for structure characterization, but all of them employ the parameter approach.[9,10]

In this work, the influence of OH- and OCH_3-groups in the positions 3, 5, 6, 7, 8, 3′, 4′, and 5′ of the flavone molecule on reversed phase HPLC retention is studied in a group of 21 flavones. The results obtained are used for evaluation of the structure of an unknown flavone found in an extract of *Teucrium polium*. The predicted substituent pattern of the unknown flavone is then experimentally confirmed by mass spectrometry and nuclear magnetic resonance of the isolated flavone, having the trivial name cirsileol.

EXPERIMENTAL

Reagents

The standards of flavones were mainly from Extrasynthèse, France; quercetin dihydrate from Merck, Germany; galangin from Aldrich, Germany. Small quantities of cirsimaritine, cirsilineol, xanthomicrol and 5,4′-OH, 6,7,8,3′-OCH_3 flavone were kindly donated by Dr. B. Voirin from Lion, France, whereas fisetin, 3,4′-dimetoxy kaempferol, 3,5,4′-trimetoxy kaempferol, 5-OH, 3,7,8,3′4′-OCH_3-flavone, and 7-OH, 5,8-OCH_3-flavone were supplied by Dr. V. Bankova from the Institute of Organic Chemistry with Centre of Phytochemistry of the Bulgarian Academy of Sciences. The substitution patterns of the 21 flavones of the studied group are presented in Table 1.

The HPLC retention data for the studied flavones were retrieved using isocratic elution with mobile phases (HPLC grade) consisting of: methanol

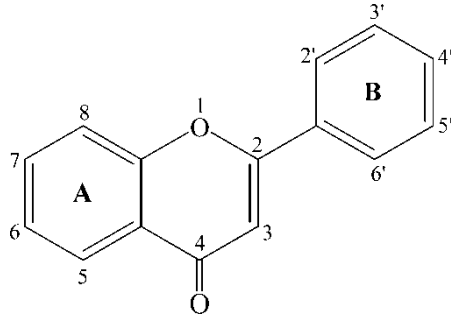

Figure 1. Structure of the "parent compound," flavone.

Table 1. Structures of the flavones used in the QSRR study

	Compound	Substituent in position							
		3	5	6	7	8	3′	4′	5′
1	Flavone								
2	Chrysine		OH		OH				
3	Apigenin		OH		OH			OH	
4	Genkwanin		OH		OCH_3			OH	
5	Acacetin		OH		OH			OCH_3	
6	Luteolin		OH		OH		OH	OH	
7	Chryseriol		OH		OH		OCH_3	OH	
8	Diosmetin		OH		OH		OH	OCH_3	
9	Galangin	OH	OH		OH				
10	Kaempferol	OH	OH		OH			OH	
11	Quercetin	OH	OH		OH		OH	OH	
12	Myricetin	OH	OH		OH		OH	OH	OH
13	Fisetin	OH			OH		OH	OH	
14	Dimethyl kaempferol	OCH_3	OH		OH			OCH_3	
15	Trimethyl kaempferol	OCH_3	OCH_3		OH			OCH_3	
16	7-OH, 5,8-OCH_3 Flavone		OCH_3		OH	OCH_3			
17	5-OH, 3,7,8,3′,4′-OCH_3 Flavone	OCH_3	OH		OCH_3	OCH_3	OCH_3	OCH_3	
18	Cirsimaritine		OH	OCH_3	OCH_3			OH	
19	Cirsilineol		OH	OCH_3	OCH_3		OCH_3	OH	
20	Xantomicrol		OH	OCH_3	OCH_3	OCH_3		OH	
21	5,4′-OH, 6,7,8,3′-OCH_3 Flavone		OH	OCH_3	OCH_3	OCH_3	OCH_3	OH	

and 5% water solution of acetic acid in different ratios (55; 60; 65; 70; 75, and 80% CH_3OH, v/v), acetonitrile and 5% water solution of acetic acid in different ratios (45; 55; 60, and 70% CH_3CN, v/v).

Instrumentation

A Varian HPLC system equipped with a ternary pump model 9012 and UV-diode array detector Polychrom model 9065 and a reverse phase column C_{18} (250 × 4.6 mm, 5 μm particles) from Varian were used. The flow rate was 1.0 mL/min and the temperature was set to 30ºC.

A semipreparative column LiChroprep® RP8 (40–63 μm) with dimensions 240 × 10 mm from Merck, Darmstadt, was used with the same HPLC system for isolation of the unknown flavone.

Mass spectra were obtained using a Hewlett Packard 5972 mass spectrometer system, 70 eV, and NMR spectra were recorded in DMSO-d_6 using the Brucker 250 (250 MHz for ^1H and 62.5 MHz for ^{13}C) instrument.

The software package STATISTICA for Windows 5.0 was used for performing the multiple linear regression, or more precisely linear regression of multiple variables, as well as for calculation of the statistical parameters for evaluation of the regression results.

Isolation of an Unknown Flavone

The unknown flavone was detected in the extract of *Teucrium polium* prepared by extraction of a dried and powdered plant material in 70% ethanol, filtration and then evaporation of the ethanol, and re-extraction of the aqueous phase in diethyl ether. This diethyl ether extract contains free flavone aglycones and is evaporated to dryness, the residue dissolved in methanol, filtrated, and analyzed by HPLC with UV DAD. In our previous investigations,[11,12] the flavones luteolin, apigenin, and cirsimaritin have been identified in the extracts of *Teucrium polium*. Another flavone, with an UV spectrum very similar to the one of cirsilineol, but with less retention time has been detected, and the lack of an authentic standard did not enable its identification. A semipreparative method for isolation of the unknown was developed using gradient elution with a mobile phase composed of solvent A (water acidified with phosphoric acid to pH = 3), B. acetonitrile, and C. methanol as follows: 0–5 min 70% A and 30% B (*V*/*V*); 10–20 min 65% A and 35% B; 25-30 min 50% A and 50% B, and from 35–45 min 100% C. The elution with this method was performed 20 times on extract volumes of 200 μL and the fractions containing the unknown flavone were collected and subjected to further spectroscopic methods (MS, NMR) for identification.

RESULTS AND DISCUSSION

The non-parameter approach is here employed for quantitative evaluation of the contribution of two substituents (OH- and OCH_3) in positions 3, 5, 6, 7, 8, 3′, 4′, and 5′ on the retention of flavones and flavonols. The position and type of substituent are designated as indicator variables, which can acquire values of **1** or **0**, for a certain substituent present in a particular position or not, respectively. Primarily, multiple linear regressions were performed with 13 indicator variables. It gave poor results for flavones containing neighboring OH groups in positions 3′ and 4′, and also for flavones containing three neighboring substituents in positions 6, 7, and 8. The presence of neighboring substituents in the aromatic ring (so called ortho-effect) was previously found to affect the retention,[7] which is the reason for introducing two additional indicator variables, one for the presence of neighboring OH-groups in positions 3′ and 4′, and one variable for the presence of three neighboring substituents in positions 6, 7, and 8. The substitution patterns of 21 flavones used in the QSRR study characterized with the 15 indicator variables are given in Table 2.

The multiple linear regressions were performed using the indicator variables as independent variables, and the logarithm of the retention factor (log k) of each solute obtained with isocratic elution with each of the 10 mobile phases as dependent variables. Experimental data for the log k values are presented in Table 3.

Statistica for Windows 5.0® was employed for calculation using the four parameters for evaluation of the statistical significance of the results. The multiple linear regression models gave a good correlation of the log k values with the 15 indicator variables, with the following values of the statistical parameters: coefficient of determination, R^2 from 0.9884 to 0.9984; adjusted coefficient of determination, $\boldsymbol{R^2}$ from 0.9535 to 0.9937, and F-test values in the range 28.361–212.16, t-test values >1.5, which confirms the existence of linear dependence.

The values of the regression coefficients are given in Table 4. They are a good indicator of the influence of the substituent in the different position on retention in the reversed phase HPLC. Namely, the intercept shows the logarithm of the retention factor for an unsubstituted flavone, whereas the regression coefficients for each indicator variable are a quantitative measure for the decrease (if negative) or increase (if positive) in retention caused by introduction of a hydroxy or methoxy group in the corresponding position. Using these regression coefficients, a logarithm of the retention factor of any flavone with a specific substitution pattern can be calculated just by summarizing the obtained regression coefficients of the unsubstituted flavone and the ones for the substituents present. A graph presenting the experimentally obtained versus the calculated log k values of the flavones used for establishing the model is presented in Figure 2.

Table 2. Indicator variables for the 21 flavones (numbered as in Table 1), where X1: 3-OH; X2: 3-OMe; X3: 5-OH; X4: 5-OMe; X5: 7-OH; X6: 7-OMe; X7: 3′-OH; X8: 3′-OMe; X9: 4′-OH; X10: 4′-OMe; X11: 5′-OH; X12: 3′,4′-OH; X13: 8-OMe; X14: 6-OMe; X15: 6,7,8-OMe

	Indicator variable														
Compound	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
3	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0
4	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0
5	0	0	1	0	1	0	0	0	0	1	0	0	0	0	0
6	0	0	1	0	1	0	1	0	1	0	0	1	0	0	0
7	0	0	1	0	1	0	0	1	1	0	0	0	0	0	0
8	0	0	1	0	1	0	1	0	0	1	0	0	0	0	0
9	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0
10	1	0	1	0	1	0	0	0	1	0	0	0	0	0	0
11	1	0	1	0	1	0	1	0	1	0	0	1	0	0	0
12	1	0	1	0	1	0	1	0	1	0	1	1	0	0	0
13	1	0	0	0	1	0	1	0	1	0	0	1	0	0	0
14	0	1	1	0	1	0	0	0	0	1	0	0	0	0	0
15	0	1	0	1	1	0	0	0	0	1	0	0	0	0	0
16	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0
17	0	1	1	0	0	1	0	1	0	1	0	0	1	0	0
18	0	0	1	0	0	1	0	0	1	0	0	0	0	1	0
19	0	0	1	0	0	1	0	1	1	0	0	0	0	1	0
20	0	0	1	0	0	1	0	0	1	0	0	0	1	1	1
21	0	0	1	0	0	1	0	1	1	0	0	0	1	1	1

Table 3. Experimental log k values obtained with isocratic elution with the corresponding mobile phases

	log k									
	Volume fraction of methanol						Volume fraction of acetonitrile			
	55%	60%	65%	70%	75%	80%	45%	55%	60%	70%
1	1,2150	1,0872	0,8819	0,6933	0,5152	0,3600	1,1198	0,7769	0,6742	0,4525
2	1,1337	0,9389	0,7782	0,5801	0,4038	0,2370	0,8366	0,4995	0,3691	0,1249
3	0,7012	0,5305	0,3897	0,2073	0,0459	−0,1108	0,3149	0,0274	−0,1011	−0,3471
4	1,2691	1,0648	0,8919	0,6834	0,4971	0,3201	0,8956	0,5399	0,4041	0,1536
5	1,2624	1,0565	0,8425	0,6390	0,4721	0,2826	0,8698	0,5209	0,3690	0,1050
6	0,4554	0,2940	0,1664	−0,0122	−0,1677	−0,3233	0,0669	−0,1731	−0,2789	−0,4585
7	0,7406	0,5658	0,4113	0,2073	0,0459	−0,1108	0,3798	0,0812	−0,0264	−0,2763
8	0,7990	0,6091	0,4152	0,2310	0,0908	−0,0874	0,3882	0,1027	−0,0228	−0,2979
9	1,2105	1,0090	0,8157	0,6058	0,4332	0,2347	0,8698	0,5209	0,3690	0,1050
10	0,6816	0,4987	0,3284	0,1375	0,0349	−0,1923	0,3447	0,0554	−0,0841	−0,3727
11	0,3937	0,2204	0,0554	−0,1222	−0,2821	−0,4347	0,0841	−0,1838	−0,2553	−0,4592
12	0,0756	−0,0816	−0,2346	−0,4017	−0,4848	−0,6825	−0,2218	−0,3305	−0,4175	−0,5363
13	0,1527	0,0335	−0,0902	−0,2118	−0,3468	−0,4958	−0,0844	−0,2457	−0,2889	−0,4787
14	0,9562	0,8181	0,6299	0,4397	0,2586	0,1018	0,5309	0,1939	0,1043	−0,0937
15	0,9563	0,8097	0,6318	0,4526	0,2655	0,0811	0,5304	0,2612	0,1230	−0,0772
16	0,6867	0,5760	0,4034	0,2189	0,1080	−0,0736	0,4333	0,1523	0,0553	−0,0888
17	1,3640	1,2098	1,0038	0,8115	0,6108	0,4153	1,2099	0,8553	0,6847	0,4159
18	0,8570	0,6999	0,5174	0,3425	0,1809	0,0300	0,6834	0,3514	0,2295	0,0121
19	0,8915	0,7304	0,5369	0,3638	0,1934	0,0366	0,7520	0,3978	0,2881	0,0806
20	1,0592	0,8979	0,6871	0,5073	0,3421	0,1748	0,8116	0,4144	0,3151	0,0948
21	1,1100	0,9413	0,7365	0,5359	0,3692	0,2039	0,8345	0,5209	0,3762	0,1626

Table 4. Regression coefficients obtained for each dependent variable[a]

Indicator variable	Substituent and position	Elution with a given volume fraction of									
		CH_3OH						CH_3CN			
		55%	60%	65%	70%	75%	80%	45%	55%	60%	70%
Intercept	None	1.2150	1.0872	0.8819	0.6933	0.5152	0.3600	1.1198	0.7769	0.6742	0.4525
X2	3-OCH_3	−0.2513	−0.1998	−0.1724	−0.1547	−0.1739	−0.1590	−0.2407	−0.2222	−0.1930	−0.1694
X3	5-OH	0.2718	0.2237	0.2011	0.1446	0.1219	0.1167	0.1599	0.0673	0.0219	0.0199
X4	5-OCH_3	0.1896	0.1573	0.1428	0.0906	0.0694	0.0633	0.0120	−0.0227	−0.0670	−0.0076
X5	7-OH	−0.3284	−0.3466	−0.2960	−0.2561	−0.2284	−0.2464	−0.4510	−0.3602	−0.3449	−0.3648
X6	7-OCH_3	0.2919	0.2317	0.2653	0.2886	0.2628	0.2382	0.1926	0.2296	0.2079	0.1691
X7	3′-OH	−0.4634	−0.4474	−0.4273	−0.4080	−0.3812	−0.3700	−0.4816	−0.4183	−0.3918	−0.4029
X8	3′-OCH_3	0.0672	0.0581	0.0576	0.0468	0.0299	0.0357	0.0867	0.1037	0.0910	0.0860
X9	4′-OH	−0.4821	−0.4584	−0.4364	−0.4208	−0.3829	−0.3839	−0.5275	−0.4814	−0.4640	−0.4732
X10	4′-OCH_3	0.1040	0.0922	0.0555	0.0572	0.0635	0.0523	0.0411	0.0369	0.0179	−0.0026
X11	5′-OH	−0.3490	−0.3388	−0.3455	−0.3345	−0.2599	−0.3035	−0.2973	−0.1521	−0.1504	−0.0775
X12	3′,4′-OH	0.2117	0.1987	0.1876	0.1798	0.1306	0.1445	0.2559	0.2372	0.2376	0.3096
X13	8-OCH_3	−0.3620	−0.3026	−0.3053	−0.2866	−0.2283	−0.2395	−0.1985	−0.1893	−0.1711	−0.1543
X14	6-OCH_3	−0.4559	−0.3980	−0.4136	−0.3759	−0.3447	−0.3156	−0.2703	−0.2696	−0.2267	−0.1650
X15	6,7,8-OCH_3	0.5723	0.5070	0.4899	0.4551	0.3968	0.3955	0.3038	0.2824	0.2579	0.2367
R^2		0.9952	0.9962	0.9941	0.9933	0.9938	0.9937	0.9915	0.9865	0.9931	0.9984

[a]The intercept is the value of log k for an unsubstituted flavone, whereas the log k value for any substituted flavone can be calculated by adding the values of the regression coefficients for the indicator variables for the corresponding substituents.

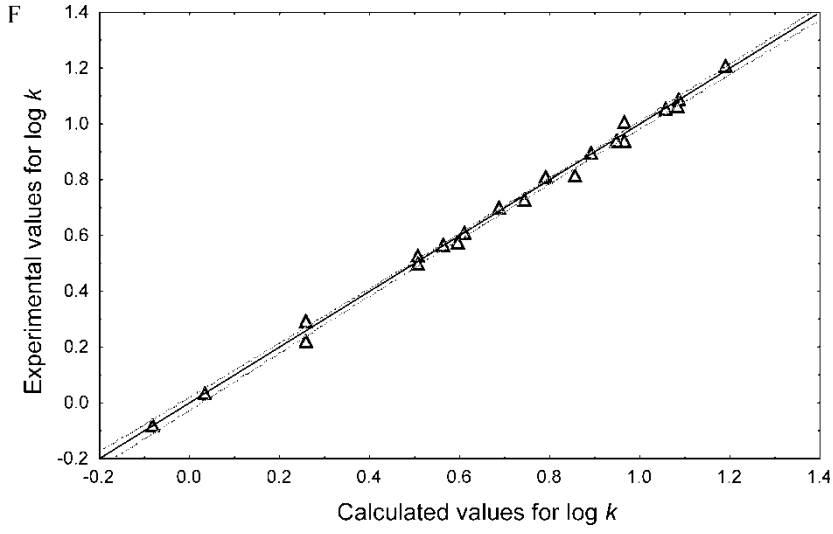

Figure 2. Correlation graph between the experimental and calculated values of log k obtained for isocratic elution with 60% (V/V) methanol.

A careful inspection of the obtained values within a series of substituted flavones leads to some interesting outcomes:

The indicator variable X1 can be excluded as statistically insignificant (<0.001), which means that the OH-group in position 3 in the flavone ring does not affect the retention owing to formation of an intramolecular hydrogen bond with the oxygen from the neighboring carbonyl group. This hydrogen bond is relatively weak owing to formation of a heterocyclic ring of five atoms, which decreases the influence of the hydroxyl group on the retention, but at the same time, does not significantly change the effect of the carbonyl group on retention. This is a quantitative proof for the experimental problem of separating the flavone-flavonol pairs: apigenin-kaempferol and luteolin-quercetin, differing only in the 3-OH group.[11]

The OH-group in position 5 leads to stronger retention, which is also due to formation of a hydrogen bond with the carbonyl oxygen, but this hydrogen bond is stronger because of geometrical reasons (a six membered heterocyclic ring is formed), which causes significant decrease of the influence of the carbonyl group and the 5-OH group and, thus, increased retention.

OH-groups in positions 7, 3′, 4′, and 5′ decrease retention on the reverse phase column owing to increasing polarity, according to the expectations. The obtained values of the regression coefficients imply that greater decrease in the retention is caused by introduction of 3′ or 4′ hydroxy group, but on the other hand, simultaneous presence of these groups has a net effect, which is smaller than the sum of their individual effects due to formation of an intramolecular

hydrogen bond (that is the reason for introducing the indicator variable for neighboring OH-groups in positions 3′ and 4′).

The negative regression coefficient for X2 implies that the OCH_3 group in position 3 decreases retention owing to its inductive effect, which increases the influence of the polar carbonyl group. Methylation of the hydroxyl group in position 3 results in a lack of hydrogen bonding between 3-OH and the carbonyl group. This leads to increased polarity of the molecule.

Substitution of methoxy groups in ring B and methylation of the hydroxy groups causes slightly increased retention. Still, such substitution in ring A does not give the same effect. As can be seen from the obtained regression coefficients, introducing a methoxy group in position 5 or 7 considerably increases retention, whereas in position 6 or 8 it produces a significant decrease in retention (comparable with the ones caused by an OH-group in position 3′ and 7, respectively), which has previously been observed.[13]

An interesting fact is that the introduction of a third methoxy group in ring A brings a net effect, which differs considerably from the sum of the effects of the individual groups. Namely, the presence of three methoxy groups in positions 6, 7, and 8, according to the obtained regression coefficients would produce a significant decrease in retention, but this does not agree with the obtained experimental values. For this reason, as previously mentioned, the indicator variable has been introduced for the simultaneous presence of three methoxy groups in ring A of the flavone structure. In this way, their net effect, which is a slight increase in retention, is satisfactorily explained by this model.

The obtained results are very useful for studies of flavones because they explain the retention as a function of the substitution pattern. As can be seen from the regression coefficients, introduction of the OH-group in the molecule does not necessarily bring decreased retention and OCH_3-group an increased retention, and their environment must always be considered for possible interactions. Similar qualitative data on the relationship between structure and retention of different classes of flavonoids have been previously reported by Bankova and co-workers.[13] Also, Dondi and co-workers have examined the effects of the substitutions of hydroxyl and glucosyl groups onto flavonoid structures.[14,15] They have found that $\Delta \log k'$ is largely independent of the stationary phase or of the acid used to alter the pH of the eluent and, for the substituents, that "the effects were not strictly constant and additive, but are complex results of intramolecular interactions".[15]

Testing of the Model for Prediction of the Substitution Pattern of an Unknown Flavone

This quantitative approach for the contribution of the substituents in various positions on retention is especially helpful in identification of flavones for

Table 5. Calculated and experimental values of log k for cirsilineol and the supposed substitution patterns of the unknown flavone

		60% Methanol		70% Methanol	
Compound	Substitution	Calculated	Experimental	Calculated	Experimental
Cirsilineol	5,4′-OH 6,7,3′-OCH_3	0,7442	0,7304	0,3765	0,3638
Cirsiliol	5,3′,4′-OH 6,7-OCH_3	0,4375	0,4977	0,1015	0,1079
	5,7,4′-OH 6,3′-OCH_3	0,1659		−0,1681	
		45% Acetonitrile		55% Acetonitrile	
Cirsilineol	5,4′-OH 6,7,3′-OCH_3	0,7611	0,7520	0,4264	0,3978
Cirsiliol	5,3′,4′-OH 6,7-OCH_3	0,4487	0,3392	0,1417	0,1086
	5,7,4′-OH 6,3′-OCH_3	0,1176		−0,1633	

which no authentic samples are available. In these cases, the first indication of a flavone structure is obtained by its UV-spectrum, which also implicates the probable substitution patterns.[16] The second step includes calculation of the log k values for all possible structures and comparison with the experimentally obtained data. The best match gives the most probable structure of the unknown analyte, whose positive identification must be made by further analyses (isolation and spectroscopy).

We have tested our model on an unknown flavone found in the extract of *Teucrium polium*. According to its UV-spectrum (very similar to the one of cirsilineol: 5,3′-OH, 6,7,4′-OCH_3 flavone) and retention (smaller than the one of cirsilineol), an analogous substitution pattern to the one of cirsilineol is supposed, but, with at least one OH-group more than cirsilineol. It is known that methylation of the OH-groups in positions 3, 5, and 4′ produces hypsochromic shifts of the characteristic absorption bands in the UV-spectra of flavones, but does not significantly affect the shape of the UV-spectra when it occurs in the other positions.[16] Applied to our problem, this means that in the structure of cirsilineol a methoxy group is substituted with a hydroxy group, producing a shorter retention of the unknown flavone. At this point, the previously described non-parameter approach for the quantitative effect of the substituent in the different position on retention in the reversed phase HPLC can be employed for prediction of the structure of the unknown flavone. Namely, the log k values are calculated from the obtained regression coefficients for the supposed two substitution patterns for flavones with OH-group instead of OCH_3-group in position 7 and 4′ (Table 5). Compared to the experimentally obtained log k for isocratic elution at the same conditions, it can be seen that the calculated value for the flavone with OH-group instead of OCH_3-group in position 4′, i.e., 5,3′,4′-OH, 6,7-OCH_3 flavone (trivial name cirsiliol), is in good agreement with the experimentally obtained ones.

The predicted structure of the unknown flavone was then confirmed by spectral analysis. The mass spectrum of the isolated compound was obtained after sylilation and the molecular mass (546) corresponds to a flavone with three [free] OH-groups and two OCH_3-groups. The ^{13}C NMR spectrum, compared to the literature data,[17] confirmed the presence of two OCH_3-groups in positions 6 and 7 (at δ 60.2 and 56.6, respectively). The signal for a 3′-OCH_3-group is missing (at δ 56.0), and the signal at δ 12.0 in the ^1H-NMR spectrum indicates the presence of a 5-OH group. This means that there are two methoxy groups in positions 6 and 7, OH-group in position 5, and the remaining two OH-groups are in positions 3′ and 4′ in the B ring of the flavone structure, and the unknown flavone is cirsiliol with the substitution pattern: 5,3′,4′-OH, 6,7-OCH_3 flavone.

The spectral confirmation of the predicted substitution pattern of the unknown flavone suggests that the proposed non-parameter approach satisfactorily describes the functional group contribution to retention. It

can be used for prediction of the retention of unknown substances belonging to a group of analogues, and *vice versa* for prediction of the possible substitution pattern of an unknown from its retention data, which can be a valuable tool in identifying substances for which authentic samples are not available.

ACKNOWLEDGMENTS

The authors are grateful for the financial support provided by Bulgarian Ministry of Education and Science and by the Ministry of Education and Science of the Republic of Macedonia in the frame of bilateral cooperation in the period of 2005–2006.

REFERENCES

1. Free, S.M., Jr.; Wilson, J.W. A mathematical contribution to structure-activity studies. J. Med. Chem. **1964**, *7*, 395–399.
2. Molnár, I.; Horváth, C. Catecholamines and related compounds. Effect of substituents on retention in reversed-phase chromatography. J. Chromatogr. **1978**, *145*, 371–381.
3. Chen, B.K.; Horváth, C. Evaluation of substituent contributions to chromatographic retention: quantitative structure-retention relationships. J. Chromatogr. **1979**, *171*, 15–28.
4. Gill, R.; Alexander, S.P.; Moffat, A.C. Group-contribution approach to the behaviour of 2-Phenylethylamines in reversed-phase high-performance liquid chromatography. J. Chromatogr. **1981**, *218*, 639–646.
5. Assenza, S.P.; Brown, P.R. Quantitative structure retention relationships for purine compounds on reversed-phase packings. J. Chromatogr. **1983**, *282*, 477–486.
6. Sadlej-Sosnowska, N. Structure retention relationship for steroid-hormones-functional-groups as structural descriptors. J. Liq. Chromatogr. **1994**, *17*, 2319–2330.
7. Smith, R.M. Functional group contributions to the retention of analytes in reversed-phase high-performance liquid chromatogr. J. Chromatogr. A **1993**, *656*, 381–415.
8. Wollenweber, E.; Dietz, V.H. Occurrence and distribution of free flavonoid aglycones in plants. Phytochemistry **1981**, *20*, 869–932.
9. Payares, P.; Díaz, D.; Olivero, J.; Vivas, R.; Gómez, I. Prediction of the gas chromatographic relative retention times of flavonoids from molecular structure. J. Chromatogr. A **1997**, *771*, 213–219.
10. Montaña, M.P.; Pappano, N.B.; Debattista, N.B.; Raba, J.; Luco, J.M. Prediction of the gas chromatographic relative retention times of flavonoids from molecular structure. Chromatographia **2000**, *51*, 727–35.
11. Stefova, M. PhD Thesis, University Sts Cyril and Methodius, Skopje, 2001.
12. Stefova, M.; Kulevanova, S.; Stafilov, T.; Stefkov, G. Quantitative determination of total flavones in *Teucrium polium* using HPLC-UV-DAD. Maced. Pharm. Bul. **2003**, *49*, 184–85.
13. Bankova, V.S.; Popov, S.S.; Marekov, N.L. High-performance liquid chromatographic analysis of flavonoids form propolis. J. Chromatogr. **1982**, *242*, 135–43.

14. Dondi, F.; Kahie, Y.D.; Lodi, G.; Blo, G.; Pietrogrande, C.; Reschiglian, P. Solvent selectivity effects in reversed-phase high-performance liquid chromatography of flavonoid compounds. J. Chromatogr. **1989**, *461*, 281–92.
15. Pietrogrande, M.C.; Reschiglian, P.; Dondi, F.; Kahie, Y.D.; Bertolasi, V. Correlations between high-performance liquid chromatographic retention, X-ray structural and ^{13}C NMR spectroscopic data of flavonoid compounds. J. Chromatogr. **1992**, *592*, 65–73.
16. Mabry, T.J.; Markham, K.R.; Thomas, M.B. *The Systematic Identification of Flavonoids*; Springer-Verlag: New York-Heidelberg-Berlin, 1970.
17. Zhang, Q.W.; Zhang, Y.X.; Zhang, Y.; Xiao, Y.-Q.; Wang, Z.M. Studies on chemical constituents in buds of *Artemisia scoparia*. Zhongguo Zhongyao Zazhi **2002**, *27* (3), 202–204.

Received December 12, 2005
Accepted October 14, 2006
Manuscript 6788