

# High-Performance Liquid Chromatography of Chalcones: Quantitative Structure-Retention Relationships Using Partial Least-Squares (PLS) Modeling

M. P. Montaña<sup>1</sup> / N. B. Pappano<sup>1</sup> / N. B. Debattista<sup>1</sup> / J. Raba<sup>2</sup> / J. M. Luco<sup>3\*</sup>

<sup>1</sup>Area de Química Física, <sup>2</sup>Area de Química Analítica, <sup>3</sup>Laboratorio de Alimentos, Facultad de Química, Bioquímica y Farmacia, Universidad Nacional de San Luis, Chacabuco y Pedernera 5700 San Luis, Argentina.

## Key Words

Column liquid chromatography  
Partial least squares (PLS) projections  
Chalcones  
Topological descriptors  
QSRR

## Summary

In this study, the multivariate partial least squares projections to latent structures (PLS) technique was used for modeling the RP-HPLC retention data of 17 chalcones, which were determined with methanol-water mobile phases of different compositions. The PLS model was based on molecular descriptors which can be calculated for any compound utilizing only the knowledge of its molecular structure. The PLS analysis resulted in a model with the following statistics:  $r = 0.976$ ,  $Q = 0.933$ ,  $s = 0.076$ , and  $F = 43.63$ . The adequacy of the developed model was assessed by means of cross-validation and also, by PLS modeling of the retention data of several chalcones reported by Walczak et al. [J. Chromatogr. 353, 123, (1986)], which were obtained using stationary phases of different polarity (-NH<sub>2</sub>, DIOL, -CN, ODS, C<sub>8</sub>). The structural interpretation of the developed PLS model was accomplished by means of comparative correlations between the nonempirical descriptors used in the model and the solvation parameters developed by Abraham. The results obtained in this work provides evidence for the great potential of the topological approach for the development of quantitative structure-retention relationship (QSRR) models.

## Introduction

Chalcones have been reported to display a wide variety of biological activities [1]. For example, several derivatives have shown antibacterial [2], antiviral [3], hepato

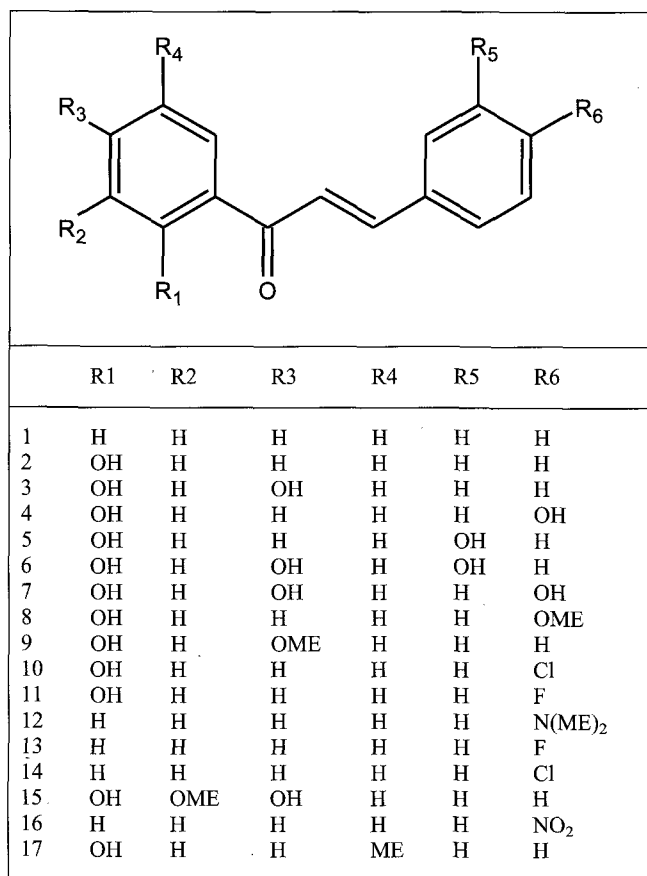
protectant [4], antimutagenic [5], gastric protectant [6], antiinflammatory [7], antileishmanial [8] and anticancer [9,10] activities.

Reversed-phase liquid chromatography (RPLC) has been widely recognized as a valuable alternative method to extract and quantify information about the structure and physicochemical properties of organic compounds [11], particularly hydrophobicity parameters that are extensively used in studies of quantitative structure-activity relationships (QSAR) [12-14]. In previous studies [15-17], hydrophobicity parameters obtained by RPLC as well as nonempirical descriptors based on the chemical graph theory, have been found useful in establishing quantitative structure-activity and structure-property relationships (QSAR, QSPR).

It is generally accepted that fortuitous or artifactual QSPR/QSAR models may be obtained when there exists multicollinearity among predictor variables and the models are derived by means of multiple linear regression (MLR) techniques. The disadvantage of the MLR method has recently been overcome through the development of a partial least squares (PLS) method [18] which circumvents the problems of collinearity among the variables and also offers the advantage of handling data sets where the number of independent variables is greater than the number of observations.

Thermodynamic parameters and magnitudes related to simple flavonoid structure, microbiological and chemical reactivity, and bioactivity have already been determined [19-25].

In the present study, the PLS technique was used for modeling the RPLC retention data of 17 *E-s-cis* chalcones, which were determined with methanol:water mobile phases of different compositions. The model here developed (PLS<sub>selec</sub>) was based on molecular descriptors which can be calculated for any compound utilizing only the knowledge of its molecular structure, and included several topological and constitutional descriptors as well as calculated hydrophobicity parameters. The adequacy of the PLS<sub>selec</sub> model was assessed by means of crossvalidation and also, by PLS modeling of the retention data of 24 (4,4')-*E-s-cis* chal-



**Figure 1**  
Structures of chalcones included in the study.

cones reported by Walczak and co-workers [26]. The data were obtained using stationary phases of different polarity (-NH<sub>2</sub>, DIOL, -CN, ODS, C<sub>8</sub>), and the new developed PLS model was based on the descriptor variables used in the PLS<sub>selec</sub> model. Finally, the structural interpretation of PLS<sub>selec</sub> was accomplished by means of comparative correlations between the selected non-empirical descriptors and the solvation parameters developed by Abraham [27], such as dipolarity/polarizability ( $\pi_2^H$ ), hydrogen-bond acceptor basicity ( $\Sigma\beta_2^H$ ), hydrogen-bond donor acidity ( $\Sigma\alpha_2^H$ ), and excess molar refraction ( $R_2$ ).

## Experimental

### Materials

The compounds described (1-17) were synthesized according to Claisen-Schmidt and Reichal-Müller procedures [28]. They were purified by LH 20 Sephadex column chromatography with methanol as eluent. The identities of the compounds were verified by UV, IR, and <sup>1</sup>H NMR spectroscopy. The structures are given in Figure 1.

## Instrumental

The experiments were performed with a Beckman (model 332) liquid chromatograph equipped with a variable wavelength detector (model 164) operated at 300 nm. The retention times were measured with a Varian 4290 integrator. A Phenosphere 5  $\mu$ m ODS-2 C<sub>18</sub> column (250 x 4.6 mm) was used in all experiments.

## HPLC Procedure

Chalcones were dissolved in the mobile phase at concentrations ranged from 0.05 mM to 0.12 mM, depending on the solubility of the compound. The chromatography was carried out at room temperature ( $26 \pm 1$  °C) and the injection volume was 50  $\mu$ L for all experiments. The flow-rate was 1 mL min<sup>-1</sup>, and the mobile phases consisted of different volume fractions of methanol and water. The column dead time ( $T_0$ ) was estimated from the retention time of deuteromethanol measured at 220 nm with methanol as eluent. The obtained retention data ( $Tr$ ) were measured at 0.05-increments of  $\phi$ MeOH in the range of  $0.50 \leq \phi\text{MeOH} \leq 0.90$  for compounds 3, 4, 5, 6, 7, 13, 15, and 16; and  $0.60 \leq \phi\text{MeOH} \leq 0.90$  for the other compounds. These data were used to derive the values of capacity factor ( $k$ ), which was calculated in the usual manner; that is,  $k = (Tr - T_0)/T_0$ . All capacity factors given represent the mean of 2-3 determinations of each sample solution. The reproducibility of retention times varied from 0.2%-1.6% (RSD) within a period of 6 weeks of experiments.

Finally, the values of  $\log kw$  and  $S$  were obtained by linear regression using Eq. (1):

$$\log k\phi = \log kw - S\phi \quad (1)$$

The correlation coefficients from  $\log k\phi$  vs.  $\phi\text{MeOH}$  regressions were  $>0.990$  for all compounds assayed. All chromatographic parameters included in the analysis are given in Table I.

## Structural Descriptors

The following indexes based on molecular topology were considered: the Wiener index [29], the valence and connectivity molecular indexes [30], the kappa shape indexes [31], and the charge and geometrical indexes recently introduced by Gálvez et al. [32,33]. The other group of variables included: molar volume, molecular weight, calculated octanol-water partition coefficient and several constitutional descriptors. To quantify the total H-bond capacity of the compounds under study, the parameters  $HBA$  and  $HBD$  as defined by Magee [34] were used. A list of all structural descriptors considered in this study is given in Table II.

## Statistical Methods

The PLS method was employed to search for relationships between the retention data and the structural de-

**Table I.** Chromatographic parameters for chalcones

Comp.	logk90 <sup>a</sup>	logk 85	logk 80	logk 75	logk 70	logk 65	logk 60	logk 50	logk <sub>w</sub>
1	-0.0132	0.1139	0.2689	0.4807	0.6665	0.8834	1.1004		3.47
2	0.2014	0.3608	0.5502	0.8048	1.0261	1.2777			4.27
3	-0.1675	-0.0996	0.0434	0.2601	0.4624	0.7020	0.8543	1.4041	3.25
4	-0.1549	-0.0605	0.0828	0.2636	0.4646	0.6808	0.8865	1.4186	3.17
5	-0.1457	-0.0339	0.1021	0.3107	0.4836	0.6942	0.9020	1.4473	3.16
6	-0.5017	-0.3768	-0.2757	-0.1051	0.0810	0.2559	0.4433	0.9843	2.98
7	-0.4881	-0.4034	-0.3098	-0.1278	0.0394	0.2583	0.4078	0.9244	2.90
8	0.1703	0.3314	0.5165	0.7563	0.9786	1.2394			4.18
9	0.1875	0.3551	0.5472	0.7627	1.0247	1.2889			4.31
10	0.2162	0.3997	0.6160	0.8316	1.1331				3.92
11	0.0673	0.2292	0.4133	0.6091	0.8780	1.0779			3.73
12	0.1072	0.2541	0.4273	0.6196	0.9390	1.0887			3.74
13	-0.0731	0.0682	0.2214	0.3953	0.6352	0.8225	1.0667	1.5861	3.56
14	0.0737	0.2343	0.4191	0.6138	0.8851	1.0955			3.77
15	-0.1675	-0.0630	0.0626	0.2095	0.5435	0.7288	0.9193	1.4672	3.63
16	-0.1163	0.0086	0.1492	0.3181	0.5575	0.7328	0.9687	1.4693	3.50
17	0.2330	0.4166	0.6191	0.8341	1.1191				4.15

<sup>(a)</sup> Percentage of MeOH in the mixture**Table II.** Symbols and definitions for the molecular descriptors applied in this study.

${}^m\chi^nP$	simple and valence ( $n = v$ ) path connectivity index of order $m = 0-6$
${}^m\chi^nC$	simple and valence ( $n = v$ ) cluster connectivity index of order $m = 3$ and $4$
${}^m\chi^nPC$	simple and valence ( $n = v$ ) path-cluster connectivity index of order $m = 4$
${}^m\kappa$	simple kappa shape index of order $m = 1-3$
${}^m\kappa\alpha$	alpha kappa shape index of order $m = 0-4$
${}^m\Delta x$	differential connectivity index of order $m = 0-3$ defined as ( ${}^m\chi - {}^m\chi^n$ )
${}^nG_k^{(a)}$	simple and valence ( $n = v$ ) charge index of order $k = 1-5$
${}^nJ_k^{(a)}$	simple and valence ( $n = v$ ) mean charge index of order $k = 1-5$ defined as $J_k = G_k / N - 1$
$W^{(a)}$	Wiener index
$L^{(a)}$	topological molecular length defined as the counted distance in the number of edges between the molecule's two most separate atoms by the shortest means
$S^{(a)}$	molecular surface parameter calculated as the sum of the predetermined $S$ values for several molecular fragments
$E^{(a)}$	shape index defined as $E = S / L^2$
$Pr_1^{(a)}$	number of pairs of ramifications separated by one edge
$Pr_2^{(a)}$	number of pairs of ramifications separated by two edges
$Pr_3^{(a)}$	number of pairs of ramifications separated by three edges
$V_m$	molar volume
$M_w$	molecular weight
$HBA^{(b)}$	count of electron pairs on O and N
$HBD$	count of O-H and N-H bonds
$\log P_{oct}$	calculated octanol-water coefficient

<sup>(a)</sup> Charge and geometrical indexes were calculated according to Gálvez et al. [32,33].<sup>(b)</sup> According to Magee [34] only one available acceptor pair is considered for each oxygen in the nitro group.

scriptors. This principal component-like method is based on the projection of the original multivariate data matrices down onto smaller matrices (T,U) with orthogonal columns, which relates the information in the

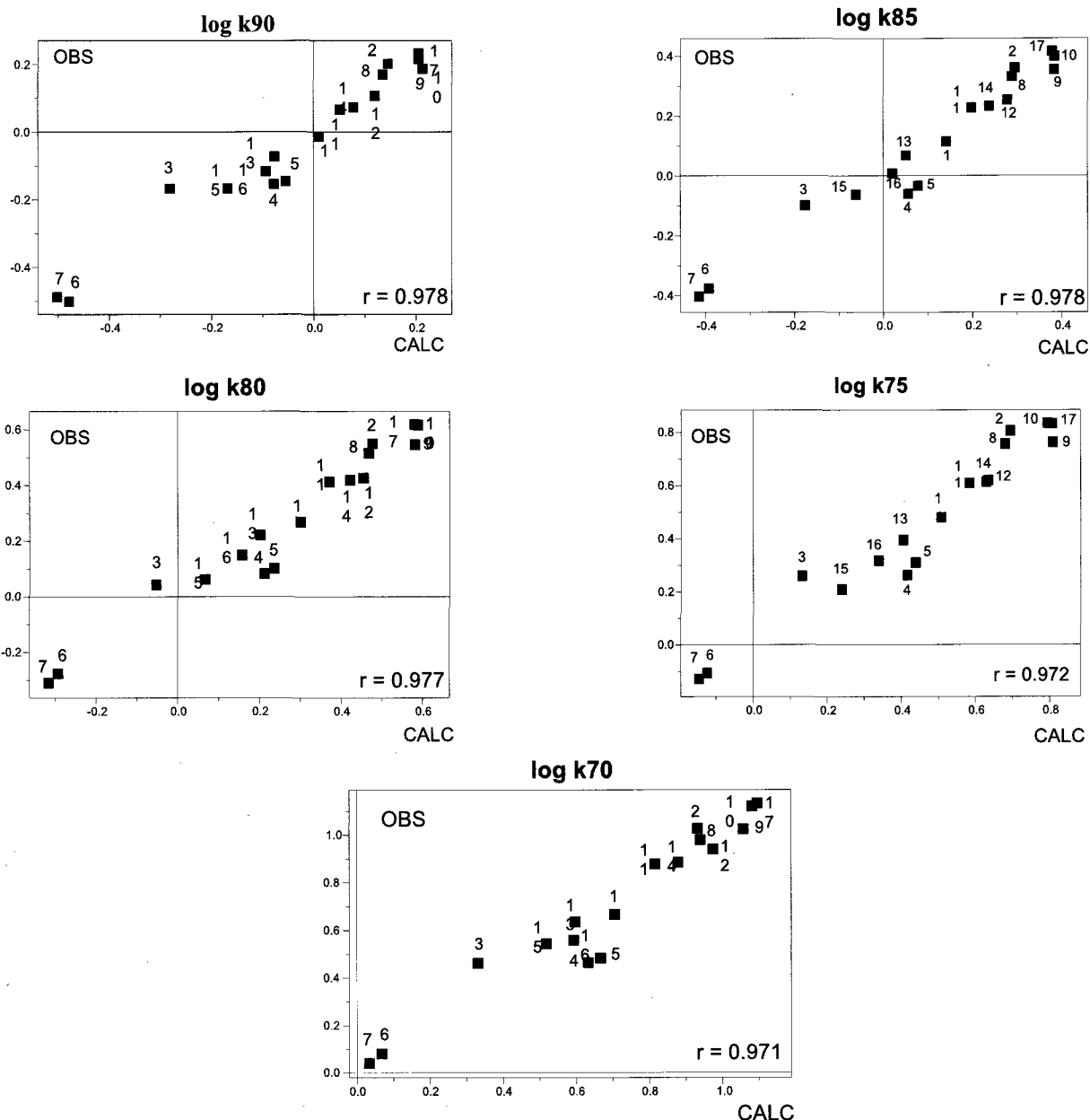
response matrix Y to the systematic variance in the descriptor matrix X, as shown below:

$$\begin{aligned} X &= \bar{X} + TP' + E \\ Y &= \bar{Y} + UC' + F \\ U &= T + H \text{ (the inner relation)} \end{aligned}$$

where  $\bar{X}$  and  $\bar{Y}$  are the corresponding mean value matrices, T and U are the matrices of scores that summarize the  $x$  and  $y$  variables respectively, P is the matrix of loadings showing the influence of the  $x$  in each component, C is the matrix of weights expressing the correlation between Y and T(X) and E, F, and H are the corresponding residuals matrices. The PLS calculations also give an auxiliary matrix W (PLS weights), which expresses the correlation between U and X and is used to calculate T. In the present work, the response matrix Y consisted of five dependent variables ( $\log k_{90}$ ,  $\log k_{85}$ ,  $\log k_{80}$ ,  $\log k_{75}$ ,  $\log k_{70}$ ) while the matrix X consisted of several structural descriptors. Determinations of the significant number of model dimensions was made by crossvalidation [35].

The computer software (Molconn-X) used to calculate the molecular connectivity-type topological indexes was obtained from L.H. Hall, Eastern Nazarene College, Quincy, MA, USA. Calculations of the charge and geometrical indexes were performed with the INDIS program kindly offered by Dr. Jorge Gálvez, University of Valencia, Spain. Molar volume and the octanol-water partition coefficient were calculated with the ACD/LogP 3.5 software obtained from Advanced Chemistry Development Inc. (Toronto, Canada).

PLS analysis was carried out using the SIMCA-S 5.1a software package obtained from Umetri AB, Box 7960, 907 19 Umea, Sweden.



**Figure 2**  
Relationships between experimental and calculated retention data for each of the dependent variables of chalcones included in the study.

## Results and Discussion

### PLS Analysis

All variables used in the PLS calculations were initially autoscaled to zero mean and unit variance to give each descriptor equal importance in the PLS analysis. The statistical significance of the screened models was judged by the correlation coefficient ( $r$ ), the standard deviation ( $s$ ) and the F-statistic. The predictive ability was evaluated by the crossvalidation coefficient ( $Q$ ) which is based on the prediction error sum of squares (PRESS). Because of the large number of structural descriptors considered in this study, the VIP (variable importance for the projection) parameter [35] was used

to unravel which descriptor variables were the most relevant to explain the response matrix  $Y$ .

After careful analysis of the obtained results using different combinations of the molecular descriptors shown in Table II, the PLS analysis resulted in a significant five-component model (PLS<sub>selec</sub>) with the following statistics:  $r = 0.976$ ,  $Q = 0.933$ ,  $s = 0.076$  and  $F = 43.63$ . The  $Q$  value was calculated using three crossvalidation groups and represent the highest obtained value. Figure 2 shows the correlation plots of observed versus calculated retention data for each of the modeled dependent variables. At this point, it is very important to indicate that it is not possible to obtain a suitable QSRR model (whether MLR or PLS derived) when  $\log kw$  is used as

**Table III.** Peak asymmetry factors for chalcones<sup>(b)</sup>

Comp.	90:10 <sup>(a)</sup>	85:15	80:20	75:25	70:30	65:35	60:40	50:50
1	1.0	1.0	1.0	1.1	1.2	1.4	2.0	
2	1.1	1.4	1.7	2.7	3.0	4.4		
3	1.0	1.1	1.2	1.3	1.4	1.6	2.0	2.5
4	1.0	1.0	1.0	1.2	1.3	1.3	1.4	1.8
5	1.0	1.0	1.0	1.1	1.2	1.2	1.4	2.5
6	1.1	1.2	1.2	1.3	1.4	1.5	1.6	1.8
7	1.0	1.0	1.0	1.0	1.1	1.2	1.3	1.8
8	1.1	1.2	1.3	1.7	2.0	2.4		
9	1.2	1.3	1.5	1.5	3.0	3.6		
10	1.0	1.1	1.4	2.7	2.8			
11	1.0	1.1	1.1	1.5	1.8	2.0		
12	1.0	1.2	1.0	1.0	1.2	1.4		
13	1.0	1.0	1.0	1.1	1.2	1.3	1.5	1.5
14	1.0	1.1	1.2	1.2	1.3	1.4		
15	1.0	1.0	1.2	1.6	2.0	2.7	2.8	4.2
16	1.0	1.0	1.0	1.0	1.0	1.0	1.2	1.4
17	1.1	1.3	1.5	2.5	3.0			

<sup>(a)</sup>Percentage of MeOH:H<sub>2</sub>O in the mixture. <sup>(b)</sup>Calculated to the 10 % of peak height.

**Table IV.** Pseudoregression coefficients of selected PLS model<sup>(a)</sup>

Descriptor	logk 90	logk 85	logk 80	logk 75	logk 70
log Poct	0.1085	0.1284	0.1506	0.1665	0.1626
<i>HBD</i>	-0.1490	-0.1621	-0.1793	-0.1856	-0.2157
$\Delta G_1^{(b)}$	0.0600	0.0659	0.0731	0.0769	0.0867
$\Delta G_2$	0.0875	0.0955	0.1059	0.1091	0.1278
$\Delta G_3$	-0.0737	-0.0790	-0.0871	-0.0908	-0.1086
$\Delta J_1^{(c)}$	0.0415	0.0468	0.0527	0.0571	0.0591
$\Delta J_2$	0.0872	0.0948	0.1049	0.1078	0.1273
$\Delta J_3$	-0.0857	-0.0922	-0.1019	-0.1060	-0.1266
$\chi^3_{VC}$	-0.1131	-0.1201	-0.1318	-0.1415	-0.1488
$\chi^4_{PC}$	0.0930	0.1003	0.1096	0.1070	0.1485
<i>L</i>	-0.0524	-0.0567	-0.0633	-0.0658	-0.0781
Constant	-0.0336	0.1020	0.2620	0.4610	0.7010

<sup>(a)</sup> Model obtained with five PLS components. <sup>(b)</sup> Charge indexes defined as the difference between the valence and simple terms (<sup>n</sup>G<sup>v</sup> - <sup>n</sup>G) or (<sup>n</sup>J<sup>v</sup> - <sup>n</sup>J) of order n = 1-3.

dependent variable. This fact is probably related to the specific influence exerted by the silanophilic interactions (present in all alkylsilane-bonded phases) on the retention process, particularly, when the mobile phase contains high water percentages. This effect is clearly made evident with chalcones having several hydroxyl groups in the molecule. This explanation, concerning to the basic assumption of a mixed retention mechanism due to both hydrophobic and silanophilic interactions, is supported by the observation that the peak-tailing phenomenon for polar chalcones becomes more pronounced with increasing the polarity of the mobile phase (see Table III). On the other hand, taking into account that all molecular descriptors are calculated for the uncharged molecule, it is reasonable to think that acceptable correlations can only be obtained when the ionization degree of the compounds plays a minor role in the retention process, which, in principle, would oc-

cur when mobile phases contain a high percentage of MeOH in the mixture.

Table IV shows the 11 selected descriptors and the corresponding pseudoregression coefficients. From these values, it can be seen how much a single variable contributes to the modeling of the retention data. According to these values, it can be inferred that, as expected, both the hydrogen-bonding capability of the compounds analyzed (encoded by *HBD*) and the hydrophobicity, expressed by the log Poct parameter, have a predominant role in the retention behavior of these compounds. It should be noted, however, that these parameters are not enough to account for the retention behavior of these compounds, which is made evident by the following equation:

$$\log k_{80} = -1.70 + 0.507 \log Poct - 0.147 HBD$$

$$r = 0.891 \quad s = 0.140 \quad n = 17 \quad F = 26.93 \quad (2)$$

**Table V.** Solvation parameters of compounds used in Eqs (3–5).

Compounds	Descriptors <sup>(a)</sup>		$\Sigma\alpha_2^H$
	$\Sigma\beta_2^H$	$\pi_2^H$	
cyclohexanone	0.56	0.86	0.00
benzene	0.14	0.52	0.00
toluene	0.14	0.52	0.00
hexylbenzene	0.15	0.50	0.00
naphthalene	0.20	0.92	0.00
chlorobenzene	0.07	0.65	0.00
methylphenyl ether	0.29	0.75	0.00
acetophenone	0.48	1.01	0.00
benzotrile	0.33	1.11	0.00
1,4-dinitrobenzene	0.46	1.63	0.00
acetanilide	0.67	1.40	0.50
benzamide	0.67	1.50	0.49
phenol	0.30	0.89	0.60
4-chlorophenol	0.20	1.08	0.67
3,5-dichlorophenol	0.00	1.17	0.77
4-iodophenol	0.20	1.22	0.68
3-trifluoromethyl phenol	0.09	0.87	0.72
cianophenol	0.29	1.63	0.79
benzyl alcohol	0.56	0.87	0.33
indazole	0.35	1.22	0.53
caffeine	1.33	1.60	0.00
dibenzothiophene	0.20	1.31	0.00
cortisone	1.87	3.50	0.36
hydrocortisone	1.90	3.49	0.71
ethylbenzene	0.15	0.51	0.00
<i>p</i> -xylene	0.16	0.52	0.00
propylbenzene	0.15	0.50	0.00
butylbenzene	0.15	0.51	0.00
<i>p</i> -dichlorobenzene	0.02	0.75	0.00
bromobenzene	0.09	0.73	0.00
nitrobenzene	0.28	1.11	0.00
<i>p</i> -nitrotoluene	0.28	1.11	0.00
methylbenzoate	0.46	0.85	0.00
benzophenone	0.50	1.50	0.00
3-phenylpropanol	0.67	0.90	0.30

<sup>(a)</sup>The solvation parameters were taken from refs [27] and [36].

Equations of similar statistical quality can be derived for the other chromatographic parameters analyzed in this work ( $\log k_{90}$ ,  $\log k_{85}$ ,  $\log k_{75}$ ,  $\log k_{70}$ ). Although the obtained correlation is statistically significant, the correlation coefficient and standard deviation values obtained for Eq. (2) are markedly lower than those obtained in the PLS<sub>selec</sub> model with the 11 descriptors shown in Table IV. According to the results shown in this table, a suitable description of retention behavior of the compounds under study requires, in addition to the above mentioned variables ( $\log P_{oct}$ ,  $HBD$ ), two groups of topological indexes; that is, the charge and geometrical indexes.

The charge indexes used in the PLS<sub>selec</sub> model are defined as the difference between the valence and simple terms ( ${}^nG^v - {}^nG$ ) or ( ${}^nJ^v - {}^nJ$ ) of order  $n = 1-5$ . According to Gálvez et al. [32,33], these indexes are information-rich structural descriptors encoding information about the charge distribution into the molecule. Thus, one way to assess the kind of information encoded by these indexes is to perform comparative correlations with the

solvation parameters developed by Abraham [27], such as the solute dipolarity/polarizability ( $\pi_2^H$ ), the solute overall or effective hydrogen-bond basicity and acidity ( $\Sigma\beta_2^H$ ,  $\Sigma\alpha_2^H$ ), and the excess molar refraction ( $R_2$ ). Table V reports the numerical values of the solvation parameters of compounds used as a reference to derive the corresponding comparative correlations [27,36]. For the 35 compounds shown in Table V, the corresponding charge indexes were calculated and the following regression equations were obtained:

$$\Sigma\beta_2^H = 0.15 + 0.0437 \Delta G_1 + 0.237 \Delta G_3 + 1.736 \Delta J_2 - 2.820 \Delta J_3 + 0.071 HBA + 0.107 HBD$$

$$r = 0.988 \quad s = 0.077 \quad n = 35 \quad F = 187.10 \quad (3)$$

$$\pi_2^H = 0.686 - 0.207 \Delta G_2 - 0.831 \Delta J_1 + 0.818 \Delta G_4 + 0.209 HBA + 0.200 HBD$$

$$r = 0.950 \quad s = 0.232 \quad n = 35 \quad F = 53.85 \quad (4)$$

$$\Sigma\alpha_2^H = 0.006 - 0.101 \Delta G_2 + 0.385 HBD$$

$$r = 0.907 \quad s = 0.129 \quad n = 35 \quad F = 74.86 \quad (5)$$

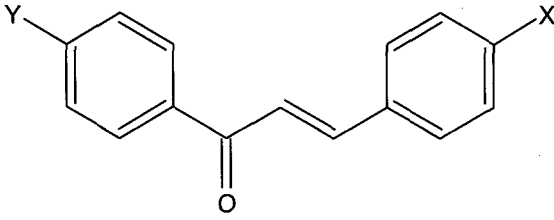
All the variables included in the developed equations are statistically significant ( $p < 0.05$ ). The correlation matrix showed that none of the charge indexes were highly correlated with any of the others. Taking into account the structural heterogeneity of the considered compounds (see Table V), Eqs. (3–5) indicate that the information contained in the charge indexes is largely electronic information.

Geometrical parameters require a separate discussion. The presence in the PLS<sub>selec</sub> model of the branching indexes, such as  ${}^3\chi^v_C$  and  ${}^4\chi^v_{PC}$ , suggest the influence of molecular stereochemistry in the retention process since both indexes carry information on degree of substitution, proximity of substituents, and length and heteroatom content in the rings. Finally, the presence of index  $L$ , the graph length, suggests the influence of steric effects on retention.

## Model Validation

It is well known that the real predictive ability of any QSAR/QSPR model cannot be judged solely by using internal validation techniques, such as crossvalidation. Therefore, a way of evaluating the PLS<sub>selec</sub> model consisted in the PLS modeling of the retention data reported by Walczak and co-workers [26] for a series of 24 (4,4')-E-*s-cis* chalcones, but using only the eleven selected descriptors shown in Table IV. The reported chromatographic parameters were obtained in normal-phase high-performance liquid chromatography (NPLC) by using five stationary phases of different polarity ( $-NH_2$ , DIOL,  $-CN$ , ODS,  $C_8$ ) and a mobile phase consisting of a *n*-heptane-tetrahydrofuran (97:3) mixture. Similarly to the analysis above described, matrix Y consisted of five dependent variables ( $\log k(NH_2)$ ,  $\log k(DIOL)$ ,  $\log k(CN)$ ,  $\log k(ODS)$  and  $\log k(C_8)$ ), while matrix X consisted of the same molecular descriptors used in the PLS<sub>selec</sub> model. Table VI shows the structures of the 24

**Table VI.** Capacity factors<sup>(a)</sup> for the (4,4')-E-*s-cis* chalcones studied by Walczak et al. [26].

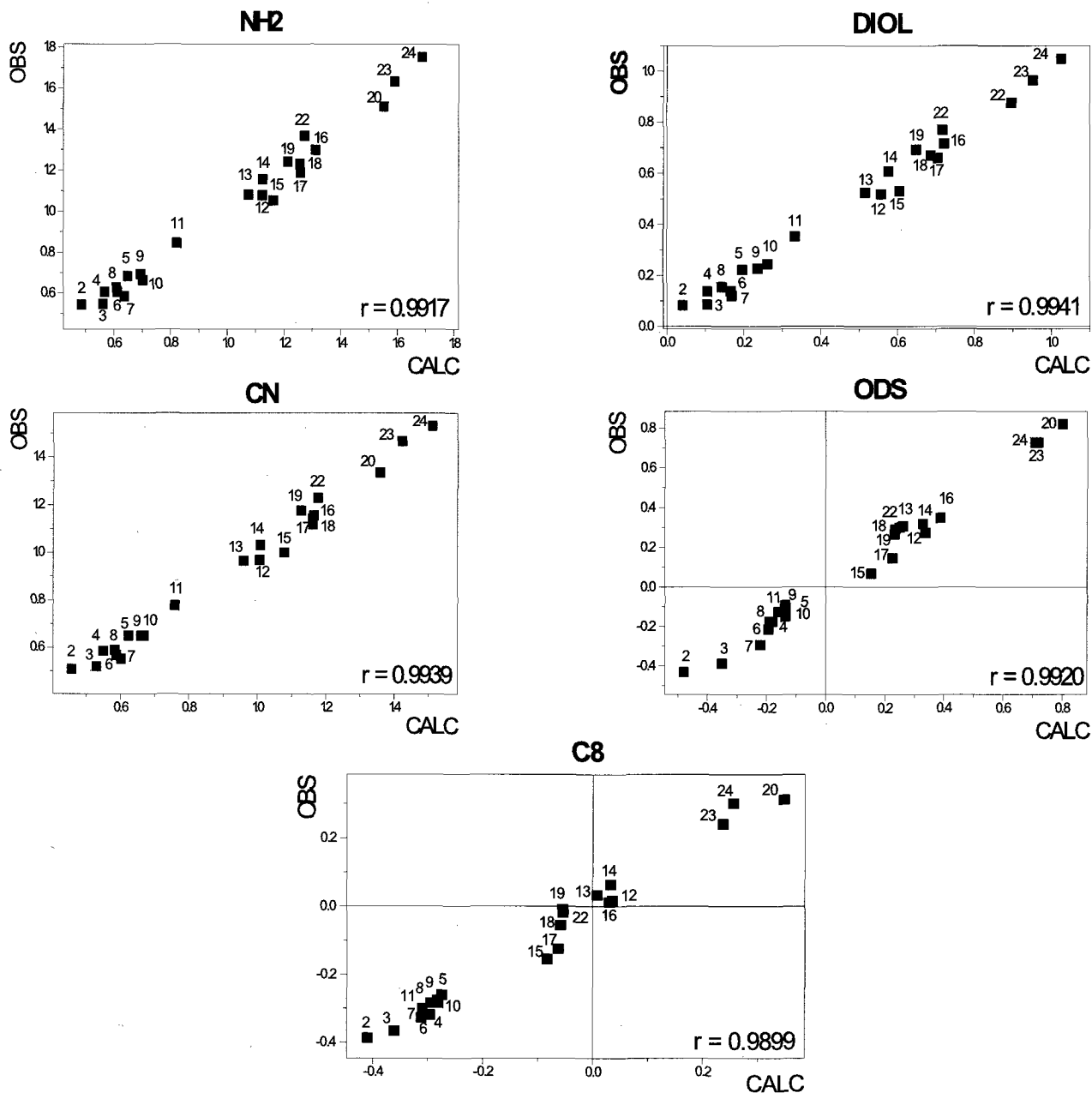
						
No	Chalcone (X-Y)	NH2	DIOL	CN	ODS	C8
1	H-CF <sub>3</sub>	3.39	1.14	2.99	0.22	0.29
2	H-tBu	3.49	1.21	3.22	0.37	0.41
3	H-iPr	3.52	1.22	3.31	0.41	0.43
4	H-H	4.04	1.37	3.85	0.67	0.52
5	F-H	4.82	1.67	4.47	0.78	0.55
6	H-F	4.03	1.38	3.70	0.61	0.48
7	H-Et	3.84	1.32	3.56	0.51	0.47
8	H-Me	4.25	1.43	3.88	0.67	0.52
9	F-Me	4.93	1.69	4.45	0.81	0.53
10	F-F	4.60	1.76	4.46	0.71	0.52
11	Me-Ph	6.98	2.25	5.98	0.75	0.50
12	MeO-Me	11.94	3.31	9.29	1.88	1.04
13	Me-MeO	12.07	3.35	9.22	2.03	1.08
14	F-MeO	14.29	4.06	10.74	2.09	1.16
15	H-NO <sub>2</sub>	11.29	3.40	9.99	1.17	0.70
16	MeO-Ph	19.84	5.19	14.32	2.23	1.03
17	F-NO <sub>2</sub>	15.40	4.57	13.14	1.40	0.75
18	NO <sub>2</sub> -Me	16.95	4.69	13.85	1.96	0.88
19	NO <sub>2</sub> -H	17.43	4.93	14.99	1.85	0.98
20	MeO-MeO	32.38	7.51	21.59	6.60	2.05
21	MeO-NO <sub>2</sub>	22.51	6.82	19.91	2.95	1.26
22	NO <sub>2</sub> -F	23.23	5.87	16.94	2.00	0.96
23	N(Me) <sub>2</sub> -NO <sub>2</sub>	42.83	9.20	29.35	5.33	1.74
24	NO <sub>2</sub> -MeO	56.67	11.21	34.02	5.33	2.00

<sup>(a)</sup> The columns used were: NH2 = Zorbax NH<sub>2</sub>; DIOL = LiChrospher 100 DIOL; CN = MicroPak CN; ODS = Zorbax ODS y C8 = Zorbax C<sub>8</sub>. The mobile phase used was a *n*-heptane-tetrahydrofuran (97:3) mixture.

**Table VII.** Pseudoregression coefficients of PLS model<sup>(a)</sup> for the retention data of chalcones studied by Walczak et al. [26].

Descriptor	NH2	DIOL	CN	ODS	C8
log Poct	-0.1451	-0.1237	-0.1401	-0.2266	-0.0827
HBA	0.1441	0.0973	0.1130	0.0891	0.1060
ΔG <sub>1</sub>	0.0566	0.0447	0.0438	0.1060	0.0587
ΔG <sub>2</sub>	0.0483	0.0332	0.0413	0.0562	0.0322
ΔG <sub>3</sub>	-0.0683	-0.0598	-0.0666	-0.0251	-0.0072
ΔJ <sub>1</sub>	-0.0077	-0.0096	-0.0155	0.0368	0.0347
ΔJ <sub>2</sub>	0.0823	0.0611	0.0726	0.1000	0.0502
ΔJ <sub>3</sub>	-0.0534	-0.0463	-0.0515	0.0070	0.0029
L	0.1991	0.1707	0.1795	0.2254	0.0765
Constant	0.9999	0.4705	0.9165	0.1114	-0.1122

<sup>(a)</sup> Model obtained with five PLS components.



**Figure 3**

Relationships between experimental and calculated retention data for each of the dependent variables of chalcones shown in Table VI.

chalcones with their corresponding retention values ( $k$ ) for each of the analyzed stationary phases. The capacity factor for compounds 1 and 21 were omitted from the PLS regression since, their  $\log k$  values were strongly overestimated for each of the dependent variables. The reason for this is not evident, especially taking into account that other structurally analogous compounds were well predicted by the selected model. The PLS analysis for the remaining 22 chalcones resulted in a significant five-component model with the following statistics:  $r = 0.993$ ,  $Q = 0.973$ ,  $s = 0.045$ ,  $F = 210.13$ . Figure 3 shows the correlation plots of observed versus calculated retention data for each of the modeled dependent variables. As can be noted, the agreement between observed and calculated values is very good. Ta-

ble VII shows the corresponding pseudoregression coefficients. As seen from this table, indexes  ${}^3\chi_C^V$  and  ${}^4\chi_{PC}^V$  were not included in the analysis. This is due to the fact that there is no variation in the substitution pattern of the analyzed chalcones, which implies that one of the basic structural features that encode these indexes is approximately constant. From the results in Table VII, it can be observed that the hydrophobicity ( $\log Poct$ ) and the ability of the chalcones to participate in hydrogen-bonding interactions ( $HBA$ ) are the major factors governing the retention of these compounds in NPLC. It should be noted, however, that the sign of coefficients for the  $\log Poct$  and  $HBA$  parameters is opposite with respect to the obtained in the PLS<sub>selec</sub> model. This implies that, in this chromatographic modality, increasing



the molecular polarity of a certain chalcone enhances the magnitude of the stationary phase-chalcone interaction (and retention) whereas, increasing solute bulkiness has the opposite effect. These results are in agreement with the findings by other authors with respect to the proposed retention mechanism for these stationary phases [37,38]. From a structural point of view, a further point to be stressed is that the influence of molecular length in NPLC is stronger than in RPLC. This is reflected in the high relative values of the coefficients of geometrical descriptor *L* for each of the analyzed stationary phases. Finally, according to the results shown in Table VII, the charge indexes serve as a fine-tuning in these multivariate relationships.

## Conclusion

The results obtained in this work clearly show the robustness and usefulness of the selected descriptors for accurately describe the retention of chalcones on different stationary-phase materials. In addition, it is important to highlight that this study provides evidence for the great potential of the charge topological indexes for the development of QSRR models.

## Acknowledgements

We wish to thank Dr. Jorge Gálvez for providing the software INDIS used for topological calculations. This work has been supported by the University National of San Luis, Argentina.

## References

- [1] J. B. Harborne, T. J. Mabry, H. Mabry, "The Flavonoids", Vol 1, Academic Press Inc., New York, 1975.
- [2] Y. B. Vibhute, S. S. Wadje, Indian J. Exp. Biol. **14**, 739, (1976).
- [3] Y. Ninomiya, N. Shimma, H. Ishitsuka, Antiviral Res. **13**, 61, (1990).
- [4] J. C. Davila, A. Lenherr, D. Acosta, Toxicology **57**, 267, (1989).
- [5] R. Edenharder, I. von Petersdorff, R. Rauscher, Mutat. Res. **287(2)**, 261, (1993).
- [6] S. Murakami, J. Pharm. Pharmacol. **42**, 723, (1990).
- [7] H. K. Hsieh, T. H. Lee, J. P. Wang, J. J. Wang, C. N. Lin, Pharm. Res. **15(1)**, 39, (1998).
- [8] S. F. Nielsen, S. B. Christensen, G. Cruciani, A. Kharazmi, T. Liljefors, J. Med. Chem. **41**, 4819, (1998).
- [9] R. De Vincenzo, G. Scambia, P. B. Panici, F. O. Ranelletti, G. Bonanno, A. Ercoli, F. D. Monache, F. Ferrari, M. Piantelli, S. Mancuso, Anticancer Drug Des. **10**, 481, (1995).
- [10] J. R. Dimmock, N. M. Kandepu, M. Hetherington, J. W. Quail, U. Pugazhenth, A. M. Sudom, M. Chamankhah, P. Rose, E. Pass, T. M. Allen, S. Halleran, J. Szydowski, B. Mutus, M. Tannous, E. K. Manavathu, T. G. Myers, E. De Clerq, J. Balzarini, J. Med. Chem. **41**, 1014, (1998).

- [11] R. Kaliszan in "High Performance Liquid Chromatography", P. R. Brown, R. A. Hartwick, Eds., vol 98 In "Chemical Analysis Series of Monographs", J. D. Winefordner, Y. M. Kolthoff, Eds., John Wiley & Sons, New York 1989, Chapter 14, p 563.
- [12] R. Kaliszan, J. Chromatogr. B, **717**, 125, (1998).
- [13] W. J. Lambert, J. Chromatogr. A, **656**, 469, (1993).
- [14] J. G. Dorsey, M. G. Khaledi, J. Chromatogr. A, **656**, 485, (1993).
- [15] J. M. Luco, M. E. Sosa, J. C. Cesco, C. E. Tonn, O. S. Giordano, Pestic. Sci. **41**, 1, (1994).
- [16] J. M. Luco, H. F. Ferretti, J. Chem. Inf. Comput. Sci. **37**, 392, (1997).
- [17] J. M. Luco, J. Chem. Inf. Comput. Sci. **39**, 396, (1999).
- [18] H. Kubinyi in "QSAR: Hansch Analysis and Related Approaches" VCH, Weinheim, R. Mannhold, P. Krosgaard-Larsen, H. Timmerman, Eds., VCH, Weinheim 1995, Chapter 5, p 91.
- [19] S. E. Blanco, N. B. Debattista, J. M. Luco, H. F. Ferretti, Tetrahedron Lett. **34**, 4615, (1993).
- [20] J. M. Luco, L. J. Yamin, H. F. Ferretti, J. Pharm. Sci. **84**, 903, (1995).
- [21] N. B. Debattista, N. B. Pappano, Talanta **44**, 1967, (1997).
- [22] M. P. Montana, N. B. Pappano, N. B. Debattista, Talanta **47**, 729, (1998).
- [23] N. B. Pappano, O. P. Centorbi, H. F. Ferretti, Rev. Microbiol. Sao Paulo **25**, 305, (1994).
- [24] N. B. Pappano, O. P. Centorbi, H. F. Ferretti, Rev. Microbiol. Sao Paulo **21**, 183, (1990).
- [25] N. B. Debattista, E. J. Borkowski, N. B. Pappano, J. Kavka, F. H. Ferretti, An. Asoc. Quím. Argent. **74**, 179, (1986).
- [26] B. Walczak, J. R. Chretien, M. Dreux, L. Morin-Allory, M. Lafosse, K. Szymoniak, F. Membrey, J. Chromatogr. **353**, 123, (1986).
- [27] M. H. Abraham, Chem. Soc. Rev. **22**, 73, (1993).
- [28] L. Reichel, K. Müller, Ber. **74B**, 1741, (1941).
- [29] R. Kaliszan, in "Quantitative Structure-Chromatographic Retention Relationships" J. D. Winefordner, Ed., John Wiley & Sons, New York, 1987, Chapter 8, p 138.
- [30] L. B. Kier, L. H. Hall, in "Molecular Connectivity in Chemistry and Drug Research", G. de Stevens, Ed., Academic Press, New York, 1976.
- [31] P. C. Jurs, S. L. Dixon, L. M. Egolf, in "Chemometric Methods in Molecular Design" H. van de Waterbeemd, Ed., VCH, Weinheim, 1995, Vol. 2, Chapter 2.1, p 15.
- [32] J. Gálvez, R. García-Domenech, M. T. Salabert, R. Soler, J. Chem. Inf. Comput. Sci. **34**, 520, (1994).
- [33] J. Gálvez, R. García-Domenech, V. de Julián-Ortiz, R. Soler, J. Chem. Inf. Comput. Sci. **35**, 272, (1995).
- [34] P. Magee, in "Classical and Three-Dimensional QSAR in Agrochemistry", C. Hansch, T. Fujita, Eds., American Chemical Society Symposium Series 606, Washington DC, 1995, Chapter 9, p.120.
- [35] S. Wold, in "Chemometric Methods in Molecular Design", H. van de Waterbeemd, Ed., VCH, Weinheim 1995, Chapter 4.4, p 195.
- [36] C. M. Du, K. Valko, C. Bevan, D. Reynolds, M. H. Abraham, Anal. Chem. **70**, 4228, (1998).
- [37] J. G. Dorsey, W. T. Cooper, Anal. Chem. **66**, 857, (1994).
- [38] J. Li, D. A. Whitman, Anal. Chim. Acta **368**, 141, (1998).

Received: Sep 28, 1999  
 Revised manuscript  
 received: Jan 3, 2000  
 Accepted: Jan 4, 2000