CROATICA CHEMICA ACTA
CCACAA **78** (2) 189–194 (2005)

ISSN-0011-1643
CCA-2997
Original Scientific Paper

A Chemical Graph Model Study of The Partition Coefficient of Halogenated Carbocompounds*

Lionello Pogliani**

Dipartimento di Chimica, Università della Calabria, 87036 Rende (CS), Italy
(E-mail: lionp@unical.it)

RECEIVED JANUARY 19, 2005; REVISED FEBRUARY 17, 2005; ACCEPTED FEBRUARY 18, 2005

A QSPR study of six sets of partition coefficients of a class of halogenated organic compounds has been performed with the aim to describe the six sets of partition coefficients with a common molecular connectivity descriptor, whose basic parameters can be derived by the aid of the three graph concepts: graph, pseudograph and complete graph. A simple odd complete graph conjecture gives rise to a set of four connectivity and pesudoconnectivity basis indices, which is, in fact, able to describe the six different cases of partition coefficients. This set of indices is, furthermore, the optimal set of indices for three cases. To underline are, also, the importance, as best single descriptor of the $^1\chi^v$ basis index, and the noticeable improvement the model undergoes when the only *cis*-compound is deleted from the model.

Keywords
chemical graphs
core electrons
partition coefficient
halogenated compound
molecular connectivity

INTRODUCTION

Recently, a complete K_p graph representation for the core electrons of any atom of any row with principal quantum number $n \geq 2$ has been successfully tested with many properties of different classes of compounds, inclusive inorganic compounds.[1–4] Till the introduction of complete graphs for the core electrons a two-valued algorithm[5,6] was used to derive the valence delta numbers, δ^v, which are the basic parameter for the valence molecular connectivity and the molecular pseudoconnectivity indices, inclusive their duals.[7–10] The interesting parallelism that has been detected between the old quantum and the complete graph algorithm allowed to improve the flexibility of the proposed δ^v.[4] The new K_p based algorithm for δ^v is a multivalued algorithm as different types of complete graphs can be considered. Up to date, only odd complete graphs showed an interesting model quality.[1–3] *i.e.*, K_p graphs with p = 1, 3, 5, …., while the search for properties which can be described with indices based on sequential K_p with p = 1, 2, 3, … is still open. Many years ago Cargas *et al.*,[11] published six sets of experimental values for the solvent/air and tissue/air partition coefficient, LogP, for twenty-five halogenated compounds (halocompounds), which can be used to check the complete graph conjecture for δ^v due to the variety of halogen derivatives of the twenty-five compounds. Furthermore, the model of LogP has always been a sort of »critical point« in model studies, as it is testified by some interesting studies,[12–15] including an interesting theorization about its high variability in different media.[16] The present work will not only try to test which

* Dedicated to Dr. Edward C. Kirby on the occasion of his 70th birthday.

** On sabbatical leave, present address: L. Pogliani, Unidad de Investigación en Diseño de Farmacos y Conectividad Molecular, Facultad de Farmacia, Dept. de Química Física, Av. V.A. Estellés s/n, 46100 Burjassot (València), Spain.

complete graph conjecture gives rise to the best model for some sets of LogP values, but also will try to test the reasonable guideline that similar properties of a common set of compounds should be described by a common descriptor.

METHOD AND ALGORITHMS

A complete graph, K_p, of order p is a graph where every pair of its vertices is adjacent.[17] A complete graph is always r-regular ($r = p - 1$), if it has all vertices with the same degree, r, this means that a regular graph is not always complete. In Figure 1 the chemical pseudograph plus complete graph of CH_3F, and CH_3Cl are shown. Here, the two carbon atoms are encoded with a K_1 complete graph, *i.e.*, a vertex, while the cores of F and Cl atoms are encoded with a K_1, and a K_3 complete graphs, respectively. Note that the dimensions of the K_3 vertex has no metrical meaning, it is just a zoom of this vertex to allow a better visualization, its dimensions are as meaningless as the dimensions of the K_1 vertices. The circle encircling the K_3 vertex has also no meaning it is just a frame for the overall multiple vertex. Note that all second row atoms are encoded with a K_1 vertex.[1–4] We remind that a graph is a set of vertices with a set of (single) edges that connect these vertices, while a chemical pseudograph allows for multiple edges and loops (*i.e.*, self-connections, mimicking non-bonding electrons). The degree of a vertex is the number of edges incident with it, self-connections contributing twice to the vertex degree. The three graph concepts of graph, pseudograph and complete graph allow to encode any type of hydrogen-suppressed (HS) organic molecule. The depleted hydrogens of a graph could be considered sort of null graphs, K_0, which possesses no points and no edges.[18] The following $\delta^v(K_p)$ algorithm, which encodes also the core electrons, is based on the two key parameters of complete graphs, p and r (order and regularity),

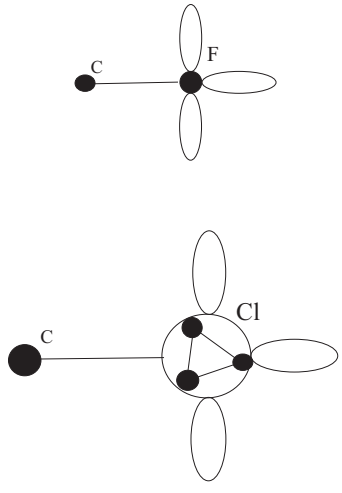

Figure 1. The chemical pseudograph-complete graph of CH_3F, and CH_3Cl, where the core electrons of C, F, and Cl are encoded with a K_1, K_1, and K_3 complete graphs, respectively.

$$\delta^v = q \cdot \delta^v(\mathrm{ps}) \,/\, [p \cdot r + 1] \quad (1)$$

Normally, two values are allowed for the q parameter, which may actually assume any value: $q = 1$ and $q = p$.[4] For second-row atoms: $\delta^v = \delta^v(\mathrm{ps})$, *i.e.*, it is the valence of a vertex in a pseudograph. Let us denote algorithm (1) for $q = 1 / p$, and $p = 1, 2, 3, 4,..$, with K_p-(p-seq) / K_p-(pp-seq), where seq means sequential, and for $q = 1 / p$ with $p = 1, 3, 5,..$, with K_p-(p-odd) / K_p-(pp-odd).

To model LogP use will be done of two subsets of basis indices, $\{\beta\} = \{\{\chi\}, \{\psi\}\}$, *i.e.*, the molecular connectivity indices, and the molecular pseudoconnectivity indices,[5–10]

$$\{\chi\} = \{\{D, {}^0\chi, {}^1\chi, \chi_t, D^v, {}^0\chi^v, {}^1\chi^v, \chi^v{}_t\},\ \{\psi\} = \{{}^S\psi_I, {}^0\psi_I, {}^1\psi_I, {}^T\psi_I, {}^S\psi_E, {}^0\psi_E, {}^1\psi_E, {}^T\psi_E\}\} \quad (2)$$

Due to the formal similarity between the χ and ψ indices, they are here collected in pairs,

$$D = \Sigma_i \delta_i \qquad {}^S\psi_I = \Sigma_i I_i \quad (3)$$

$${}^0\chi = \Sigma_i (\delta_i)^{-0.5} \qquad {}^0\psi_I = \Sigma_i (I_i)^{-0.5} \quad (4)$$

$${}^1\chi = \Sigma (\delta_i \delta_j)^{-0.5} \qquad {}^1\psi_I = \Sigma (I_i I_j)^{-0.5} \quad (5)$$

$$\chi_t = (\Pi \delta_i)^{-0.5} \qquad {}^T\psi_I = (\Pi I_i)^{-0.5} \quad (6)$$

Sums in Eqs. (3) and (4), as well as products (Π) in Eqs. (6), are taken over all vertices of the hydrogen-suppressed chemical graph. Sums in Eqs. (5) are over all edges of the chemical graph (σ bonds in a molecule). Replacing δ with δ^v the corresponding valence basis χ^v indices are obtained, and replacing I_i with S_i the corresponding ψ_E indices are obtained. Basis ψ indices are indirectly related to the δ^v number, defined in Eq. (1), through the I-State (ψ_I subset) and S-State (ψ_E subset) indices:[7]

$$I = (\delta^v + 1) / \delta,\ S = I + \Sigma \Delta I, \Delta I = (I_i - I_j) / r^2_{ij} \quad (7)$$

Here, r_{ij} counts the atoms in the minimum path length separating two atoms i and j. Factor $\Sigma\Delta I$ incorporates the information about the influence of the remainder of the molecular environment, and, as it can be negative, S can also be negative. To avoid imaginary ψ_E values, every S value of our classes of compounds has been rescaled to the S value of C in CF_4, *i.e.*, $S(C) = -5.5$. It has been shown elsewhere that the rescaling procedure has minor effects on the model.[8] The regression parameters and the different statistics have been obtained from the least-squares procedure of Excel. Here, descriptors and regression parameters of a linear combination are collected into a vector, as this will allow to write the model

equation as a dot product: $P = \boldsymbol{\beta} \cdot \boldsymbol{C}$., where $\boldsymbol{C} = (c_1, c_2, \ldots, c_0)$, and $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, U_0)$, while $U_0 \equiv 1$ is the unitary index, and c_0 is its regression parameter. The components of vector $\boldsymbol{\beta}$ are chosen by the aid of a combinatorial procedure, which searches the best basis β_i indices within the subsets of connectivity and pseudoconnectivity indices. To check how far the regression parameters, c_i, are meaningful, their utility will be given as $u_i = |c_i/e_i|$ where e_i is the estimated error of c_i. The u_i values, $\boldsymbol{u} = (u_1, u_2, \ldots, u_0)$, will be shown only for the best descriptor.

RESULTS AND DISCUSSION

The six sets of LogP values for twenty-five halocompounds are collected in Table 1. The original values are taken from Ref. 11. It will be tried to model them both with single molecular connectivity or pseudoconnectivity basis indices or with combinations of them. The basis indices, which give rise to the best descriptors are the ones, which encode the core electrons by the aid of the simple odd complete graphs conjecture, *i.e.*, in Eq. (1) we should have $q = 1$, and $p = 1, 3, 5$, for F, Cl, and Br, respectively. This representation is normally called the K_p-(p-odd) representation for the core electrons. Throughout the present model the following parameters will also be used: $D_F = \Sigma_i \delta_F$, and $D_{Cl} = \Sigma_i \delta_{Cl}$, over all F, and Cl atoms of a molecule. Actually, they are a count of the overall number of fluorine and/or chlorine atoms in a molecule.

Saline

The description of the saline partition coefficient is rather unsatisfactory, and the best single and multilinear descriptions are the following,

$\{{}^{T}\psi_I\}$: $F = 5.1$, $r = 0.427$, $s = 0.5$, $n = 25$
$\{{}^{0}\chi^{v}, {}^{1}\chi^{v}\}$: $F = 18$, $r = 0.785$, $s = 0.4$, $n = 25$
$\{{}^{0}\chi, {}^{0}\psi_I, D_F, D_{Cl}\}$: $F = 26$, $r = 0.917$, $s = 0.2_5$, $n = 25$

$\boldsymbol{C}$ = (–3.7718, 3.70213, 1.36892, 0.35451, 1.93033),
$\boldsymbol{u}$ = (9.3, 9.6, 7.0, 3.7, 6.7)

Note the ${}^{1}\chi^{v}$ index, which is, here, a critical index for the two-index combination. This index will be met frequently as best single descriptor throughout the model of

TABLE I. Liquid and rat tissue air partition coefficients, LogP, at 37 °C for twenty-five halocompounds (from Ref. 11)

Molecule[(a)]	L o g P					
	Saline	Ol. Oil	Blood	Liver	Mus(rat)	Fat(rat)
CH_3Cl	–0.056	0.933	0.393	0.540	–0.013	1.130
CH_2Cl_2	0.775	2.117	1.288	1.152	0.899	2.079
$CHCl_3$	0.529	2.604	1.318	1.324	1.143	2.307
CCl_4	–0.456	2.573	0.655	1.152	0.657	2.555
CH_2=CHCl	–0.367	1.387	0.225	0.204	0.342	1.301
CCl_2=CH_2	–0.456	1.808	0.699	0.645	0.312	1.836
CHCl=CHCl(c)	0.512	2.444	1.334	1.185	0.785	2.356
CHCl=CHCl(t)	0.149	2.250	0.981	0.952	0.547	2.170
CCl_2=CHCl	–0.081	2.743	1.340	1.435	1.004	2.744
CCl_2=CCl_2	–0.102	3.329	1.276	1.847	1.301	3.214
CH_3-CH_2Cl	0.037	1.590	0.611	0.558	0.508	1.587
$CHCl_2$-CH_3	0.389	2.270	1.049	1.033	0.709	2.215
CH_2Cl-CH_2Cl	1.057	2.563	1.483	1.553	1.369	2.537
CCl_3-CH_3	–0.125	2.470	0.760	0.934	0.498	2.420
$CHCl_2$-CH_2Cl	1.124	3.249	1.763	1.863	1.360	3.158
$CHCl_2$-$CHCl_2$	1.369	3.803	2.152	2.292	2.004	3.576
CCl_3-CH_2Cl	0.548	3.429	1.620	1.945	1.597	3.332
CH_2F_2	0.117	0.678	0.204	0.439	0.158	0.155
CH_2FCl	0.489	1.348	0.706	0.537	0.391	1.188
CH_2BrCl	0.937	2.558	1.618	1.465	1.045	2.512
CH_2Br_2	1.158	2.981	1.870	1.833	1.607	2.899
CF_3-CHClBr	–0.301	2.297	0.721	0.882	0.649	2.260
CH_2=CHBr	–0.357	1.748	0.607	0.522	0.354	1.692
CH_2Br-CH_2Cl	0.950	2.755	1.722	1.631	1.405	2.982
CF_3-CH_2Cl	–0.377	1.380	0.104	0.265	0.090	1.326

[(a)] c stands for *cis* and t for *trans*.

the following sets of partition coefficients. Keep also an eye to the four-index combination, which will reappear here end there.

Excluding *cis*-1,2-dichloroethylene the model improves slightly: $F = 27$, $r = 0.922$, $s = 0.2_5$, $n = 24$. A detectable F worsening is, instead, observed with the exclusion of the *trans*-compound (note that *cis* and *trans* compounds have the same set of basis index values): $F = 25$, $n = 24$.

Olive Oil

The model here improves a lot relatively to the saline case. The best single basis index is $^1\chi^v$, which plays no role in the best LCBI (Linear Combination of Basis Indices),

$\{^1\chi^v\}$: $F = 99$, $r = 0.901$, $s = 0.3_5$, $n = 25$
$\{D^v, {}^0\psi_I\}$: $F = 174$, $r = 0.970$, $s = 0.2$, $n = 25$
$\{^0\chi, D^v, {}^0\psi_I\}$: $F = 135$, $r = 0.975$, $s = 0.2$, $n = 25$,

$\boldsymbol{u}$ = (5.7, 2.1, 8.5, 5.5)
$\boldsymbol{C}$ = (–1.8722, 0.04176, 2.58623, 1.05843)

Deleting *cis*-1,2-dichloroethylene the F value deteriorates a bit: $F = 132$, $n = 24$, but deleting *trans*-1,2-dichloroethylene the F worsening grows: $F = 130$, $n = 24$.

Blood

Here we have the satisfactory four-index LCBI, which we already met for the saline case, and, again, the $^1\chi^v$ basis index as the best single descriptor,

$\{^1\chi^v\}$: $F = 27$, $r = 0.737$, $s = 0.4$, $n = 25$
$\{^1\chi^v, {}^T\psi_E\}$: $F = 37$, $r = 0.879$, $s = 0.3$, $n = 25$
$\{^0\chi, {}^0\psi_I, D_F, D_{Cl}\}$: $F = 38$, $r = 0.941$, $s = 0.2$, $n = 25$

$\boldsymbol{C}$ = (–2.60992, 2.92293, 0.70755, 0.14607, 1.37927),
$\boldsymbol{u}$ = (7.6, 9.1, 4.3, 1.8, 5.7)

Furthermore, the 5-indices LCBI, $\{D, {}^0\chi, {}^1\chi^v, D_{Cl}, S_F\}$, shows a very interesting model quality, with $F = 39$, $r = 0.954$, $s = 0.2$, $n = 25$, but five indices for twenty-five points is a bit too much. Excluding *cis*-1,2-dichloroethylene the model improves: $F = 41$, $r = 0.947$, $s = 0.2$, $n = 24$, deleting, instead, *trans*-1,2-dichloroethylene the F worsens a bit: $F = 37$, $n = 24$.

Liver

The following are the optimal descriptors for this case, where the $^1\chi^v$ index is, again, the best single basis index, and the 4-index combination has already be seen for the saline and blood cases,

$\{^1\chi^v\}$: $F = 48$, $r = 0.822$, $s = 0.3_5$, $n = 25$
$\{^0\chi, {}^0\psi_I\}$: $F = 46$, $r = 0.899$, $s = 0.3$, $n = 25$
$\{^0\chi, {}^0\psi_I, D_F, D_{Cl}\}$: $F = 52$, $r = 0.955$, $s = 0.2$, $n = 25$

$\boldsymbol{C}$ = (–2.4135, 2.74399, 0.70903, 0.22563, 1.10583),
$\boldsymbol{u}$ = (7.9, 9.5, 4.8, 3.2, 5.1)

Excluding either *cis*- or *trans*-1,2-dichloroethylene from the model F worsens a bit in both cases, *i.e.*, $F = 50$, $n = 24$.

Muscle

Here the best single-index is again the $^1\chi^v$ index, but the four-index combination is completely new, and includes, as well as the two-index combination the $^1\chi^v$ index,

$\{^1\chi^v\}$: $F = 44$, $r = 0.809$, $s = 0.3$, $n = 25$
$\{^1\chi^v, {}^1\psi_E\}$: $F = 40$, $r = 0.884$, $s = 0.3$, $n = 25$
$\{D, {}^0\chi^v, {}^1\chi^v, {}^1\psi_E\}$: $F = 47$, $r = 0.951$, $s = 0.2$, $n = 25$

$\boldsymbol{C}$ = (0.35769, –0.94262, 2.08192, –2.80391, –0.38763),
$\boldsymbol{u}$ = (4.9, 4.5, 7.3, 6.2, 2.8)

Excluding *trans*-1,2-dichloroethylene, this time, the model improves a bit: $F = 48$, $r = 0.954$, $s = 0.2$, $n = 24$, while excluding *cis*-1,2-dichloroethylene F deteriorates: $F = 44$, $n = 24$.

Fat Tissue

Here we have an optimal three-index combination, while the best single-index is always the $^1\chi^v$ index,

$\{^1\chi^v\}$: $F = 124$, $r = 0.918$, $s = 0.3$, $n = 25$
$\{D^v, {}^0\psi_I\}$: $F = 264$, $r = 0.980$, $s = 0.2$, $n = 25$
$\{^0\chi, {}^1\chi^v, {}^0\psi_I\}$: $F = 264$, $r = 0.987$, $s = 0.1$, $n = 25$,

$\boldsymbol{u}$ = (10, 3.5, 9.8, 4.5)
$\boldsymbol{C}$ = (–1.8722, 2.58623, 0.04176, 1.05843)

The best three-index linear combination for LogP-olive oil, $\{^0\chi, D^v, {}^0\psi_I\}$, is also a good descriptor, but with inferior quality, for LogP-Fat, with: $F = 210$, $r = 0.984$, $s = 0.2$, $n = 25$. Practically this linear combination is a rather good descriptor for $n = 50$ points. The contrary is not quite true: the linear combination for the fat tissue is not a quite good descriptor for LogP-olive oil.

Deleting the *cis*-compound the model improves more than a bit: $F = 298$, $r = 0.989$, $s = 0.1$, $n = 24$, while excluding the *trans*-compound F deteriorates: $F = 253$, $n = 24$.

An Overall Model for LogP

From what has been said the K_p-(p-odd) conjecture for the core electrons gives rise to a common optimal combination, $\{^0\chi, {}^0\psi_I, D_F, D_{Cl}\}$, for the saline, blood, and

liver cases, for a total of $n = 75$ points. In this combination the only basis index which is able to differentiate among the different K_p representations is the $^0\psi_I$ index, as the other indices are based only on δ and not on δ^v. Further, while the $^0\psi_I$ and $^0\chi$ indices are somewhat interrelated[19] ($r = 0.94$), the interrelations among the other indices of this combination are $r \leq 0.74$, *i.e.*, practically, our indices are orthogonal to each other. Now, this vector, $\boldsymbol{\beta} = (^0\chi, {}^0\psi_I, D_F, D_{Cl}, U_0)$ is also able to model in a satisfactory way the three remaining cases, olive oil, muscle, and fat for a total of $n = 150$ points. The statistics of the six cases all together are,

Saline: $F = 26$, $r = 0.917$, $s = 0.2_5$, $n = 25$,
$\boldsymbol{C} = (-3.7718, 3.70213, 1.36892, 0.35451, 1.93033)$

Blood: $F = 38$, $r = 0.941$, $s = 0.2$, $n = 25$,
$\boldsymbol{C} = (-2.60992, 2.92293, 0.70755, 0.14607, 1.37927)$

Liver: $F = 52$, $r = 0.955$, $s = 0.2$, $n = 25$,
$\boldsymbol{C} = (-2.4135, 2.74399, 0.70903, 0.22563, 1.10583)$

Olive Oil: $F = 99$, $r = 0.976$, $s = 0.2$,
$\boldsymbol{C} = (-1.8217, 2.56576, 0.29914, 0.06943, 1.06648)$

Muscle: $F = 41$, $r = 0.944$, $s = 0.2$,
$\boldsymbol{C} = (-2.34576, 2.69296, 0.68124, 0.14847, 0.86830)$

Fat: $F = 150$, $r = 0.984$, $s = 0.1_5$,
$\boldsymbol{C} = (-1.62782, 2.47560, 0.12546, -0.0317, 0.80659)$

Four of the six cases, liver, oil, muscle, and fat, show a very satisfactory model. Now, five compounds chosen at random (CH_2=CHCl, CH_2Cl–CH_2Cl, CCl_3–CH_2Cl, CH_2Br_2, and CF_3–CH_2Cl), will be left-out from the model and the model equation derived by the aid of twenty points will be used to externally model the LogP of these five compounds for each of the four cases, liver, muscle, oil, and fat. This means that twenty points will be externally validated. For the twenty points of saline, blood and liver cases this combination, continues to be the optimal combination. The model, for the hundred points only is shown in Figure 2, where the twenty externally validated points are in open squares.

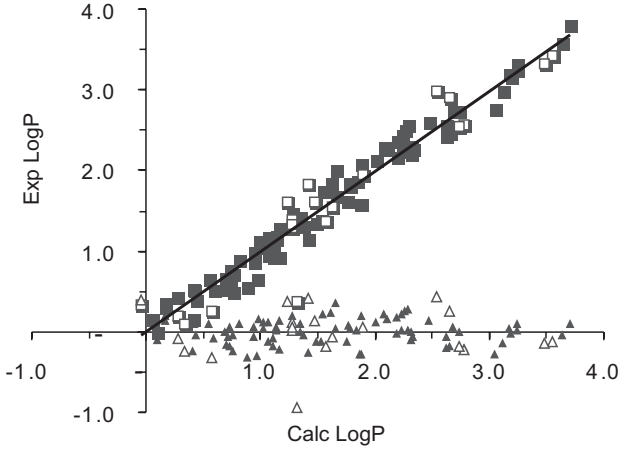

Figure 2. Plot of the experimental *vs.* calculated LogP for Liver, Olive oil, Muscle, and Fat (■), with twenty externally validated points (open squares), and the corresponding residual plot (▲, and open triangles).

CONCLUSIONS

This study confirms the usefulness of complete graphs, and especially of odd K_p, in deriving a useful valence delta number, δ^v, which includes information about the core electrons. Not always the model of a single case is optimal, like the saline case, but the central point here is that the suggested guideline, *similar properties of a common set of compounds should be described with a common descriptor* holds in a quite satisfactory way. In fact the found combination of four basis indices achieves a very interesting model of the six sets of partition coefficients, and especially of four of them. The unsatisfactory saline case seems to tell that the solvent here plays an incisive role in determining the outcome of the LogP values. The fact that for the six cases of LogP, and for the overall case, the K_p-(*p*-odd) representation for the core electrons is the only valid is rewarding as it strikes in the right direction: a homogeneous description for the core electrons for one and same set of compounds, and for the same type of property. That sequential complete graphs play no primary role in encoding the core electrons for these and other sets of compounds[1–5] rises the suspect of a preference for odd complete graphs by many sets of compounds, which till now has only been contradicted for a property of inorganic compounds.[20]

The importance of some basis indices should be underlined: $^1\chi^v$, as a single descriptor, $^0\psi_I$ in multiple linear combinations together with $^0\chi$, and with D_F, an D_{Cl}. The importance of $^1\chi^v$ as single basic descriptor and as an important index in the two- and four-index combination for the blood, fat and the muscle cases, is highly rewarding. This additive bond parameter, which encodes information on *p*, non-bonding, and core electrons, may be seen as the basic descriptor for LogP. Gass and Schwarzenbach have recently proposed some rules of thumb to understand qualitatively the trends observed with the partition coefficient.[16] These rules are mainly centered on concepts like apolar, monopolar, and dipolar, which describe the hydrogen-bond polarity. The basis subgraph indices, D_F, and D_{Cl}, are nothing else than the number of fluorine, and chlorine atoms and they, practically, describe an asymmetry towards the halogen atoms in the given optimal combination for the overall case. The detected nearly overall worsening of the model due to the *cis*-ethylene compound confirms a previous result that molecular connectivity indices work better with *trans*-compounds.[7]

REFERENCES

1. L. Pogliani, *Croat. Chem. Acta* **77** (2004) 193–201.
2. L. Pogliani, *New J. Chem* **27** (2003) 919–927.
3. L. Pogliani, *J. Comput. Chem.* **24** (2003) 1097–1109.
4. L. Pogliani, *J. Chem. Inf. Comput. Sci.* **44** (2004) 42–49.
5. L. B. Kier and L. H. Hall, *Molecular Connectivity in Structure-Activity Analysis*, Wiley, New York, 1986.
6. L. Kier and L. H. Hall, *Molecular Structure Description. The Electrotopological State,* Academic Press, New York, 1999.
7. L. Pogliani, *Chem. Rev.* **100** (2000) 3827–3858.
8. L. Pogliani, *J. Chem. Inf. Comput. Sci.* **41** (2001) 836–847.
9. L. Pogliani, *J. Phys. Chem.* **104** (2000) 9029–9045.
10. L. Pogliani, *J. Mol. Struct. (Theochem)* **581** (2002) 87–100.
11. M. L. Cargas, P. G. Seybold, and M. E. Andersen, *Toxicol Letts* **43** (1988) 235–256.
12. S. C. Basak, D. Mills, D. M. Hawkins, and H. A. El-Masri, *SAR&QSAR Environ. Res.* **13** (2002) 649–665.
13. S. C. Basak, B. D. Gute, D. Mills, and D. M. Hawkins, *J. Mol. Struct.(Theochem)* **622** (2002) 127–145.
14. P. G. Seybold, *Adv. Quant. Struct. Prop. Rel.* **3** (2002) 109–136.
15. B. Ren, *J. Comput-Aid. Mol. Des.* **17** (2003) 607–620.
16. K.-U. Gass and R. P. Schwarzenbach, *J. Chem. Ed.* **80** (2003) 450–455.
17. N. Trinajstić, *Chemical Graph Theory*, CRC Press, Boca Raton, Fl, 1992.
18. F. Harary and R. Read, *Proc.Graphs and Combinatorics Conference*, George Washington University, Springer, New York, 1977, cited in: J. D. Barrow, *The Book of Nothing*, Vintage Books, New York, 1977, p.155.
19. Z. Mihalić, S. Nikolić, and N. Trinajstić, *J. Chem. Inf. Comput. Sci.* **32** (1992) 28–37.
20. L. Pogliani, *Int. J. Quant. Chem.* **102** (2005) 38–52.

SAŽETAK

Modeliranje particijskih koeficijenata halogeniranih ugljikovih spojeva pomoću kemijske teorije grafova

Lionello Pogliani

Modelirano je šest vrsta particijskih koeficijenata za skup od 25 halogeniranih ugljikovih spojeva pomoću molekularnoga indeksa povezanosti, čiji su parametri dobiveni pomoću standardnih graf-teorijskih koncepcija: graf, pseudograf i potpuni graf. Pomoću ideje o neparnome potpunome grafu dobivena su četiri indeksa povezanosti i dva indeksa pseudopovezanosti s kojima je provedeno modeliranje particijskih koeficijenata. Pokazalo se da je navedeni skup indeksa optimalan za tri vrste particijski koeficijenata. Autor je također istaknuo da se najbolji pokazao valentni indeks povezanosti prvoga reda ($^1\chi^v$), a da se model znatno poboljša uklanjanjem *cis*-spoja iz skupa studiranih modela.