



## SOME THEORETICAL CONSIDERATIONS ON AN EMPIRICAL HYDROPHOBICITY INDEX IN REVERSED-PHASE LIQUID CHROMATOGRAPHY\*\*

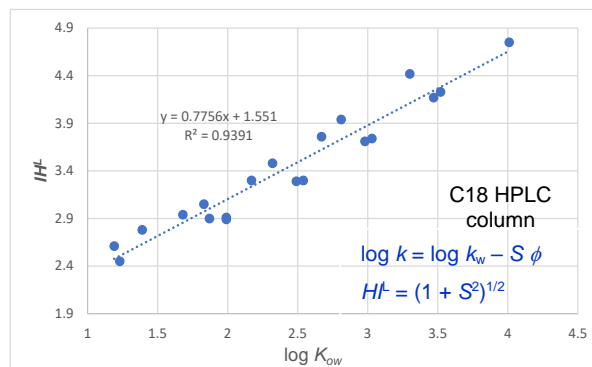
Victor DAVID,<sup>a\*</sup> Toma GALAON,<sup>a</sup> Edvin CAIALI<sup>a</sup> and Serban C. MOLDOVEANU<sup>b</sup>

<sup>a</sup>University of Bucharest, Faculty of Chemistry, Department of Analytical and Physical Chemistry,  
Sos. Panduri, no. 90, sector 5, Bucharest, 050663, Roumania

<sup>b</sup>SM Consulting, LLC, Winston-Salem NC, USA (smoldov@aol.com)

Received October 17, 2023

This study is analyzing the chromatographic hydrophobicity index (denoted by  $HI^V$ ) in reversed-phase liquid chromatography based on the linear dependence of the logarithm of experimental retention factor ( $\log k$ ) on the content of the organic component from mobile phase (retention function). In the literature  $HI^V$  is given by the ratio between the logarithm of the extrapolated retention factor for water as mobile phase ( $\log k_w$ ) and the slope ( $S$ ) of the linear equation describing the dependence of  $\log k$  on the organic solvent fraction ( $\phi$ ) from mobile phase composition. The paper demonstrates that the parameter  $HI^V$  is related to the extrapolated values of the retention factor for the extreme compositions of the mobile phase, namely to 100% aqueous component ( $\log k_w$ ) and 100% organic component ( $\log k_{org}$ ), respectively. A new empirical hydrophobicity descriptor (denoted by  $HI^L$ ) is proposed as an alternative to  $HI^V$ , which is related only to slope  $S$  as a square root of  $(1 + S^2)$ . Some examples and correlations are discussed and compared using retention data acquired for different HPLC columns and mobile phase compositions, showing that  $HI^L$  correlates better than  $HI^V$  with octanol-water partition constant ( $\log K_{ow}$ ) of studied solutes.



### INTRODUCTION

Mobile phase composition plays a main role in modulating the retention of solutes in liquid chromatography (LC).<sup>1</sup> The choice of the organic component in reversed-phase mechanism (RP) depends on solutes and the characteristics of stationary phase, and the most frequent choice is between methanol (MeOH) and acetonitrile (AcCN).<sup>2</sup> The retention factor ( $k$ ) calculated for

solutes with formula  $k = (t_R - t_0)/t_0$  (where:  $t_R$  being the retention time of solute and  $t_0$  the dead time of the RP column) depends on the characteristics of the stationary phase,<sup>3,4</sup> and the composition of mobile phase described by the volume fraction of the organic solvent content ( $\phi$ ).

The literature reports a large number of studies focused on the dependence between  $k$  and  $\phi$  in RP-LC,<sup>5</sup> and the retention modeling in this domain is applied to the estimation of various physico-

\* Corresponding author: victor.david@chimie.unibuc.ro

\*\*Supplementary Information on <https://www.icf.ro/rrch/> or <https://revroum.lew.ro>

chemical parameters that are otherwise difficult to be measured by other methods.<sup>6</sup> These parameters can be related to the nature of studies solutes, such as to the hydrophobicity/lipophilicity descriptors, solubility, dissociation constants, or thermodynamic,<sup>7</sup> but they can also be useful in stationary-phase characterization and comparison.<sup>8</sup> Experimental data from retention studies are used in different quantitative structure-retention relationships (QSRR) in order to find out molecular characteristics<sup>9</sup> or features of the stationary phases.<sup>10-12</sup> One of the molecular descriptors useful for characterizing both solutes and stationary phases has been introduced by Valko and coworkers, and this is designated for estimating the hydrophobic character of these entities. This descriptor is known as hydrophobicity index (*HI*),<sup>13,14</sup> and is usually utilized in different theoretical studies for understanding the retention mechanism in RP-LC.<sup>15-17</sup> It is the aim of this paper to enlarge and discuss in detail the meaning of this empirical concept of hydrophobicity index and to correlate it with the extrapolated values of the retention factor for extreme compositions of the mobile phase, namely to 100% aqueous component and 100% organic component.

## THEORETICAL FRAMEWORK

The dependence of the retention on the mobile phase composition (retention function) in many retention studies is described by a linear equation between the ten-base logarithm of the retention factor,  $\log k$ , and the fraction ( $\phi$ ) of the organic component in the mobile phase composition, written in this form:<sup>18</sup>

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

where:  $\phi$  is situated between [0; 1],  $k_w$  is the extrapolated value of the retention factor for 100% aqueous component ( $\phi = 0$ ), and  $S$  represents the slope of this dependence (solvent strength parameter),<sup>19</sup> *i.e.*  $d(\log k)/d\phi$ . In practice, this equation is valid only on a partial domain of  $\phi$ , but when the measurements are extended to the larger or entire domain of  $\phi$ , then the dependence between  $\log k$  and  $\phi$  can be described by a polynomial equation.<sup>5,20-24</sup> This situation is however dependent on the nature of the organic modifier from mobile phase.<sup>25</sup>

In the current applications a linear regression is applied to the dependence between  $\log k$  and  $\phi$ , written by a simple equation in the form:

$$\log k = a + b\phi \quad (1')$$

From the comparison between equations 1 and 1' the parameters of the retention function,  $\log k_w$  and  $S$ , can be calculated as:

$$\log k_w = a \quad \text{and} \quad S = -b \quad (1'')$$

Empirical observations have indicated that  $S$  can be estimated in a good approximation by means of the molecular weight,  $M_w$  (Da):<sup>26</sup>

$$S \approx \frac{1}{4} \sqrt{M_w} \quad (2)$$

With the aid of these two regression parameters,  $\log k_w$  and  $S$ , an empirical index describing the hydrophobicity of a solute in RP-LC has been advanced in the literature, which is specific for each studied solute. This empirical descriptor is known as RP-LC hydrophobicity index (*Valko et al.*)<sup>13,14</sup> and denoted by *HI*, which is given by the ratio:

$$HI = \frac{1}{S} \log k_w \quad (3)$$

or, with the aid of the regression parameters,  $a$  and  $b$ , the formula 3 becomes:

$$HI = -\frac{a}{b} \quad (3')$$

The formula for  $S$  can be deduced from eq. 1, by extrapolating  $\log k$  for  $\phi = 1$ , when  $\log k$  becomes  $\log k_{org}$ .<sup>27</sup> Consequently:

$$\log k_{org} = \log k_w - S \quad (4)$$

$$S = \log k_w - \log k_{org} = \log \frac{k_w}{k_{org}} \quad (5)$$

Formula (5) shows that the slope of the dependence (1) is given by the difference between  $\log k_w$  and  $\log k_{org}$ . Consequently, eq. 3 can be written as following:

$$HI = \frac{\log k_w}{(\log k_w - \log k_{org})} = \frac{\log k_w}{\log \frac{k_w}{k_{org}}} \quad (6)$$

or, written in this form:

$$HI = \frac{1}{1 - \frac{\log k_{org}}{\log k_w}} \quad (6')$$

According to the previous discussion, in case  $\log k_w > 0$  we may have the following situations: a)  $HI < 1$  (if  $\log k_{org}$  is negative);

b)  $HI = 1$  (if  $\log k_{org}$  is 0), and c)  $HI > 1$  (if  $\log k_{org}$  is positive).

In conclusion, the expression of hydrophobicity index proposed by Valko *et al.* depends on the extrapolated values of the logarithm of retention factor for the two extreme compositions of the mobile phase, namely  $\log k_w$  for 100% aqueous component ( $\phi = 0$ ), and  $\log k_{org}$  for 0% aqueous component ( $\phi = 1$ ).

Geometrically, the triangle describing the linear retention function given by eq. 1 and illustrated in Fig. 1 has the angles  $\alpha$  and  $\beta$ . The slope  $S$  is  $\text{tg } \alpha$ , which is equal to the difference previously discussed ( $\log k_w - \log k_{org}$ ), since the length of geometrical abscissa is 1. Also, the value of  $HI$  can be calculated from the ratio:

$$HI = \frac{A}{B} \quad (7)$$

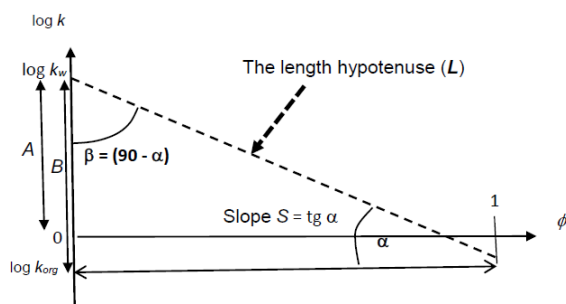


Fig. 1 – Graphical representation of the linear retention function in liquid chromatography ( $\log k$  versus  $\phi$ ).

Another index for hydrophobicity could be considered the length of hypotenuse of the retention triangle, denoted by  $L$ , whose formula can be deduced by the transformation:

$$\sin \alpha = \frac{\text{tg } \alpha}{\sqrt{1 + \text{tg}^2 \alpha}} = \frac{S}{\sqrt{1 + S^2}} \quad (8)$$

A relation obtained from Fig. 1 is following:

$$\sin \alpha = \frac{\log k_w - \log k_{org}}{L} = \frac{S}{L} \quad (9)$$

Thus, the expression of  $L$  from eq. 9 becomes:

$$L = \frac{\sqrt{1 + S^2}}{S} S = \sqrt{1 + S^2} = \sqrt{1 + \left(\log \frac{k_w}{k_{org}}\right)^2} \quad (10)$$

Unlike  $HI$  from formula 6 that can be smaller than 1, the values of  $L$  are always  $> 1$  as can be observed from eq. 10.

Both indexes can be considered as characterizing the hydrophobicity in RP-LC for studied solutes, as depending on the two mentioned extrapolated values of the retention factor. As a summary, using the linear model of the retention

dependence on the mobile phase composition, we can identify these two hydrophobicity indexes by following notation and with their new formulas, written again:

$$HI^V = \frac{\log k_w}{(\log k_w - \log k_{org})} \quad (11)$$

$$HI^L = \sqrt{1 + \left(\log \frac{k_w}{k_{org}}\right)^2} \quad (12)$$

These two descriptors could be used also in case the retention function is described by a polynomial function instead of the linear eq. 1. In QSRR studies both indexes can be correlated with molecular descriptors, mainly  $\log K_{ow}$ .

For a polynomial dependence, the expression of  $\log k$  can be written as:

$$\log k = a_1 + a_2 \phi + a_3 \phi^2 \quad (13)$$

From eq. 13, the formulas for extrapolated  $\log k_w$  and  $\log k_{org}$  become:

$$\log k_w = a_1 \quad (14)$$

$$\log k_{org} = a_1 + a_2 + a_3 \quad (15)$$

with the condition:  $a_2 + a_3 < 0$ . In this case, instead of using eq. 2', the index  $HI^V$  can be calculated with the regression parameters  $a_1$ ,  $a_2$ , and  $a_3$  by the ratio:

$$HI^V = -\frac{a_1}{a_2 + a_3} \quad (16)$$

For a polynomial dependence, the alternative index represented by eq. 12 is no longer valid, this time the slope being a variable on the entire retention domain.

## APPLICATION TO EXPERIMENTAL DATA SETS

Experimental data ( $t_R$  and  $t_0$ ) were acquired for two HPLC columns (Ultisil XB-C18, Welch, with dimensions 150 mm length, 4.6 mm inner diameter, and 5  $\mu\text{m}$  particle diameter; Gemini 5 u C18 110A, Phenomenex, 150 mm  $\times$  4.6 mm  $\times$  5  $\mu\text{m}$ ), by using mobile phase compositions with  $\phi$  between 0.45 – 0.60 for methanol, and 0.6 – 0.75 for acetonitrile as organic modifier. HPLC separations were carried out at 25°C using a Perkin Elmer LC system, model Flexar (Waltham, USA), with UV detection at 254 nm. Other experimental conditions have been already reported and used for the evaluation of methylene selectivity of studied solutes in RP-LC based on solvophobic theory.<sup>28</sup> The retention factor  $k$  was calculated for each studied solute with experimental values of  $t_R$  and  $t_0$

( $t_0$  being measured from the retention time of uracil). Then representation graphs between  $\log k$  and  $\phi$  were performed, and linear regressions were applied, from which the regression parameters  $a$  and  $b$  according to eq. 1' were calculated (see Tables 1S-4S from Supplementary Data).

The set of 20 studied solutes were following: benzene, toluene, ethylbenzene, propylbenzene, butylbenzene, benzoic acid and its esters from methyl to butyl, p-hydroxybenzoic acid and its esters from methyl to butyl (known as parabens), p-hydroxybenzaldehyde, p-hydroxyacetophenone, p-hydroxypropiofenone, p-hydroxybutyrophenone, and p-hydroxyvalerophenone (purchased from Sigma-Aldrich, Steinheim, Germany). Data on their molecular characteristics can also be found in Tables 1S-4S from Supplementary Data.

## RESULTS AND DISCUSSION

The two hydrophobicity indexes  $HI^V$  and  $HI^L$  were calculated with the presented formulas, starting from the linear dependences between  $\log k$  and  $\phi$ , obtained for the two different HPLC columns and mobile phases with one of the two frequently used organic solvents (methanol and acetonitrile). The data resulted from experimental values of the retention time as an average for three determinations for the set of 20 studied solutes were processed here in the view of calculating the main parameters involved in the described

procedure for the retention indexes. Thus, the extrapolated values  $\log k_w$  and  $\log k_{org}$  were calculated from the regression parameters of linear dependences between  $\log k$  and  $\phi$ . With the aid of these values, the values of  $HI^V$  and  $HI^L$  were computed (Tables 1S-4S). They were correlated to the theoretical values of octanol-water partition constant ( $\log K_{ow}$ , or  $\log P$ ), estimated by fragment methodology.<sup>29</sup>

Data processing for the results obtained for each HPLC column and each organic modifier of mobile phase composition has led to several useful trends, for the following correlations:

i)  $S$  versus  $\frac{1}{4}\sqrt{M_w}$  ;

ii)  $HI^V$  versus  $\log K_{ow}$ ;

iii)  $HI^L$  versus  $\log K_{ow}$ .

For ii) and iii) correlations there were preferred with the theoretical values of  $\log K_{ow}$ , due to the lack of experimental  $\log K_{ow}$  values for two of the studied compounds (*i.e.* p-hydroxy butyrophenone and p-hydroxy valerophenone). But the two descriptors, theoretical  $\log K_{ow}$  and experimental  $\log K_{ow}$ , are fairly well correlated, with the determination coefficient  $R^2 > 0.97$  (as can be seen from Fig. 2), such that this choice does not influence the discussion of the mentioned retention analysis of this study, and the two descriptors can be used interchangeable. Therefore, in the following correlations the theoretical  $\log K_{ow}$  has been used.

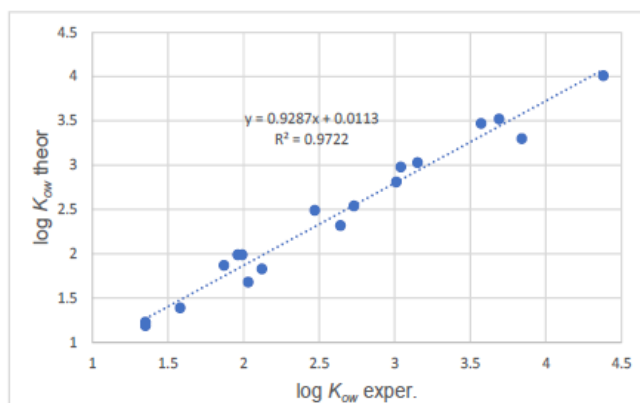


Fig. 2 – Correlation graph between theoretical  $\log K_{ow}$  and experimental  $\log K_{ow}$ .

First correlation between  $S$  and  $1/4\sqrt{M_w}$  was characterized by a rather poor determination coefficient ( $R^2$  approximately 0.83), the graph illustrated in Fig. 2 showing that this empirical rule is not always valid. This remark can be seen from

the different retention data obtained for the two solutes having the same  $M_w$ , namely methyl benzoate and 4-hydroxy acetophenone. The high difference between their regression parameters, and consequently with high difference between their

calculated parameters, induces low correlation between  $S$  and  $1/4\sqrt{M_w}$ . Also, according to the graph given in Fig. 2, the slope of the linear dependence between  $S$  and  $1/4\sqrt{M_w}$  (calculated as 0.3002) was much lower than the expected value of

this correlation (theoretically 1.00, according to formula 2). Therefore, this rule according to which the slope  $S$  of the retention function (eq. 1) is correlated to the square root of the molecular weight ( $M_w$ ) has to be used with caution.

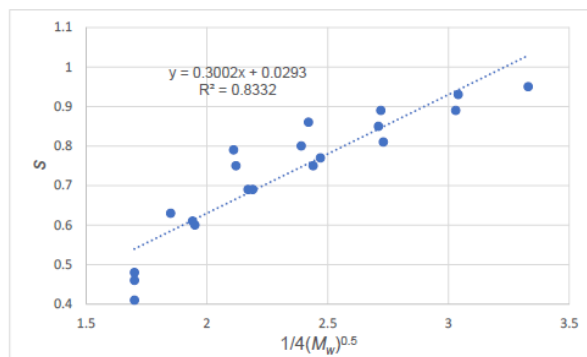


Fig. 3 – Correlation graph between  $S$  and  $\frac{1}{4}\sqrt{M_w}$ .

The main used hydrophobicity index  $HI^V$ , known also as *Valko* index, was studied for the two HPLC columns and the two organic modifiers. Its correlation with  $\log K_{ow}$  is characterized by  $R^2$  between 0.67 and 0.76, with the remark that better  $R^2$  is obtained for methanol

as organic modifier (Figs. 4 and 5). There are no significant differences between the two used columns, in both cases for organic modifier, and a conclusion of this study is this hydrophobicity index is rather poorly correlated to  $\log K_{ow}$  of the studied solutes.

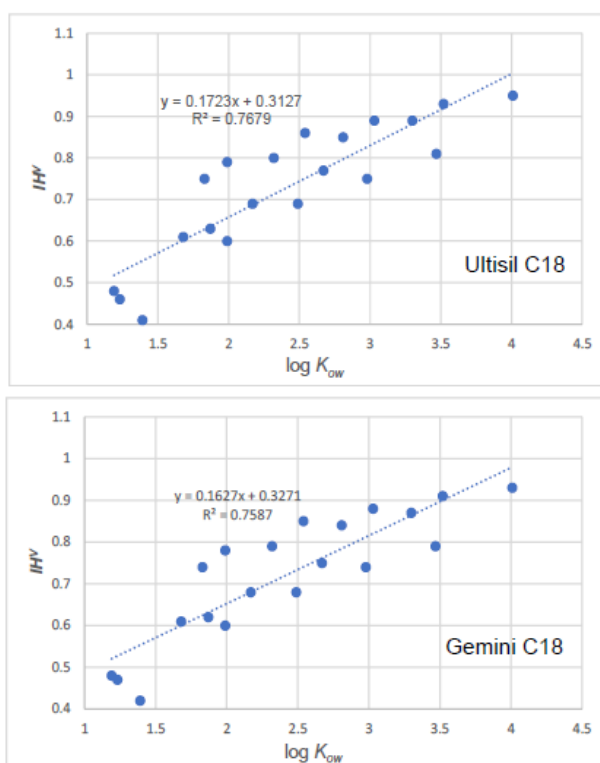


Fig. 4 – Comparison of the correlations between  $HI^V$  and  $\log K_{ow}$  for the two indicated columns and MeOH used as organic solvent in mobile phase.

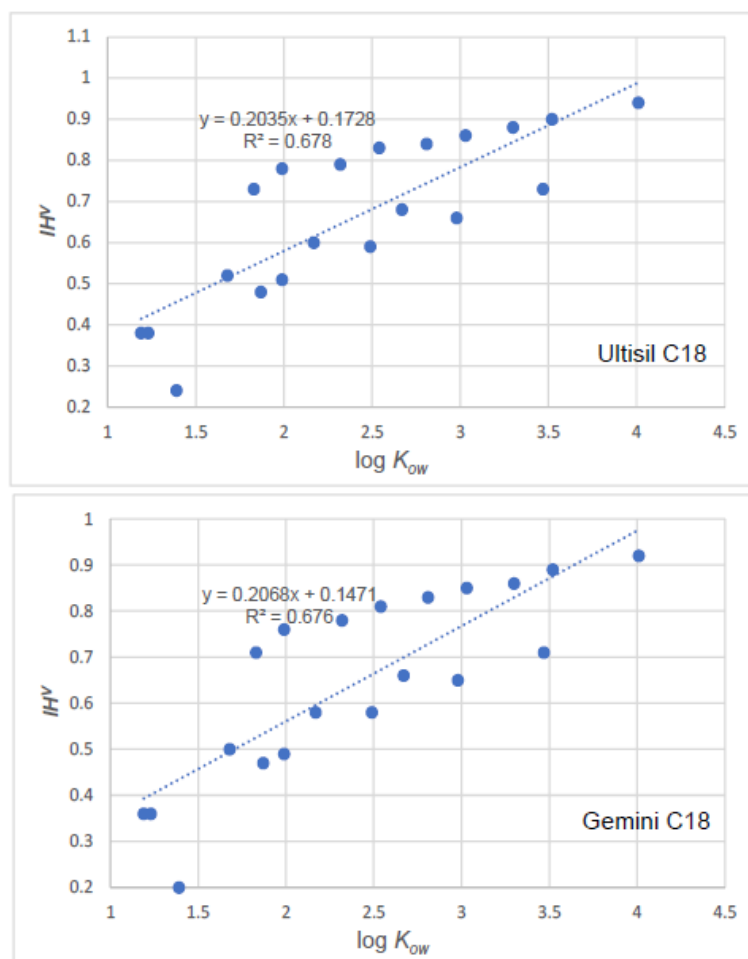


Fig. 5 – Comparison of the correlations between  $HI^V$  and  $\log K_{ow}$  for the two indicated columns and AcCN used as organic solvent in mobile phase.

The correlations are significantly improved when studying the dependence between  $HI^L$  and  $\log K_{ow}$ . As can be seen from Figs. 6 and 7, the determination coefficients for these correlations are

higher than 0.9, with the remark that  $R^2$  has close values for both organic solvents. The corresponding values for  $HI^V$  and  $HI^L$  are summarized in Table 1.

Table 1

Comparison of  $R^2$  for correlations between  $HI^V$  and  $HI^L$  indexes with  $\log K_{ow}$

<b>RP-HPLC column and mobile phase</b>	<b><math>HI^V</math></b>	<b><math>HI^L</math></b>
Ultisil C18, MeOH	0.7679	0.9391
Ultisil C18, AcCN	0.6780	0.9182
Gemini C18, MeOH	0.7587	0.9207
Gemini C18, AcCN	0.6760	0.9320

Taking into consideration that  $HI^L$  is a function of only slope  $S$ , we may conclude that this one may represent another option for hydrophobicity index

for retention modeling in RP-LC. So far, this descriptor has not been studied in the literature, and any comparison with other attempts is not possible.

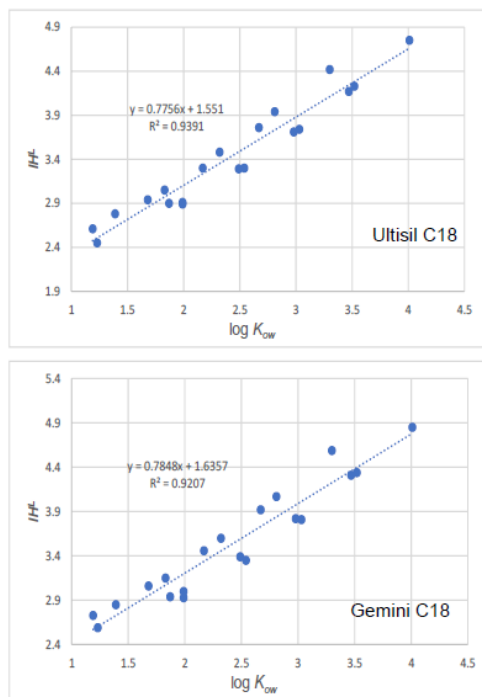


Fig. 6 – Comparison of the correlations between  $HI^L$  and  $\log K_{ow}$  for the two indicated columns and MeOH used as organic solvent in mobile phase.

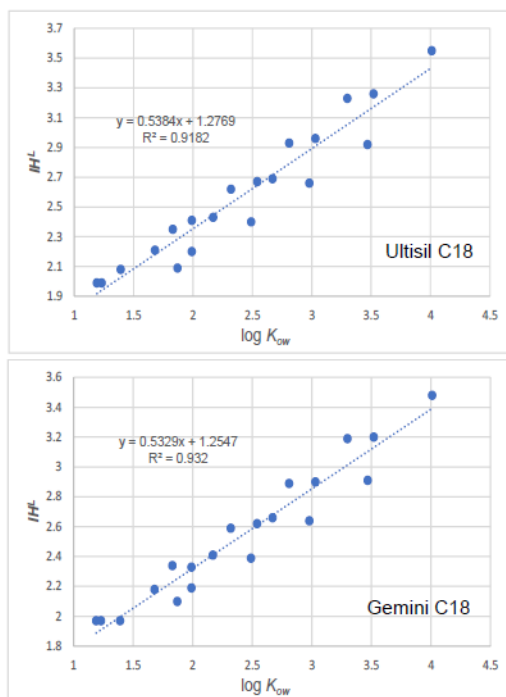


Fig. 7 – Comparison of the correlations between  $HI^L$  and  $\log K_{ow}$  for the two indicated columns and AcCN used as organic solvent in mobile phase.

A general remark of the studied correlations between  $HI^V$  and  $HI^L$  with  $\log K_{ow}$ , given in Figs. 3–7, refers to the slope of these linear regressions. For the same organic modifier, the slopes of dependences  $HI^V$  or  $HI^L$  on  $\log K_{ow}$  are almost the same for the two used HPLC columns (Ultisil C18 and Gemini C18).

The most closed values can be seen in Fig. 7 for the slope of the linear dependences between  $HI^L$  and  $\log K_{ow}$  for the two columns, with AcCN used as organic solvent in mobile phase, 0.5384 versus 0.5392, respectively. However, they differ significantly from methanol to acetonitrile in all four correlations, the

highest differences being observed in case of  $HI^L$ . In conclusion, the slope of dependences between  $HI^V$  or  $HI^L$  on  $\log K_{ow}$  does not depend on the HPLC column for the same organic modifier (MeOH or AcCN).

## CONCLUSIONS

From the analysis of hydrophobicity index ( $HI^V$ ) advanced several years ago by Valko *et al.*, a new possibility resulted for describing the retention process in RP-LC. This paper demonstrates that the parameter  $HI^V$  is related to the extrapolated values of the retention factor for the extreme compositions of the mobile phase, namely to 100% aqueous component ( $\log k_w$ ) and 100% organic component ( $\log k_{org}$ ), respectively. A new hydrophobicity index is proposed that includes in its formula only one parameter of the retention dependence between  $\log k$  and  $\phi$ , namely the slope  $S$  of this linear equation; this index is denoted by  $HI^L$ . Applied to a set of 20 studied solutes this new hydrophobicity index correlated much better with the octanol-water partition constant ( $\log K_{ow}$ ) of solutes than the hydrophobicity index introduced by Valko *et al.* This trend was observed for two HPLC columns and for both methanol and acetonitrile used as organic solvents in mobile phase. Another conclusion of this study refers to the poor correlation between the slope  $S$  of the linear dependence between  $\log k$  and  $\phi$  with molecular weight of the studied solutes, an empirical rule that has been reported by the literature.

## REFERENCES

1. S. Moldoveanu and V. David, "Essentials in Modern HPLC Separations", 2<sup>nd</sup> edition, Elsevier, Amsterdam, 2022, pp. 207–269.
2. V. Ganesh, P. Poorna Basuri, K. Sahini and C. N. Nalini, *Biomed. Chromatogr.*, **2023**, *37*, e5482.
3. N. Tanaka, K. Sakagami and M. Araki, *J. Chromatogr. A*, **1980**, *199*, 327–337.
4. L. C. Sander and S. A. Wise, *Anal. Chem.*, **1995**, *67*, 3284–3292.
5. M. J. den Uijl, P. J. Schoenmakers, B. W. Pirok and M. R. van Bommel, *J. Sep. Sci.*, **2021**, *44*, 88–114.
6. J. G. Dorsey and W. T. Cooper, *Anal. Chem.*, **1994**, *66*, 857A–867A.
7. C. F. Poole and S. N. Atapattu, *J. Chromatogr. A*, **2020**, *1626*, 461427.
8. L. C. Sander and S. A. Wise, *Anal. Chem.*, **1995**, *67*, 3284–3292.
9. C. F. Poole and S. N. Atapattu, *J. Chromatogr. A*, **2020**, *1626*, 461427.
10. R. Kalisz, *Chem. Rev.*, **2007**, *107*, 3212–3246.
11. R. Kalisz, M. A. van Straten, M. Markuszewski, C. A. Cramers and H. A. Claessens, *J. Chromatogr. A*, **1999**, *855*, 455–486.
12. T. Baczek, R. Kalisz, K. Novotna and P. Jandera, *J. Chromatogr. A*, **2005**, *1075*, 109–115.
13. K. Valkó and P. Slégel, *J. Chromatogr. A*, **1993**, *631*, 49–61.
14. K. Valkó, C. Bevan and D. Reynolds, *Anal. Chem.*, **1997**, *69*, 2022–2029.
15. A. Carotti, I. Varfaj, I. Pruscini, G. W. A. Abualzulof, L. Mercolini, E. Bianconi, A. Macchiarulo, E. Camaioni and R. Sardella, *J. Sep. Sci.*, **2023**, *46*, 2300346.
16. S. Espinosa, E. Bosch, M. Rosés and K. Valkó, *J. Chromatogr. A*, **2002**, *954*, 77–87.
17. K. Valkó, L. R. Snyder and J. L. Glajch, *J. Chromatogr. A*, **1993**, *656*, 501–520.
18. E. Soczewinski and C. A. Wachtmeister, *J. Chromatogr.*, **1962**, *7*, 311–320.
19. S. N. Atapattu, *J. Sep. Sci.*, **2023**, *46*, 2300489.
20. P. J. Schoenmakers, H. A. H. Billiet and L. De Galan, *J. Chromatogr.*, **1983**, *282*, 107–121.
21. A. R. Horner, R. E. Wilson, S. R. Groskreutz, B. E. Murray and S. G. Weber, *J. Chromatogr. A*, **2019**, *1589*, 73–82.
22. P. Nikitas, A. Pappa-Louisi and P. Agrafiotou, *J. Chromatogr. A*, **2002**, *946*, 9–32.
23. S. C. Moldoveanu, E. Bacalum, T. Galaon and V. David, *J. Sep. Sci.*, **2023**, *46*, 2300274.
24. E. Bacalum, T. Galaon, A.-C. Soare, V. David and S. C. Moldoveanu, *Rev. Roum. Chim.*, **2023**, *68*, 277–283.
25. T. Galaon, A. Medvedovici and V. David, *Sep. Sci. Technol.*, **2008**, *43*, 147–163.
26. J. W. Dolan, *LCGC North America*, **2014**, *32*, 546–551.
27. V. Voicu, I. Sora, C. Sarbu, V. David and A. Medvedovici, *J. Pharm. Biomed. Anal.*, **2010**, *52*, 508–516.
28. S. C. Moldoveanu, E. Caiali and V. David, *J. Liq. Chromatogr. Rel. Technol.*, **2018**, *41*, 24–32.
29. <http://www.epa.gov/oppt/exposure/pubs/episuite.htm>