

## HILIC Chromatography – An Insight on the Retention Mechanism

Aikaterini I Pitani, Maria G Kouskoura\* and Catherine K Markopoulou

Laboratory of Pharmaceutical Analysis, Department of Pharmaceutical Technology, School of Pharmacy, Faculty of Health Sciences, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

### Abstract

Hydrophilic interaction chromatography (HILIC) could be characterized as a complex chromatographic system that involves multiple mechanisms. These are partitioning as well as polar and ionic interactions. Among several HILIC columns, ZIC-HILIC can be used to separate small organic ionic compounds. The presence of both positive and negative charge on the stationary phase may facilitate separations of both anionic and cationic analytes. Based on the Partial Least Squares methodology, an attempt to clarify the mechanism on this column revealed that the forces dominating are mainly determined by structural features. Consequently, the physicochemical properties which are related to the analytes' structure may heighten or attenuate the process. Ionic interactions are stronger for analytes containing moieties with basic properties since the interaction with the sulfonyl group is facilitated. The partition mechanism is prevailing for those analytes that are not sufficiently ionized at the experimental conditions (mobile phase pH 3 and 6.5) and for analytes that can create halogen bonds. Moreover, the stagnant water layer on the silica bed enhances the retention of water soluble compounds due to the increased hydrophilic interactions.

**Keywords:** ZIC-HILIC; Retention mechanism; PLS; Hydrophilic interaction

### Introduction

Hydrophilic Interaction Chromatography (HILIC) is a term that was first suggested in 1990 by Alpert [1] who used the column for the separation of peptides, amino acids and other polar compounds. This methodology provides an alternative approach to separate small polar molecules efficiently using a stationary phase with increased polarity. Hydrophilic chromatography employs traditional polar stationary phases which are used in normal phase chromatography (NP-LC) and a mobile phase that is more suitable for reversed phase liquid chromatography (RP-LC) [2]. Moreover, it is suitable for highly hydrophilic and amphiphilic analytes that are too polar to be retained in RP-LC; still their charge is sufficient to allow the conditions of ion-exchange chromatography (IC) [3].

An important advantage of HILIC is that it can overcome the drawback of poor water solubility of many analytes that is often a problem in NP-LC. The most common eluent that is used in HILIC technique is acetonitrile, which is relatively polar. This has been the merit of HILIC towards NP-LC which uses mobile phases not friendly to the environment; still it seems to be a disadvantage in terms of green chemistry [4,5].

In terms of chromatography, it is known that analytes which are strongly hydrophilic but cannot attain charge in solution, do not seem to be retained on either type of stationary phases. On the contrary, in the case of highly hydrophilic compounds with functional groups allowing dipolar bonding, HILIC may retain molecules since it has been supplemented by the use of ion exchange chromatography [6].

Conceptually, HILIC could be considered as reversed phase chromatography, where a polar stationary phase has the ability to retain polar analytes which are eluted by mobile phase consisting of a mixture of an organic modifier and water.

Many researchers Buszewski [3], Bernal [7] contend that the mechanism of retention that prevails in HILIC chromatography is a difficult and complicated task that has not yet been clarified. In this vein, a large number of characteristics of numerous analytes have been employed so as to study their effect on the elution on a zwitterionic column (ZIC-HILIC) and explain the chromatographic behavior.

Based on the assertion that HILIC is partly ion exchange chromatography, stationary phases with zwitterionic functionalities (such as ZIC-HILIC) were intended for ion exchange separations [8,9] Still, the retention's mechanism complexity includes partition together with adsorption and hydrophobic interactions under specific conditions [10-12].

According to the theory of chromatography, it has been established that pH of the mobile phase is a tool that may control selectivity, since it determines the analyte's form. However, in reversed phase chromatography, molecules are not completely ionized in the experimental conditions and the presence of both ionized and neutral species leads in peak tailing or fronting [13]. For this instance, it is preferred to use a mobile phase where the aqueous phase pH is adjusted so as to induce ion formation. Based on the Henderson-Hasselbach equation [14]. The pH of the mobile phase will affect ionization of an analyte in accordance to the  $pK_a$  of the compound at a given pH value. This drawback has been surpassed using HILIC chromatography where the peak shape is improved [15].

The retention in HILIC is mainly affected by adjusting the eluent (e.g., the fraction of organic modifier), the type and concentration of the buffer and the pH value [7]. But in the case that there is high diversity in the structure of the analytes, the retention may depend on several additional factors.

The aim of this work is to investigate the mechanism of retention on a ZIC-HILIC column. Such an attempt was made through the study of a large number of analytes that belong in different groups of compounds (mainly drugs) at different pH conditions and different organic

\*Corresponding author: Maria G Kouskoura, Laboratory of Pharmaceutical Analysis, Department of Pharmaceutical Technology, School of Pharmacy, Faculty of Health Sciences, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece, Tel: +302310997665; E-mail: [mkousk@gmail.com](mailto:mkousk@gmail.com)

Received June 15, 2016; Accepted June 21, 2016; Published June 27, 2016

Citation: Pitani AI, Kouskoura MG, Markopoulou CK (2016) HILIC Chromatography – An Insight on the Retention Mechanism. J Chromatogr Sep Tech 7: 326. doi:10.4172/2157-7064.1000326

Copyright: © 2016 Pitani AI, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

modifier. The choice of ZIC-HILIC among several HILIC columns was made because it has been proved superior and supplementary to both NP-LC and RP-LC. Additionally this column can be conveniently coupled with many types of detectors such as UV and MS (especially in the ESI mode) [3].

Previous researches revealed that exploring the mechanism of retention of analytes on a stationary phase is feasible by the development of a model that takes into consideration as many parameters as possible [16,17]. There had been several studies using Quantitative Structure Retention Relationships (QSRR) [18] Principal Component Analysis (PCA) [19,20] or Multiple Linear Regression Analysis (MLR) [21] and Partial Least Squares (PLS) [22] on several columns.

According to the literature there had been studies aiming to explain the chromatographic behavior of specific categories of compounds using a limited number of descriptors [23,24]. However, there has not yet been an attempt to clarify the retention mechanism of analytes on a HILIC column using such techniques for a miscellaneous gathering of analytes using chemometric methodologies.

Therefore, this work focuses on explaining and modeling the retention behaviour with the aim of PLS. The advantage of PLS technique is that it can correlate significantly larger datasets compared to other methods [16]. The descriptors used in the current study include several physicochemical properties of the analytes along with their structural characteristics. Thus, within the results of the present study lies some perspective in the enlightenment of the complex retention mechanism on a ZIC-HILIC column.

## Materials and Methods

### Chemical and reagents

The analytes used in the experiments employ a miscellaneous group consisting of 136 compounds (Supplementary Information; Tables A and B) and analytical grade reagents. The compounds were dissolved in methanol at concentrations ranging from 5 µg/mL to 10 µg/mL depending on the solubility of each compound. 25 µL of the final dilutions were injected with a SIL-20AD auto-sampler.

The analytes were USP-grade and were obtained by i. Carlo Erbra (Milano, Italy), ii. Panreac Company (Controla, Thessaloniki, Greece), iii. Sigma-Aldrich (Life Science Chemilab), iv. Riedel-de-Haën (Life Science Chemilab), v. Fluka (Life Science Chemilab, authorized distributor in Greece), vi. Applichem (Bioline Scientific, Athens, Greece), vii. Acros-Organics (Life Science Chemilab), viii. Pierce Chemicals (Pierce Chemical Company, Rockford, Illinois), and ix. Boehringer Ingelheim (GmbH, Ingelheim am Rhein) companies.

Acetonitrile (ACN), methanol (MeOH) and water (H<sub>2</sub>O) were gradient grade for liquid chromatography and were obtained from Panreac. The adjustment of the pH value of the mobile phase was made by the addition of aqueous solution of NH<sub>4</sub>OH>25% or HCl 37%, obtained by Fluka. Mobile phases were degassed by filtering through a Millipore HV 0.45 µm pore membrane filter.

### Instrumentation

The experimental procedure was carried out on an HPLC (Shimadzu) instrument, equipped with two LC-20AD pumps, a SIL-10AD auto-sampler and a UV-diode array detector. LC Solution software was used to record and elaborate the chromatographic peaks. In order to ensure stable experimental conditions, the temperature was maintained at 30°C using a column oven (Shimadzu).

The stationary phase was a ZIC-HILIC column obtained from Merck Sequant, with dimensions 150 × 4.6 mm and particle size 5 µm.

Chromatographic characterization of the probes was studied on the basis of the pH difference of the mobile phase, as well as on the difference of the mobile phase organic modifier.

The solvent mixtures used as mobile phase were 40% MeCN/H<sub>2</sub>O adjusted at two pH values (3 and 6.5) using aqueous solutions of HCl and NH<sub>3</sub> respectively. The flow rate of the mobile phase was adjusted to 0.2 mL/min and the retention time was determined by the mean of at least duplicate measurements. The retention time (t<sub>r</sub>) of naphthol was checked at frequent time intervals in order to control the stability of the chromatographic system.

### Partial Least Squares (PLS)

The software used to develop and evaluate the models, is an independent modeling of class analogies Simca-P (Version 9; Umetrics, Uppsala, Sweden). For this purpose, three datasets were separately designed and three PLS models were compiled, labelled as total. The models designed represent data at three different experimental conditions: (a) mobile phase 40% MeCN /H<sub>2</sub>O at pH 3, (b) mobile phase 40% MeOH/H<sub>2</sub>O at pH 3 and (c) mobile phase 70% MeCN /H<sub>2</sub>O at pH 6.5. Each dataset contained 79X variables and experimental data (Y variable) for 136 observations (analytes).

Despite the fact that some of the descriptors studied were proved to be of minor interest, the range of physicochemical properties recorded was wide. The study aimed to eliminate the most important factors affecting the retention among numerous physicochemical properties as well as structural features. The major division of the dataset is shown in Figure 1.

The structural characteristics are encountered within the constitutional descriptors. They were inserted in the dataset using integer numbers and zero, which were used to denote the presence, the multiplicity or the absence of a structural feature. The chosen observations represent a diverse gathering of compounds so as to take into account as many structural characteristics as possible. This way several conclusions on the retention mechanism could be drawn concerning either the total dataset or specific classes containing groups of similar compounds.

The structure of chemical compounds is the major factor determining their physicochemical properties. Thus, the descriptors employed for the implementation of the models include lipophilicity parameters, as well as topological, geometrical and electronic properties (Supplementary Information; Table C). The above data were derived by different databases such as ACD/Labs [24] and calculation platforms

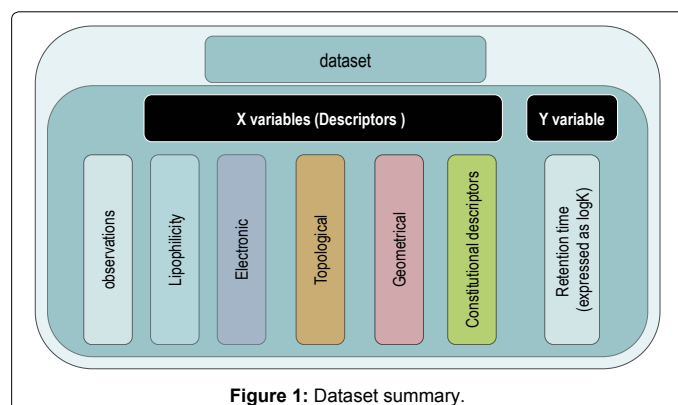


Figure 1: Dataset summary.

Dragon [25] Osiris [26], Hyperchem [27] Marvin [28]. All of them included drawing and displaying structures, as well as their geometry optimization which was made based on the Polak-Ribiere algorithm [29].

The distribution coefficient (logD) is directly related to the ionization of the compounds and had to be thoroughly studied. LogD indicates the distribution of the neutral or ionic form of a molecule in water against the non-ionized species in octanol. Based on this fact, it was necessary to calculate the corresponding logD values of the analytes studied and insert them as one of the variables in the dataset. Apart from the logD values, the ionization percentages of each analyte were also calculated and recorded (Supplementary Information; Table B).

Aside from the X variables which represent the aforesaid characteristics of the analytes a Y variable was also inserted. It refers to the logK which corresponds to the retention of each analyte on the stationary phase. The Y variable is the logK of the analytes, where K is the capacity factor which represents the relative speed of the analyte through the column and is expressed by the fraction  $K = \frac{t_r - t_o}{t_o}$  ( $t_r$  is the retention time of the analytes and  $t_o$  is the solvent front).

The observations studied were divided in two large groups (Figure 2). The first consists of chemical compounds (including drugs) which cannot ionize at the experimental conditions or their ionization percentage (IP%) is lower than 45%.

This percentage has been shown to be the turning point for ionization; when ionization is lower, the effect on the retention is negligible [30]. The second group includes molecules that are affected by the pH of the mobile phase and the ionization degree (of their acidic or basic group) is higher than 45%. These groups were also studied as separate models.

The  $pK_a$  values of the basic or acidic groups of the analytes are determining their ionization, which will vary based on the pH of the mobile phase. Therefore, an overall and more detailed view for the retention behavior of the analytes on HILIC column, as well as a clarification of the effect of their ionization, was achieved through the overall and individual models studied.

**Method validation:** Cross-validation (CV) is a method leaving out part of the observations each time, and it was used to select the number of components (model dimensions) in the PLS algorithm. Based on the CV, the response values (Y) for the observations kept out were predicted by the model and then compared with the true values. The number of rounds used by Simca-P9 for model fitting was 7. Using the appropriate number of significant components, the total models were fit according to Haaland and Thomas [31] criteria. The results of the statistical data along with the validation results are given in Table 1 (and in Supplementary Information; Figure A-C) To affirm full statistical significance of the original estimates, the intercept limits were set to be  $r^2 < 0.3$  and  $q^2 < 0.05$  [32].

Table 1 shows the statistical coefficients for all models compiled. The first model for each mobile phase is characterized as total since it includes all observations. The other two models are characterized as group A for observations whose ionization degree is  $0\% \leq (IP\%) \leq 45\%$  and group B for observations with  $IP\% > 45\%$  (as described in Figure 2).

According to the statistical data, it was obvious that the model developed when the mobile phase contained acetonitrile at pH 3 was the most reliable since it was statistically significant ( $R^2=0.914$ ,  $Q^2=0.792$ ). On the other hand, when the organic modifier of the mobile phase was replaced by methanol, although the correlation coefficient was good, the models appeared to lack predictive ability. It is worth mentioning

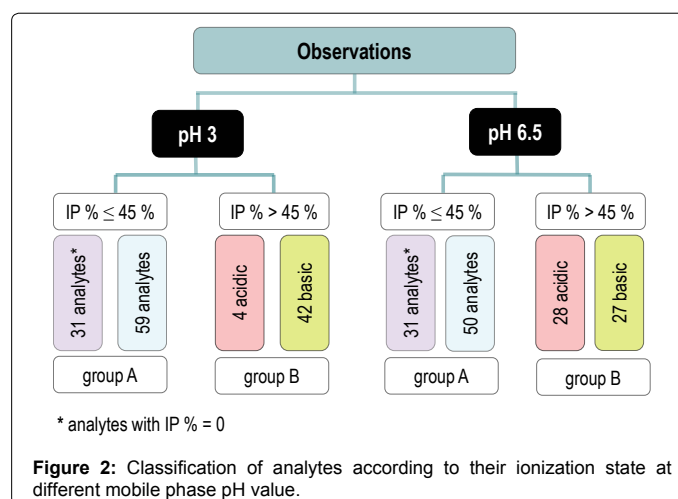


Figure 2: Classification of analytes according to their ionization state at different mobile phase pH value.

that the models at pH 6.5 also show poor predictive ability.

These results are justified due to the fact that the gathering of analytes was miscellaneous and they have many differences in structure and physicochemical properties. At pH 3 the behaviour has a more consistent pattern because the ionized analytes are those containing basic groups. On the other hand, at higher pH (6.5) there are compounds whose ionization occurs in both basic and acidic moieties in their structure and the interaction with the stationary phase is more complex. Finally, in the case of the model where the mobile phase contains MeOH at pH 3, the predictive ability is low because the mobile phase may also interact with the stationary phase shielding the charged functional groups. Therefore, methanol competes with the analytes for the active sites of the column and may interfere in the retention mechanism.

## Results and Discussion

The commercially available HILIC columns are divided in three categories; these are neutral, charged (anionic, cationic) and zwitterionic [7]. The choice of the ZIC-HILIC stationary phase was made as it is the most widely used HILIC column, because it is applicable in many categories of analytes. The structure of this zwitterionic column (Figure 3) includes a quaternary ammonium and a sulfonic group attached on the silica bed.

The chromatographs obtained showed that the majority of the chromatographic peaks of the analytes were symmetrical with no distortion at all experimental conditions.

Initially, this research focused on the correlation of the retention time of an analyte with its ionization degree. In short, the pH (3 and 6.5) of the mobile phase determines the ionization degree of a molecule, its logD value and consequently its behavior ( $t_r$ ) on a stationary phase.

A second part of the study has been an investigation of the factors affecting the retention on the HILIC column for all the observations regardless their ionization state. The effect of each variable (xk) on the Y variable in the model is labelled as VIP (Variable Importance in the Projection). VIP is the sum over all model dimensions of the contributions VIN (variable influence).

Moreover, the total PLS models were explored via the loadings' plots ( $w^*c$  Ref. [1]) which display the correlation between the X variables and the Y scores in the first dimension (component) [32,33]. The loadings' plots provide the additional information of whether the

effect of each variable is positive or negative on the logK.

### Exploring the retention mechanism of the total models

The pH value of the mobile phase has been adjusted at pH 3 by the addition of HCl (37%). This choice has been made in order to avoid any other interferences of a buffer solution on the retention which would be caused by secondary mechanisms. The retention of the analytes at this pH value can be studied based on the different classes at which the analytes belong. First of all, the compounds that cannot be ionized are not retained on the stationary phase. Most of them are eluted along with the solvent front (~ 8-9 min). Moreover, there is a number of analytes that although bearing functional groups which are susceptible to ionization, the degree of ionization at the pH value of the mobile phase is not sufficient (lower than 45%). Consequently, the retention of the majority of them is also not affected and their elution is fast when the mobile phase consists of acetonitrile in water.

When the organic modifier of the mobile phase was replaced by methanol the majority of the analytes seem to be affected by this alteration and the retention is delayed. In the case of compounds that are not ionized, the retention time is slightly longer with no exceptions. Still in both cases, this delay in the retention time is short and its variance is limited.

For analytes whose ionization is lower than 45%, the retention is also slightly delayed; but there are a few compounds that have an odd behaviour. Some examples of these compounds are shown in Table 2. According to the information provided, it is obvious that even in the case of analytes that are not ionized (most of the compounds of Table 2 have IP << 45%; Supplementary Information; Table B), the retention is slightly higher when the mobile phase contains methanol. This can be explained as acetonitrile is an aprotic polar solvent which does not participate in the ionization of the analytes. Moreover, methanol can act as both hydrogen bond donor and acceptor and can compete for active polar sites on the HILIC surface. Therefore, if an analyte is more hydrophilic (even slightly) due to low ionization percentage it might be more retained on the stationary phase compared to its corresponding

retention time when the mobile phase consists of acetonitrile.

The study of the elution strength of the solvents used must be based on the ionization percentage of the analytes due to the differences occurring in the mechanism retention (Table 2). Therefore, in the case of molecules from the group A (Figure 2, IP<45%, Supplementary Information; Table B), the elution of the analytes is fast (sometimes along with the solvent front) while there are no significant difference when the organic modifier is not the same. The general pattern is that the retention time for all analytes is slightly increased when the mobile phase contains methanol against acetonitrile. Still, there are specific compounds that are retained longer with acetonitrile due to special structural characteristics or physicochemical properties. An example is clofoctol whose structure results in very high values at descriptors assigned to specific structural features such as C atoms on side chains. The aforesaid results in high value for free rotatable bonds (FRB) and the Balaban index (J3D) which is a topological index increasing when the number of branches on a molecule increases [34]. Probably, the complex structure, including many branches, makes the molecules bulky and in a way it is trapped within the structure of the stationary phase, even if the ionic interactions are negligible.

A similar behaviour is observed for analytes that are highly ionized (even in their acidic moieties). Obviously, there is a steric hindrance of quaternary amine group of the column and cannot develop ionic interactions with analytes that are negatively charged.

On the other hand, highly ionized basic groups exhibit a clear delay in their retention due to the ionic interactions with the sulfonic group. Actually, the eluting strength of the solvents is different for these analytes; water enhances the elution strength while acetonitrile causes a delay in the retention.

Apart from the conditions mentioned in the present study, the retention time of the analytes was also experimentally measured with a mobile phase containing 70% MeCN in water at pH 3. The model compiled for these data gave the same results in terms of retention behaviour and the effect of the descriptors studied. The only difference

|                | 40% MeCN at pH 3 |            |             | 40% MeOH at pH 3 |              |             | 40% MeCN at pH 6.5 |             |              |
|----------------|------------------|------------|-------------|------------------|--------------|-------------|--------------------|-------------|--------------|
|                | Total            | Group A    | Group B     | Total            | Group A      | Group B     | Total              | Group A     | Group B      |
| Observations   | 136              | 90         | 46          | 136              | 90           | 46          | 136                | 81          | 55           |
| Components     | 6                | 3          | 3           | 4                | 4            | 4           | 5                  | 3           | 3            |
| R <sup>2</sup> | 0.914            | 0.875      | 0.860       | 0.753            | 0.8          | 0.926       | 0.805              | 0.669       | 0.801        |
| Q <sup>2</sup> | 0.792            | 0.818      | 0.662       | 0.577            | 0.632        | 0.493       | 0.584              | 0.518       | 0.599        |
| Validation     |                  |            |             |                  |              |             |                    |             |              |
| r <sup>2</sup> | 0.0, 0.22        | 0.0, 0.174 | 0.0, 0.314  | 0.0, 0.194       | 0.0, 0.0548  | 0.0, 0.306  | 0.0, 0.196         | 0.0, 0.19   | 0.0, 0.124   |
| q <sup>2</sup> | 0.0, -0.435      | 0.0, -0.29 | 0.0, -0.148 | 0.0, -0.174      | 0.0, -0.0746 | 0.0, -0.161 | 0.0, -0.388        | 0.0, -0.113 | 0.0, -0.0846 |

Shows the statistical coefficients for all models compiled. The first model for each mobile phase is characterized as total since it includes all observations. The other two models are characterized as group A for observations whose ionization degree is  $0\% \leq (IP\%) \leq 45\%$  and group B for observations with  $IP\% > 45\%$ .

Table 1: Statistical data for the compiled models.

| Analyte           | Retention Time |          |          | Analyte       | Retention Time |          |          |
|-------------------|----------------|----------|----------|---------------|----------------|----------|----------|
|                   | 40% MeOH       | 40% MeCN | 40% MeCN |               | 40% MeOH       | 40% MeCN | 40% MeCN |
| p-nitrophenol     | 11.97          | 8.68     | 9.45     | amitriptyline | 37.99          | 24.35    | 189.98   |
| p-nitroaniline    | 13.34          | 8.97     | 12.45    | atropine      | 32.95          | 27.31    | 285.62   |
| naphthol          | 14.70          | 8.78     | 8.41     | desipramine   | 39.04          | 25.66    | 300.00   |
| dichlorophenamide | 22.96          | 10.26    | 8.96     | dibucaine     | 31.82          | 26.17    | 132.27   |
| clofoctol         | 18.00          | 16.63    | 9.21     | dipyridamole  | 17.45          | 26.89    | 201.88   |
| furosemide        | 15.10          | 8.44     | 8.21     | ephedrine     | 31.17          | 33.38    | 300.30   |
| Sulfadiazine      | 12.11          | 9.13     | 8.99     | imipramine    | 39.04          | 24.79    | 244.53   |
| Sulfabenzamide    | 13.84          | 8.70     | 8.51     | nicergoline   | 9.31           | 25.87    | 122.47   |

Table 2: Comparison of retention times of specific analytes at methanol or acetonitrile at different mobile phases.

is that the retention time was increased for the majority of the analytes. More specifically, the increase was higher in the case of analytes containing ionisable groups but minor for the analytes that are either ionized at a lower extent or lack such moieties.

### Factors affecting the retention mechanism

The compiled models of correlation along with the VIP table (Supplementary Information; Table D) and the loadings' column plots (Figures 4-6), brought about not only the impact of the descriptors on the retention time in priority order, but also whether the effect of each descriptor was positive or negative.

- **The effect of logD**

It is known that the logD value is related to the distribution of the ionic form of a molecule in water against the non-ionized species in octanol. Among the analytes studied, there are compounds that lack both acidic and basic groups and their logD value coincides with their partition coefficient (logP).

On the other hand there are compounds that have different logD and their elution is usually delayed. As it is shown in the loadings' plots (Figures 4-6), this delay is longer when the increase in the logD is higher which means that the ionization degree is higher (IP%). The phenomenon is even more intense in the case that the ionization exceeds 45%. The above is imprinted in the VIP (Variable Importance) (Supplementary Information; Table D) and its effect on the retention

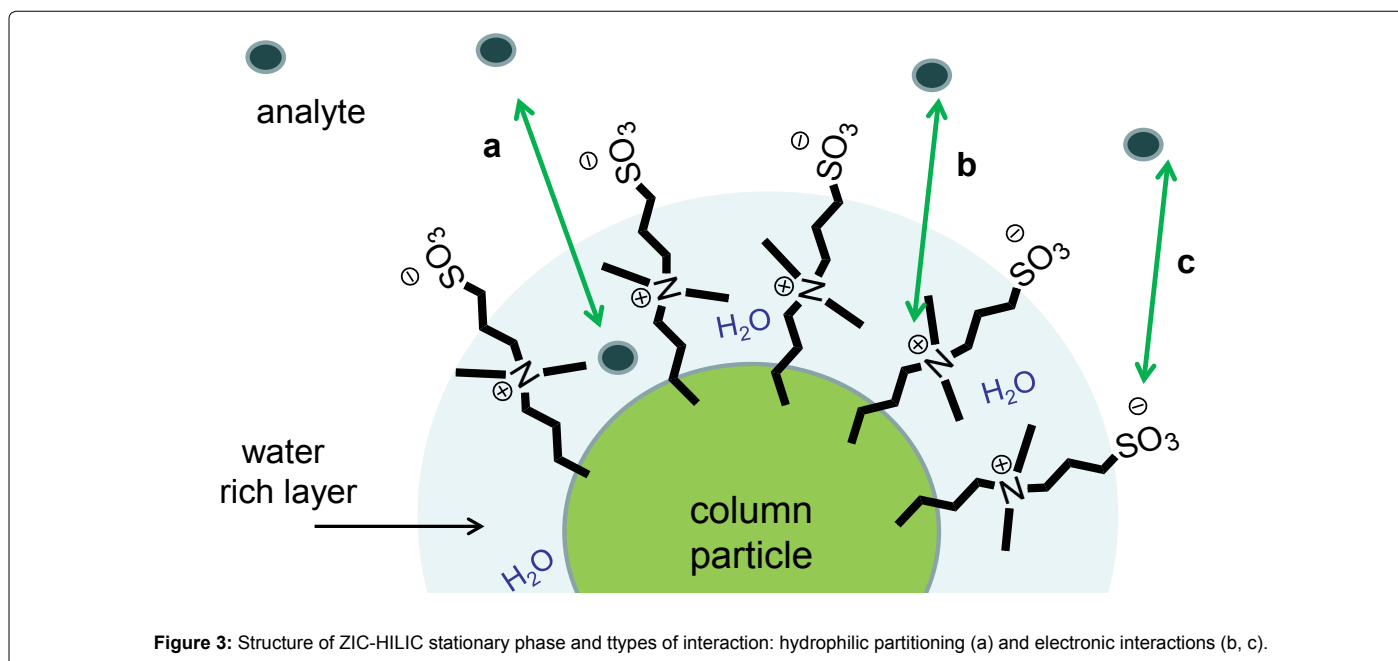


Figure 3: Structure of ZIC-HILIC stationary phase and types of interaction: hydrophilic partitioning (a) and electronic interactions (b, c).

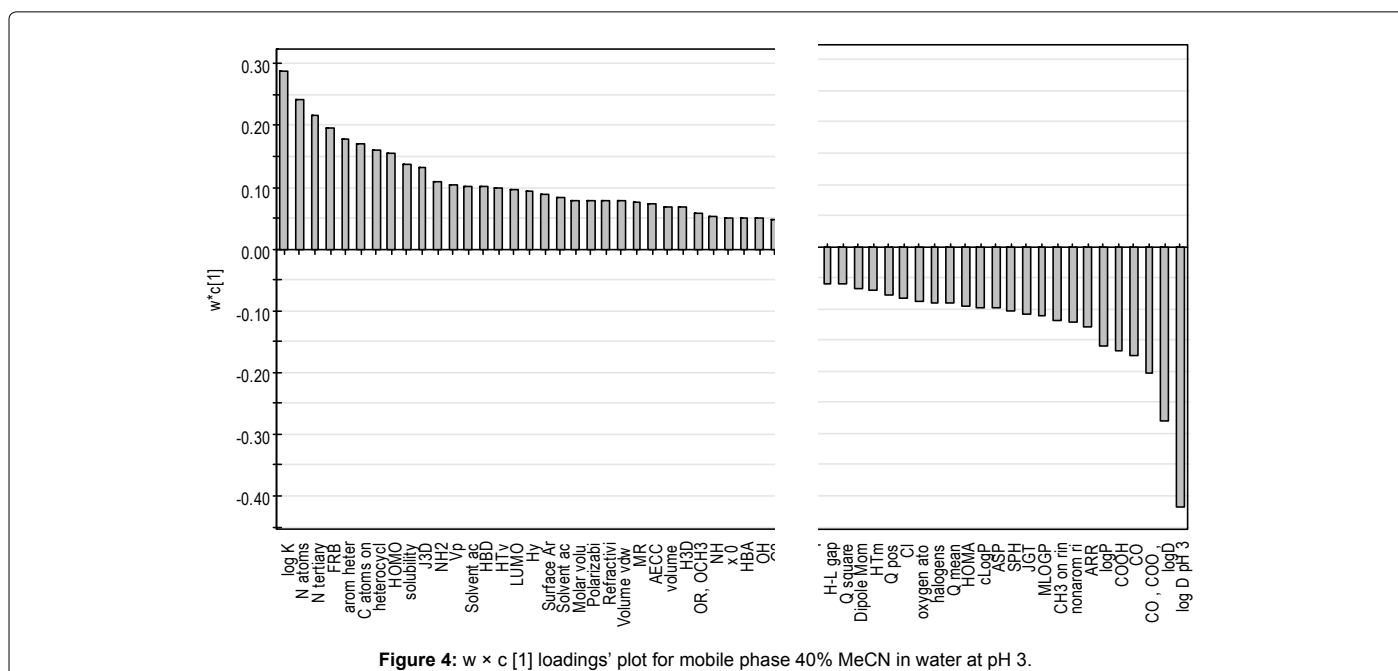


Figure 4: w × c [1] loadings' plot for mobile phase 40% MeCN in water at pH 3.

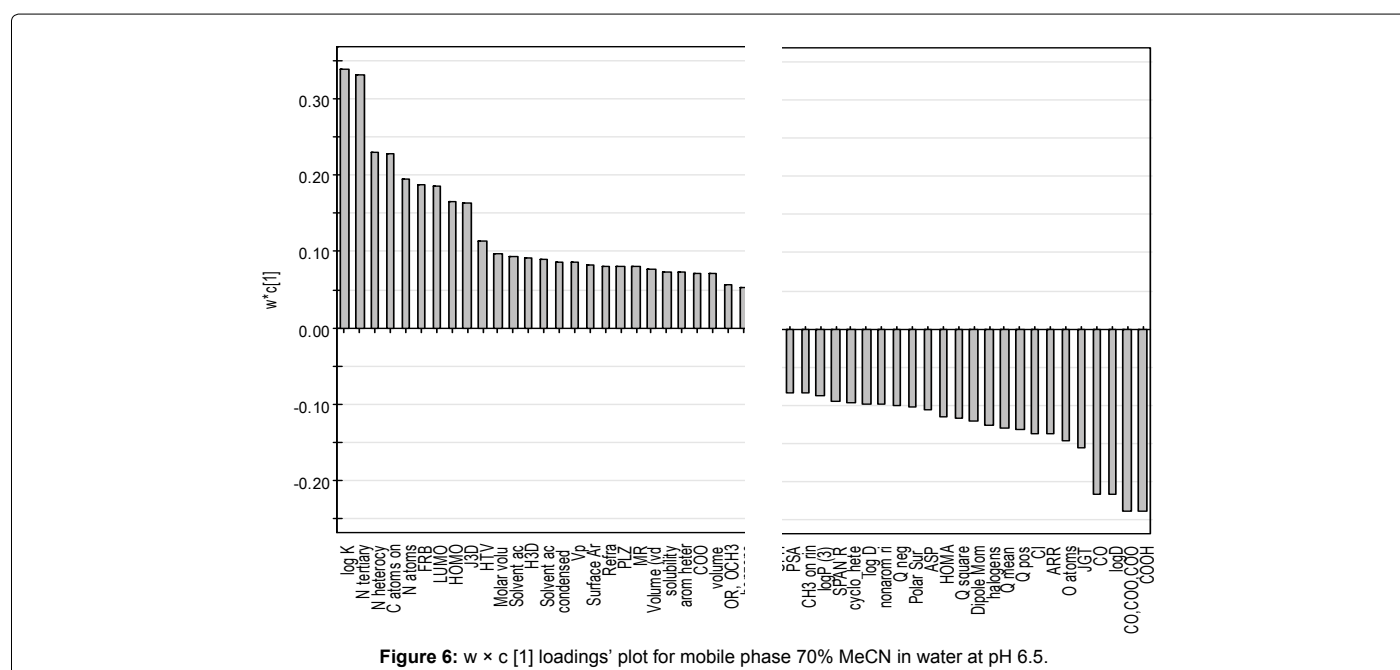
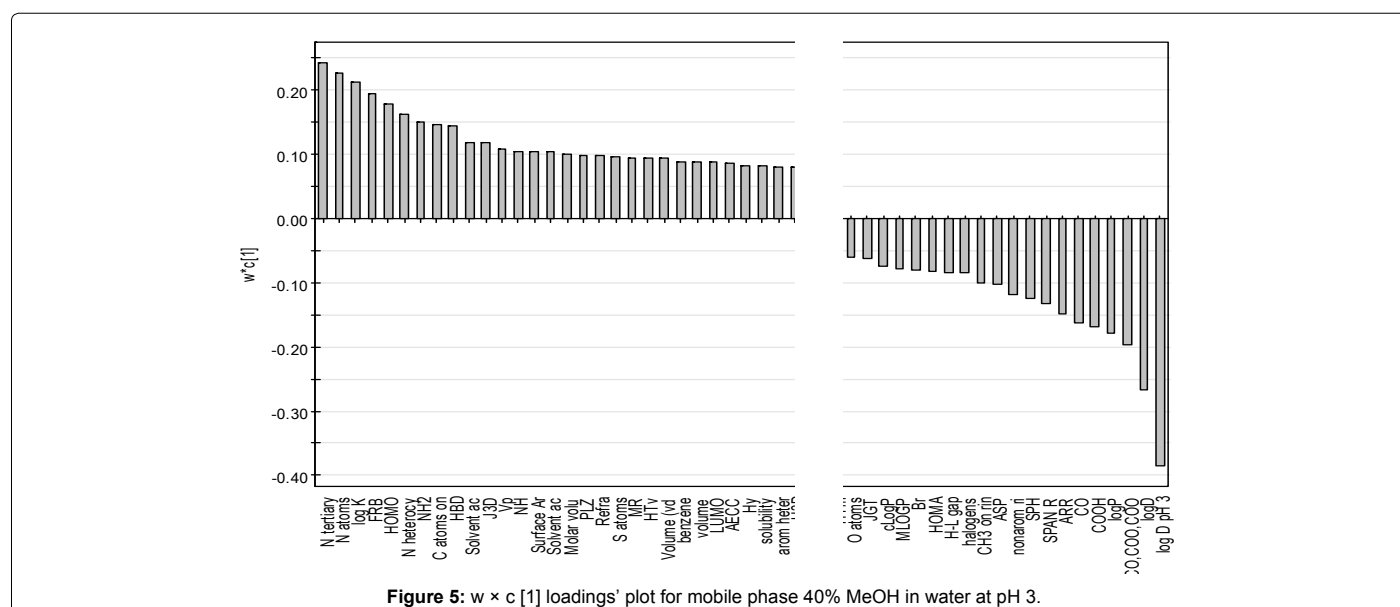
behaviour is similar for both mobile phases, i.e., regardless the organic modifier (methanol or acetonitrile). At pH 6.5, the effect of logD is similar, although the impact of this descriptor (VIP value) is relatively lower compared to the effect observed at pH 3. This can be explained because the groups ionized at pH 3 are mainly functional groups with basic properties which interact with the sulfonyl group of the stationary phase which is more active. On the other hand, at pH 6.5 more analytes are in their ionic form, but at least for half of them the ionized group has acidic properties and the ionic interaction with the ammonium group of the stationary phase is partly hindered.

• **The effect of the analytes' structure**

The presence of nitrogen atoms (N tertiary, NH, NH<sub>2</sub> etc.) as well as the presence of functional groups containing oxygen atoms (such as carbonyl, ester, carboxylic groups) are also important for the retention

of the analytes. These descriptors have the same effect when the pH value is 3 at both mobile phases (MeOH or MeCN) as shown in the figures illustrating the loadings' plots of the total models (Figures 4 and 5) as well as at pH 6.5 (Figure 6).

More specifically, the presence of N atoms (e.g., N atoms, N tertiary, N heterocyclic) cause an increase in the retention time, because these groups are affected by the pH value of the mobile phase. At pH 3 the percentage of ionic micro species of analytes containing N atoms is higher because they increase the basicity of such molecules. Consequently, ionic interaction with the stationary phase (Figure 3c) is extended and the polar compounds will be delayed due to this mechanism, which is prevailing. Many of the analytes studied have functional groups with basic properties which are charged and therefore will be retained longer.



However, one should not overlook, that the presence of a traditionally basic group (e.g.,  $\text{NH}_2$ ,  $\text{NH}$ ) will cause an increase in the retention time, but this effect is less important compared to the other N bearing groups that were previously discussed. A more detailed study of the molecules' structure could give the answer to this disagreement.

There are analytes that contain amine groups which will increase their ionization and therefore their delay. Apart from them, there are also compounds which contain amine groups but neighbouring carbonyl groups (e.g.,  $\text{CO}$ ,  $\text{COO}$ ). In this case, the basic character of these moieties ( $\text{NH}$ ,  $\text{NH}_2$ ) will diminish and in some cases they will act as weak acids. Hence, according to the loadings' plots, the descriptors  $\text{NH}$ ,  $\text{NH}_2$ , are of lower importance in the increase of the elution time while at the same time the presence of groups containing oxygen will favour faster elution. Obviously, the structure of a molecule is of vital importance, since it is the main reason for its ionization ability.

#### • The effect of solubility

Another descriptor that appears to increase the retention of analytes on the HILIC column is the solubility of compounds in water. Explanation of the solubility effect leads to the conclusion that on HILIC column there is another mechanism that works, which is partition. It is known that the solubility of compounds is the minimum amount of a solute that can be dissolved in a given amount of a solvent (in this case water).

At pH value equal to 3 some analytes are encountered in their ionic form. Meanwhile, the solubility of compounds is pH dependent [35]. The solvation of analytes in water is better when a compound is in its ionic form; the solubility of charged species increases. Even in the case of molecules whose ionization is not sufficient (<45%), if their solubility is high the retention is increased. The analytes are dissolved in the water portion of the mobile phase and they are in a way trapped in the water rich layer surrounding the column particle enhancing the partition mechanism and resulting in the delay of the analytes' elution.

Hence, there is higher hydrophilic interaction of analytes with the water layer on the column bed (Figure 3a). The analyte partitions into and out of the adsorbed water layer (partition mechanism) and these forces work along with the ionic interactions and the retention of the analytes will increase.

#### • Other descriptors

Other descriptors that also affect the retention causing a delayed elution are the volume of analytes which is directly related to the carbon atoms on side chains and the number of free rotatable bonds (FRB). Large molecules with long side chains are ponderous but their effect on the retention time is secondary since the ionic interactions prevail.

The  $w^*c$  loadings' plot in the case of the alcoholic mobile phase shows that the presence of halogen atoms is responsible for faster elution, although this effect is not one of the most important factors. The presence of halogen atoms seems to be a determining factor compared to older studies in lipophilic columns where the effect of halogen atoms was of secondary importance in the retention mechanism and cause an increase in the retention of the analytes [22]. Probably, such molecules are able to form halogen bonds (halogen interactions) with methanol. At the same time the solubility of molecules is working at the same direction and the solvation of analytes in the mobile phase is facilitated not only due to their low solubility in water, but also because of their ability to interact with molecules of methanol. The halogen bond is relatively strong and may favour faster elution.

Another contributing factor to halogen bond strength comes from

the short distance between the halogen containing Lewis acid (donor) and the Lewis base (acceptor), where X is a halogen. The attractive forces developed result in the distance between the donor and acceptor to be shorter than the sum of van der Waals radii. The interaction becomes stronger as the distance decreases between the halogen Lewis acid and the Lewis base.

The model compiled in the case that the mobile phase contains acetonitrile but the pH value is adjusted to 6.5 shares several similarities with the previous two models. Among the most important is the effect of the presence of the N atoms in the structure of the analyte which is responsible for the retention of the analytes. The molecules containing nitrogen atoms which are not attached to a carbonyl group seem to enhance the ionization and cause an increase in the retention time.

It is vital to mention at this point that due to the presence of the stagnant water layer on the surface of the column bed, the ionic interactions with the sulfonic group are stronger compared to the interactions of the quaternary ammonium group because of the hindrance caused by the water molecules. Therefore, the sulfonyl group is more active and will be able to retain positively charged molecules longer.

The above is verified due to the fact that in the VIP values the effect of the charge is more determinant as well as the global topological charge (JGT). This increase in the impact of the charge is also based on the fact that the number of analytes that are in their ionic form is bigger; hence the effect in the total mechanism of this descriptor is more vital.

**Individual models:** Apart from the total models developed for the sum of the observations, individual models were compiled as well (Supplementary Information; Figure D-1). According to the statistical results of Table 1, the most reliable model which corresponds to analytes that have IP% <45% are more reliable in acidic conditions. Of course this result is not associated to the ionization of the analytes but mainly to the ability of MeCN to have a more consistent behaviour. This consistency is determined due to the fact that MeCN is polar but aprotic and does not participate in any type of interactions such as hydrogen bonding between the analytes and the mobile phase or the stationary phase and the mobile phase. On the other hand, MeOH may participate in such interactions and compete with the analyte for the free silanol groups of the stationary phase or create weak bonds with the analyte molecules.

Especially, in the case of MeOH the presence of halogen atoms is a determinant factor causing a decrease in the retention time of analytes. This phenomenon has been explained in a previous paragraph based on the halogen bonds developed.

Concerning the effect of the HOMO-LUMO energies on the retention time, it seems to be a bit complicated. First of all, for ionized species the effect of LUMO is more important causing an increase in the retention time. An analyte with high LUMO value acts as a Lewis base ( $\pi$ -base) which interacts more easily and is retained on the sulfonic group of the stationary phase.

The same effect was observed for the total models but the impact was not of the same importance. As mentioned, a factor affecting the retention of these molecules, despite the fact that their ionization is not high, is their structure and mainly the presence of nitrogen atoms. This can be explained since the presence of such groups promotes the ionization of the analytes. Even when ionization is limited, there are cases that the molecule is able to be retained due to the stabilization of the polar molecules which occurs due to the high energies of HOMO orbitals. At the same time, high HOMO energies will enhance stability

[36,37]. This effect is obvious for compounds whose ionization at this pH is high and it is more probable for compounds that are not ionized but have high polar surface areas due to the presence of moieties where delocalization of electrons can easily occur. The difference in the priority order is logical because the total models contain analytes which are not ionized and they are eluted in short time despite their HOMO - LUMO values.

In the case of group A (Figure 2), high HOMO seems to increase the retention time but the impact of this descriptor is relatively low. The impact of the LUMO energies also seems to be of secondary importance, except in the case where the mobile phase contains methanol. For this model, analytes with increased LUMO energies appear to have faster elution times, more probably because the LUMO of the analyte interacts with the HOMO of MeOH and this electron interaction facilitates faster elution.

When the mobile phase has pH value equal to 6.5, then there are more analytes that are encountered in their ionic form. In this instance, apart from the solubility of the analytes in water which is determinant for their delay, the retention time is also increased when the analyte is a hydrogen bond donor (high HBD value) and in the case that it contains moieties that favor this behavior.

The hydrophilicity of the analytes, due to the presence of hydroxyl groups along with their ability to accept protons (high HBA), facilitates hydrogen bonding between free silanol groups or the water molecules covering the stationary phase and the analyte, or the interaction between the analyte and the charged groups of the stationary phase. The above is directly related to the increased values of the polar surface area of the analytes which is also responsible for the retention time increase.

An important observation is that the lipophilicity of the analytes is not the major factor affecting the retention in the total models. Moreover, in all cases the effect is opposite compared to the reversed phase chromatography. It was obvious that when the molecules could be ionized then the logP has minor effect on the retention (causing a decrease on the retention time). For analytes that are not ionized, logP is also responsible for the decrease on the retention time but this is not the rule because at high pH (6.5) logP is of secondary importance.

## Conclusion

The retention behaviour on a ZIC-HILIC column is mainly affected by the ionization of the analytes. Still, a more thorough investigation of several characteristics of the analytes revealed that apart from the structure of a molecule, other parameters affect the retention as well. According to the VIP values and the loadings' plots for both mobile phases at pH 3, the most important factors that increase the retention are the presence of N atoms, the solubility of the compounds in water, as well as the solvent accessible surface area of molecules. Importantly, high logD values of the analytes decreases the elution time since the main mechanism is based on ionic interactions. The stagnant water rich layer that is covering the surface of the stationary phase may limit the interaction of the inner quaternary ammonium group with acidic groups of the analytes. At pH 3 these groups are not charged and the electrostatic interactions are very weak resulting in negligible delay if any.

## References

- Alpert AJ (1990) Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds. *J Chromatogr* 499:177-196.

- Heaton J, Gray N, Cowan DA, Plumb RS, Legido-Quigley C, et al. (2012) Comparison of reversed-phase and hydrophilic interaction liquid chromatography for the separation of ephedrine. *J Chromatogr A* 1228: 329-337.
- Buszewski B, Noga S (2012) Hydrophilic interaction liquid chromatography (HILIC)—a powerful separation technique. *Anal Bioanal Chem* 402: 231-247.
- Tache F, Udrescu S, Albu F, Micala F, Medvedovici A (2013) Greening pharmaceutical applications of liquid chromatography through using propylene carbonate-ethanol mixtures instead of acetonitrile as organic modifier in the mobile phases. *J Pharm Biomed Anal* 75: 230-238.
- Dos Santos Pereira A, David F, Vanhoenacker G, Sandra P (2009) The acetonitrile shortage: is reversed HILIC with water an alternative for the analysis of highly polar ionizable solutes? *J Sep Sci* 32: 2001-2007.
- Dai J, Carr PW (2009) Effect of mobile phase anionic additives on selectivity, efficiency, and sample loading capacity of cationic drugs in reversed-phase liquid chromatography. *J Chromatogr A* 1216: 6695-6705.
- Bernal J, Ares AM, Pól J, Wiedmer SK (2011) Hydrophilic interaction liquid chromatography in food analysis. *J Chromatogr A* 1218: 7438-7452.
- Hemstrom P, Irgum K (2006) Hydrophilic interaction chromatography. *J Sep Sci* 29: 1784-1821.
- Jandera P (2011) Stationary and mobile phases in hydrophilic interaction chromatography: a review. *Anal Chim Acta* 692: 1-25.
- McCalley DV (2007) Is hydrophilic interaction chromatography with silica columns a viable alternative to reversed-phase liquid chromatography for the analysis of ionisable compounds? *J Chromatogr A* 1171: 46-55.
- McCalley DV (2008) Evaluation of the properties of a superficially porous silica stationary phase in hydrophilic interaction chromatography. *J Chromatogr A* 1193: 85-91.
- Hao Z, Xiao B, Weng N (2008) Impact of column temperature and mobile phase components on selectivity of hydrophilic interaction chromatography (HILIC). *J Sep Sci* 31: 1449-1464.
- Buckenmaier SM, McCalley DV, Euerby MR (2002) Overloading study of bases using polymeric RP-HPLC columns as an aid to rationalization of overloading on silica-ODS phases. *Anal Chem* 74: 4672-4681.
- Po HN, Senozan NM (2001) Henderson-Hasselbalch Equation: Its History and Limitations. *J Chem Educ* 78: 1499-1503.
- McCalley DV (2010) Study of the selectivity, retention mechanisms and performance of alternative silica-based stationary phases for separation of ionised solutes in hydrophilic interaction chromatography. *J Chromatogr A* 1217: 3408-3417.
- Kouskoura MG, Hadjipavlou-Litina D, Markopoulou CK (2014) Elucidation of the retention mechanism on a reverse-phase cyano column by modeling. *J Sep Sci* 37: 1919-1929.
- Kouskoura MG, Kachrimanis KG, Markopoulou CK (2014) Modeling the drugs' passive transfer in the body based on their chromatographic behavior. *J Pharm Biomed Anal* 100: 94-102.
- Baczek T, Kalisz R, Novotna K, Jandera P (2005) Comparative characteristics of HPLC columns based on quantitative structure-retention relationships (QSRR) and hydrophobic-subtraction model. *J Chromatogr A* 1075: 109-115.
- Buszewski B, Kowalska S, Kowalkowski T, Rozpedowska K, Michel M, et al. (2007) HPLC columns partition by chemometric methods based on peptides retention. *J Chromatogr B Analyt Technol Biomed Life Sci* 845: 253-260.
- Visky D, Vander Heyden Y, Ivanyi T, Baten P, De beer J, et al. (2003) Characterisation of reversed-phase liquid chromatographic columns by chromatographic tests. Rational column classification by a minimal number of column test parameters. *J Chromatogr A* 1012: 11-29.
- Ivanyi T, Vander Heyden Y, Visky D, Baten P, De beer J, et al. (2002) Minimal number of chromatographic test parameters for the characterisation of reversed-phase liquid chromatographic stationary phases. *J Chromatogr A* 954: 99-114.
- Kouskoura MG, Mitan CV, Markopoulou CK (2015) Chemometric of the retention mechanism on butyl column: effect and relation of pH and pKa. *Se Pu* 33: 1274-1286.



23. Chirita RI, West C, Zubrzycki S, Finaru AL, Elfakir C (2011) Investigations on the chromatographic behaviour of zwitterionic stationary phases used in hydrophilic interaction chromatography. *J Chromatogr A* 1218: 5939-5962.
24. Euerby MR, Hulse J, Petersson P, Vazhentsev A, et al. (2015) Retention modelling in hydrophilic interaction chromatography. *Anal Bioanal Chem* 407: 9135-9152.
25. Todeschini R, Consonni V, Pavan M (2012) DRAGON. Talete srl.
26. Sander T (2014) Osiris Property Explorer (2001-2013).
27. Hyper Chem (TM) Professional, 1115 NW 4th Street, Gainesville, Florida 32601, USA.
28. Marvin (2012) ChemAxon.
29. Polak E, Ribiere G (1969) Note sur la convergence de methodes de directions conjugees. *ESAIM* 16: 35-43.
30. Kouskoura MG, Mitan CV, Markopoulou CK (2015) Chemometric of the retention mechanism on butyl column: effect and relation of pH and pKa. *Se Pu* 33: 1274-1286.
31. Haaland DM, Thomas EV (1988) Partial Least Squares Regression for Spectral Analysis 1 Relation to other Quantitative Calibration Methods and the Extraction of Qualitative Information. *Anal. Chem* 60: 1193-1202.
32. Lapins M, Eklund M, Spjuth O, Prusis P, Wikberg JE (2008) Proteochemometric modeling of HIV protease susceptibility. *BMC Bioinformatics* 9: 181.
33. Umetrics AB (2011) Simca. User guide and tutorial.
34. Balaban AT (1983) Topological indices based on Topological distances in molecular graph. *Pure Appl Chem* 55: 199-206.
35. Butler JN (1998) Ionic Equilibrium - Solubility and pH calculations. John Wiley and Sons Inc., p: 578.
36. McMurry J (2000) Organic Chemistry. Brooks Cole: Pacific Grove, USA.
37. Scudder PH (2013) Electron flow in Organic Chemistry - A decision-based guide to organic mechanisms, 2nd edn. Wiley: New Jersey and Canada.