Mechanisms of retention in HPLC

María Celia García-Álvarez-Coque
José Ramón Torres-Lapasió

Department of Analytical Chemistry
University of Valencia
Valencia, Spain

https://sites.google.com/site/fuschrom/

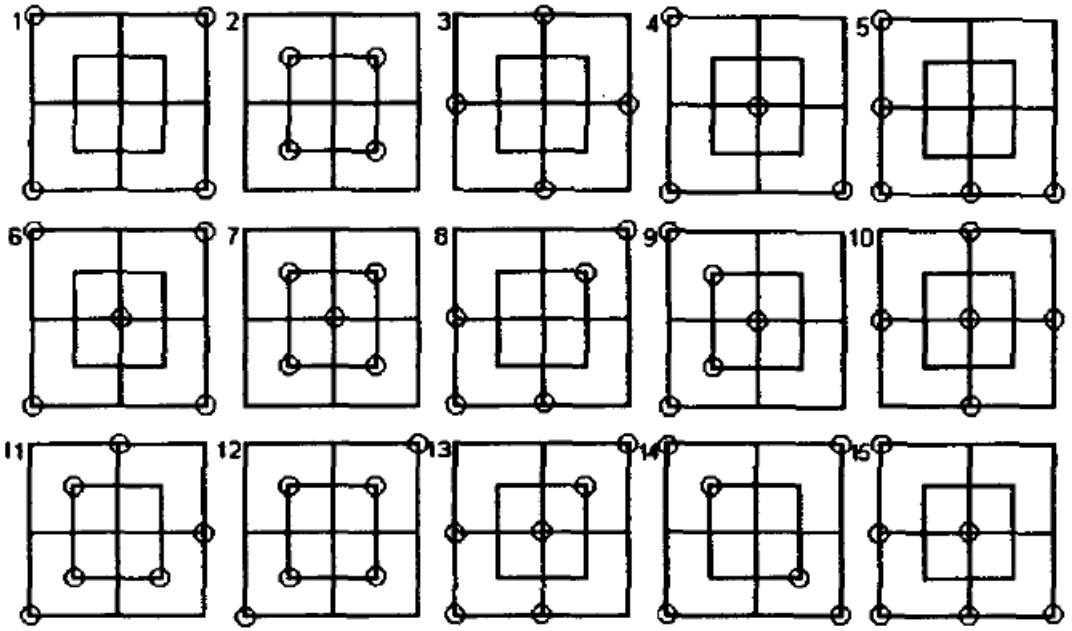
Part 4

Index

4.1. Introduction

Interpretive optimisation strategies are based on the **accurate description (modelling) of chromatographic behaviour.** The first step in these strategies consists of **gathering information** about the chromatographic behaviour of the compounds in the sample, covering a reasonably wide factor region usually attending only to retention. For this purpose, either **isocratic or gradient experiments** can be used.

Isocratic elution



Gradient elution

Experimental design : distribution of the experiments within the factor domain

For each solute, a relationship that describes the retention is obtained as a function of the experimental factors. This function will allow the prediction of isocratic or gradient retention times at different conditions.

Alternatives

- convenient algebraic expression (empirical or mechanistic)

  parameters searched on a least-squares basis

- black-box algorithm (as neural networks)

Quality (accuracy and precision) of predictions is decisive for the reliability of optimisations or any other applications, and depends on:

- information richness provided by the training data in the experimental design
- distribution of the experimental points
- equation selected to fit the training data
- fitting procedure
- ranges of experimental factors
- elution mode (isocratic or gradient)

4.2. Isocratic elution in RPLC at fixed pH

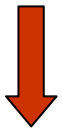
4.2.1. Use of polynomial models with hydro-organic mobile phases

The concentration of organic solvent in the mobile phase is the factor most frequently optimised in RPLC.

Advantages

- large impact on both elution strength and selectivity
- a very convenient property:

  flexibility and accuracy in the implementation of changes in this factor



Hence, the importance of obtaining reliable retention models involving the concentration of organic solvent as factor !!!

RPLC retention (solute *i*) in terms of solubility parameters (Schoenmakers et al.):

$$\ln k_i = \frac{t_{Ri} - t_0}{t_0} = \frac{v_i}{RT}\left[(\delta_M - \delta_i)^2 - (\delta_S - \delta_i)^2\right] + \ln\frac{n_S}{n_M} \qquad (4.1)$$

k_i : solute retention factor t_{Ri} : solute retention time t_0 : dead time

R : gas constant (1.9865 cal K^{-1} $mole^{-1}$) T : absolute temperature (K)

v_i : solute molar volume (cm^3 $mole^{-1}$) δ : polarity parameter ($cal^{0.5}$ $cm^{-1.5}$)

n_M : moles of mobile phase n_S : moles of stationary phase

Water (w) / **organic solvent** (o) binary mixtures:

$$\delta_M = \sum_j \delta_j \varphi_j \qquad (1.2) \qquad \rightarrow \qquad \delta_M = (1-\varphi)\,\delta_w + \varphi\,\delta_o \qquad (4.2)$$

φ : organic solvent volume fraction

$$\ln k_i = \frac{t_{Ri} - t_0}{t_0} = \frac{v_i}{RT}\left[(\delta_M - \delta_i)^2 - (\delta_S - \delta_i)^2\right] + \ln\frac{n_S}{n_M}$$

$$\delta_M = (1-\varphi)\,\delta_w + \varphi\,\delta_o$$

$$\ln k = c_0 + c_1\,\varphi + c_2\,\varphi^2 \qquad (4.3)$$

Quadratic model

Parameters c_0 , c_1 and c_2 gather all the constants and solute/mobile phase/stationary phase properties, and have particular values for a given solute and column/solvent system.

decimal logarithms

$$\log k = c_0 + c_1\,\varphi + c_2\,\varphi^2 \qquad (4.4)$$

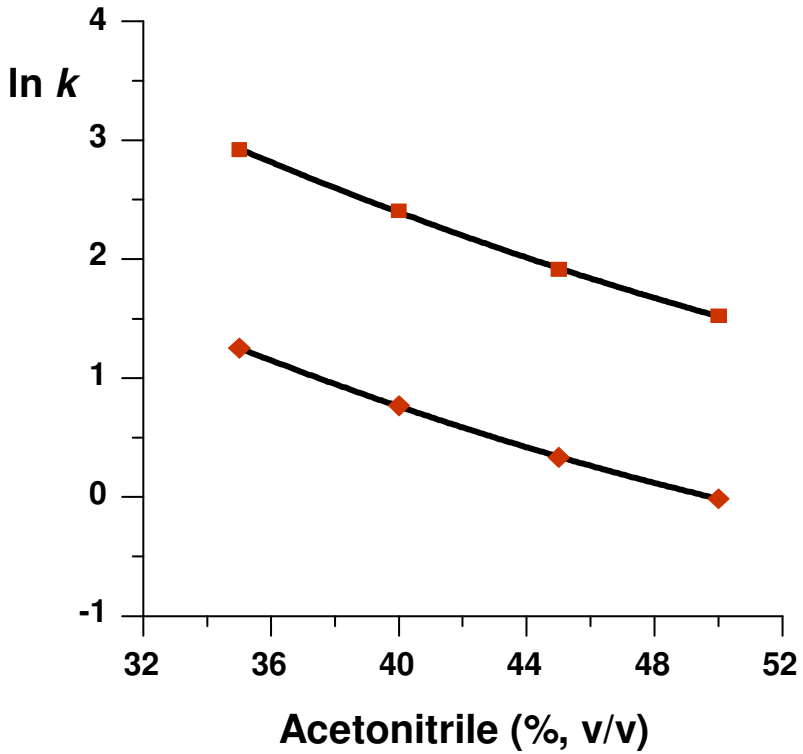
$$\log k = c_0 + c_1\,\varphi \qquad (4.5)$$

linear model

narrow concentration ranges

These equations are also valid in the presence of specific interactions in the presence of additives.

Two flavonoids eluted with acetonitrile/water mixtures



$$\log k = c_0 + c_1 \varphi + c_2 \varphi^2$$

$$\log k = c_0 + c_1 \varphi$$

Linear model

$$\log k = c_0 + c_1 \varphi = \log k_w - S \varphi \qquad (4.5)$$

c_0 (intercept): extrapolated log k value if water were the mobile phase

S (slope): sensitivity of retention to changes in the organic solvent content (elution strength)

- Simplifies the retention model to a straight-line.
- Valid for narrow concentrations of organic solvent, which is usual in RPLC.
- In wide concentration ranges, deviations from linearity are observed.
- The deviations are especially significant at the highest and lowest concentrations of organic solvent.

Ternary systems (two organic solvents and water)

$$\log k = c_0 + c_1 \varphi_1 + c_2 \varphi_2 + c_{12} \varphi_1 \varphi_2 + c_{11} \varphi_1^2 + c_{22} \varphi_2^2 \qquad (4.6)$$

φ_1 , φ_2 : concentrations of organic solvents

The search for an appropriate model has not finished here !!!

- Review Nikitas and Papa-Louisi (J. Chromatogr. 1216 (2009) 1737–1755)
- Most models are logarithmic, but there are other proposals.
- Model proposed by Jandera for NPLC, applied to RPLC:

excellent accuracy in RPLC → $\frac{1}{k} = [a + b\varphi]^n \qquad (4.7)$

More retention models !!!

$$k = e^{\ln k_w - a\varphi + b\varphi^2}$$

$$k = e^{\ln k_0 (1 - a\varphi(1-\varphi))}$$

$$k = e^{\ln k_w - a\frac{\varphi}{1+b\varphi}}$$

$$k = e^{\ln k_0 - \frac{a\varphi}{1+2\varphi}}$$

$$k = e^{\ln k_w - \ln(1+b\varphi) - c\frac{\varphi}{1+b\varphi}}$$

$$k = e^{\ln k_w - \ln(1+2\varphi) - c\frac{\varphi}{1+2\varphi}}$$

$$k = \frac{e^{\ln k_0}}{(1+b\varphi)^4}$$

$$k = \frac{e^{\ln k_o}}{1+a\varphi+b\varphi^2}$$

$$k = \frac{e^{\ln k_0}}{(1+b\varphi)^3}$$

$$k = e^{\ln k_o}\frac{1-a\varphi}{1+b\varphi}$$

$$k = \frac{e^{\ln k_0}}{1+b\varphi(1+b\varphi)}$$

$$k = e^{\ln k_0 - a\varphi}$$

$$k = e^{\ln k_0 - a\ln\varphi/\varphi_o}$$

4.2.2. Polarity models

Linear relationships between log *k* and mobile phase polarity parameters that depend on solvent content are interesting models.

- Dimroth-Reichardt parameter $E_T(30)$
- Normalised $E_T^N(30)$

limited range of linearity : 0 – 80% acetonitrile and 20 – 100% methanol

- Normalised polarity parameter proposed by Bosch and Rosés P_M^N

extends the linearity to 0 – 100% acetonitrile and methanol

For acetonitrile-water:

$$P_M^N = 1.00 - \frac{2.068\varphi}{1 + 1.341\varphi} \qquad (4.8)$$

For methanol-water:

$$P_M^N = 1.00 - \frac{1.33\varphi}{1 + 0.47\varphi} \qquad (4.9)$$

A model was developed, which isolates the polarity contributions of the three agents involved in the separation system (solute, stationary phase and mobile phase).

$$\log k = (\log k)_0 + p_S\,(P_M^N - P_S^N) \qquad (4.10)$$

Descriptors

P_M^N : mobile phase polarity

P_S^N : stationary phase polarity

$(\log k)_0$: retention in a hypothetical mobile phase with the same polarity as the stationary phase

p_S : solute polarity

The separation of the polarities of solute, mobile phase and stationary phase in the polarity model is not perfect: p_S mainly gathers the solute contributions. However, it is not an absolute value, but a relative measurement that depends on the column environment (stationary and mobile phase nature).

$$\log k = (\log k)_0 + p_S\,(P_M^N - P_S^N) \qquad (4.10)$$

fitted with the data

for a set of compounds

Advantages of the polarity model

- **A simple model: only needs knowledge of**

  one parameter per solute: p_S

  two descriptors for the column ($(\log k)_0$ and P_S^N)

  P_M^N is obtained from the mobile phase composition

- **This facilitates transference of *k* values to new columns and solvent systems by establishing simple correlations between p_S-values for small sets of selected reference compounds.**

- When $(\log k)_0$ and P_S^N are fitted individually for each solute:

$$\log k = q + p_s P_m^N = c_0 + c_1 P_m^N \qquad (4.11)$$

$$\log k = (\log k)_0 + p_S (P_M^N - P_S^N) \qquad (4.10)$$

Predictions performed by imposing common column parameters are necessarily poorer than those when all parameters in the equation are fitted for each individual solute, which are preferable for optimisation purposes !!!

$$\log k = c_0 + c_1 \varphi$$

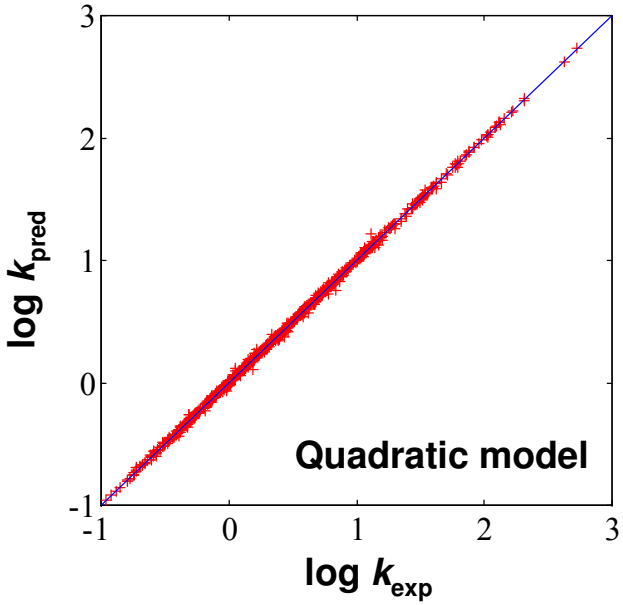
$$\log k = c_0 + c_1 P_M^N$$

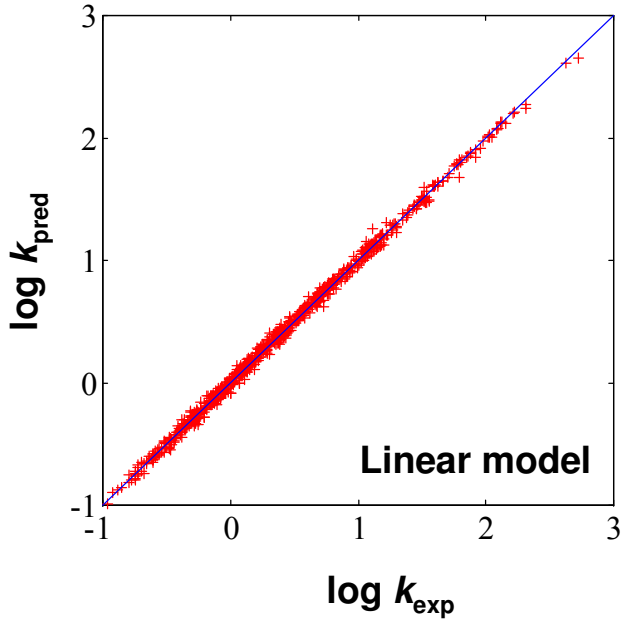
two-parameter models

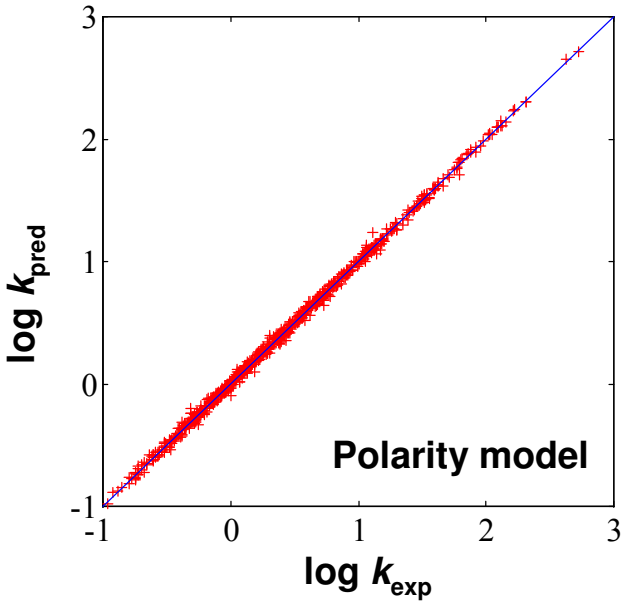
$$\log k = c_0 + c_1 \varphi + c_2 \varphi^2$$ ← three-parameter model

The two-parameter polarity model is valid for wider concentration ranges and comparable in accuracy to the quadratic model !!!

152 compounds eluted with methanol / water







4.2.3. Micelles and ionic liquids in the mobile phase

Among the chromatographic modes using mobile phases with additives, micellar liquid chromatography deserves a special mention.

Modelling

- Polynomial equations
- Mechanistic models based on equilibrium considerations
  - ⇒ gave a meaning to the coefficients in the polynomial equations
  - ⇒ improved understanding of the retention mechanism

Pure MLC (mobile phases containing only surfactant):

$$k = \frac{K_{AS}}{1 + K_{AM}[M]} \qquad (4.12)$$

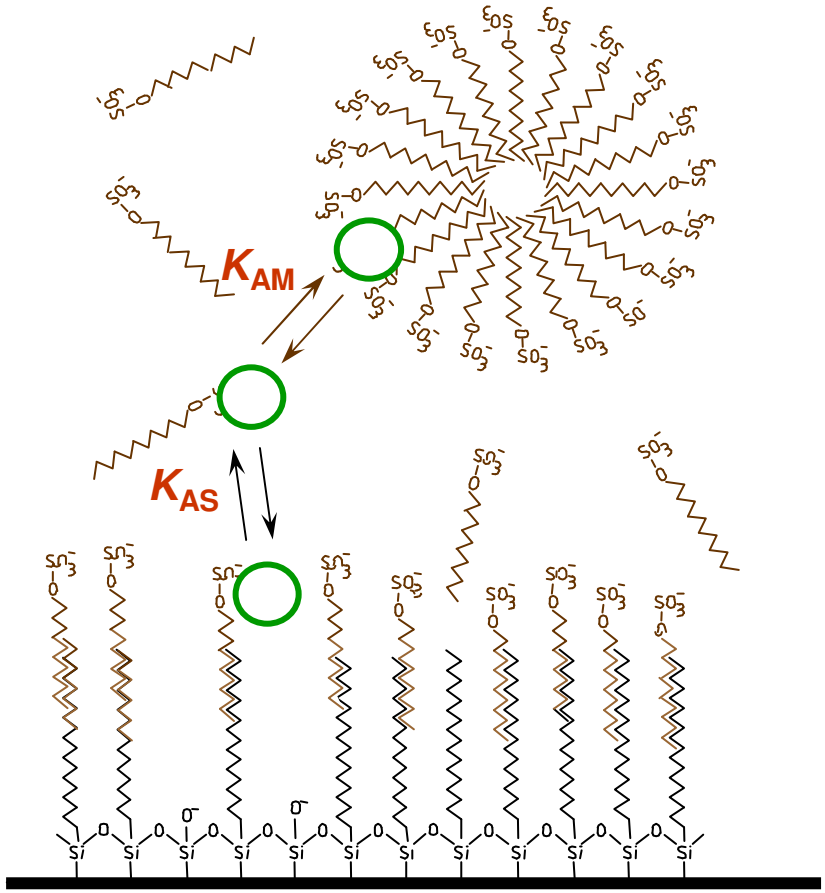
$$\frac{1}{k} = \frac{1}{K_{AS}} + \frac{K_{AM}}{K_{AS}}[M] = c_0 + c_1[M] \qquad (4.13)$$

$[M]$: concentration of surfactant molecules involved in micelle formation

K_{AS} : partition of the solute between bulk water and stationary phase

K_{AM} : partition of the solute between bulk water and micelle

Pure MLC



Hybrid MLC

Hybrid micellar medium (mobile phases containing both surfactant and organic solvent):

$$k = \frac{K_{AS}\,\dfrac{1+K_{SD}\,\varphi}{1+K_{AD}\,\varphi}}{1+K_{AM}\,\dfrac{1+K_{MD}\,\varphi}{1+K_{AD}\,\varphi}\,[M]} \qquad (4.14)$$

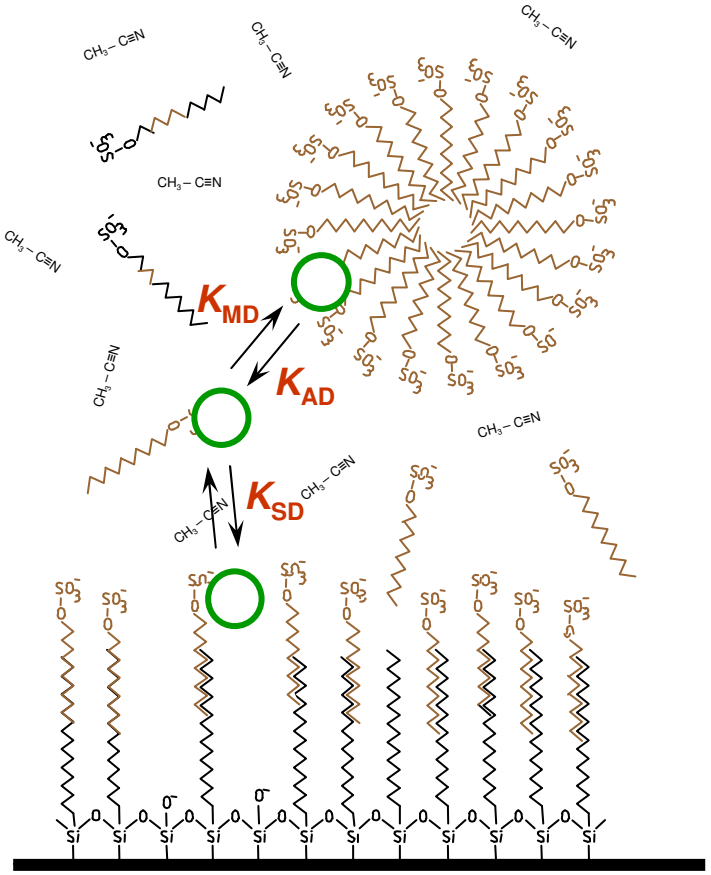
K_{SD}, K_{AD} and K_{MD} : account for the displacement of the partition equilibria by the organic solvent

K_{SD} only for highly hydrophobic solutes. In other cases:

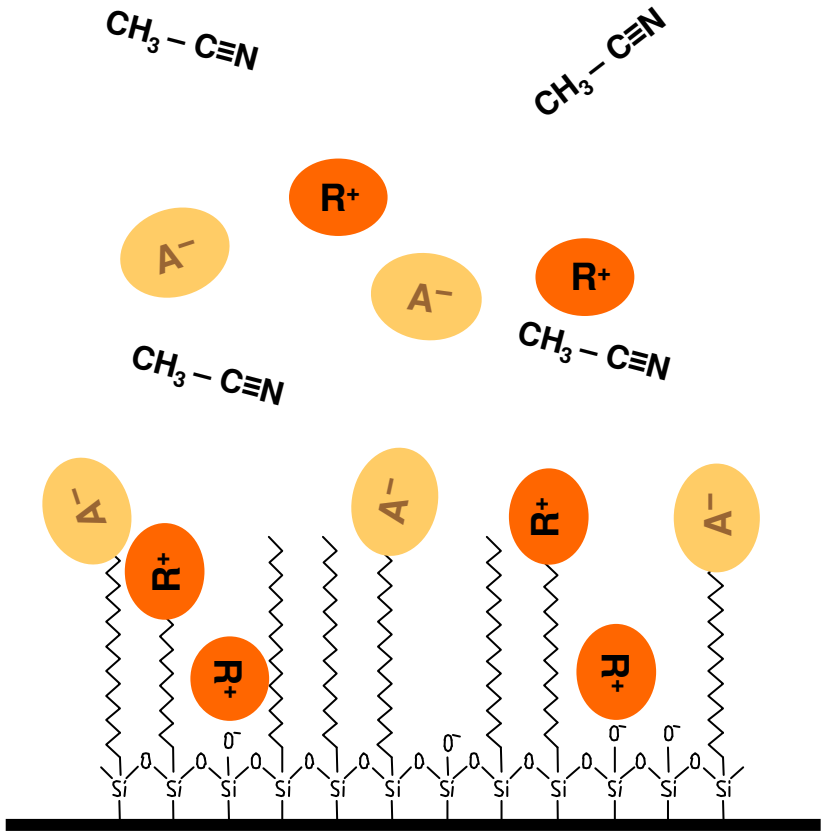
$$k = \frac{K_{AS}\,\dfrac{1}{1+K_{AD}\,\varphi}}{1+K_{AM}\,\dfrac{1+K_{MD}\,\varphi}{1+K_{AD}\,\varphi}\,[M]} \quad \rightarrow \quad \log k = c_0 + c_1\,\varphi + c_2\,[M] + c_{12}\,\varphi\,[M] \qquad (4.15)$$

Eq. (4.15) is also valid for mobile phases containing ionic liquids, where the molar concentration of ionic liquid [*IL*] is used instead of [*M*].

MLC



RPLC with ionic liquids



4.2.4. The silanol effect

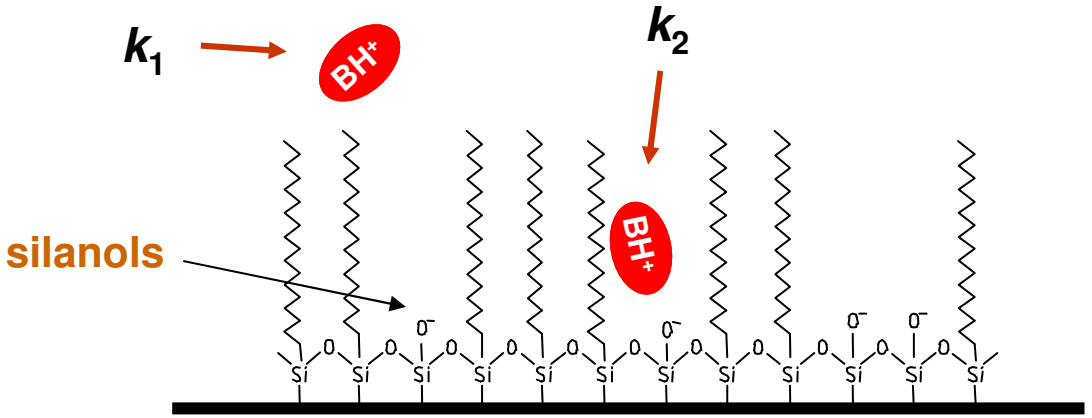
Basic compounds experience a combined retention mechanism when separated with bonded-silica stationary phases and organic solvent-water mixtures that includes **hydrophobic** (**partition into the bonded phase**), and **silanophilic interactions** (**attraction of the cationic solutes to the anionic silanols**). In the absence of any additional interaction, masking of silanols will decrease the retention.

$$k_0 = k_1 + k_2 \qquad (4.16)$$

k_0 : retention factor with silanols totally exposed for interaction

k_1 : hydrophobic contribution

k_2 : silanophilic contributions



Horváth model

Evaluation of the silanol suppressing potency by addition of amines

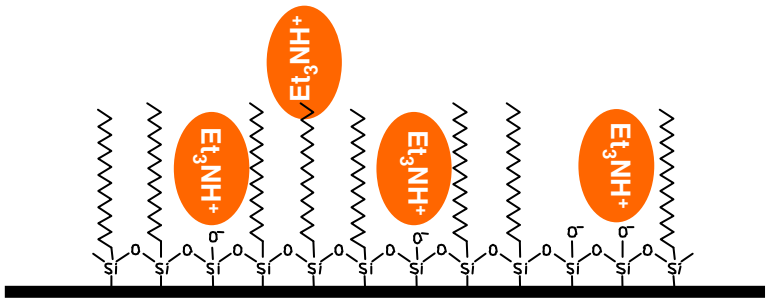
without additive → $k_0 = k_1 + k_2$ (4.16)

Secondary equilibrium $-SiO^- + A^+ \leftrightarrows -SiO^- A^+$ (4.17)

with additive → $k = k_1 + \dfrac{k_2}{1 + K_A[A]}$ (4.18)

K_A : silanol / protonated amine binding constant

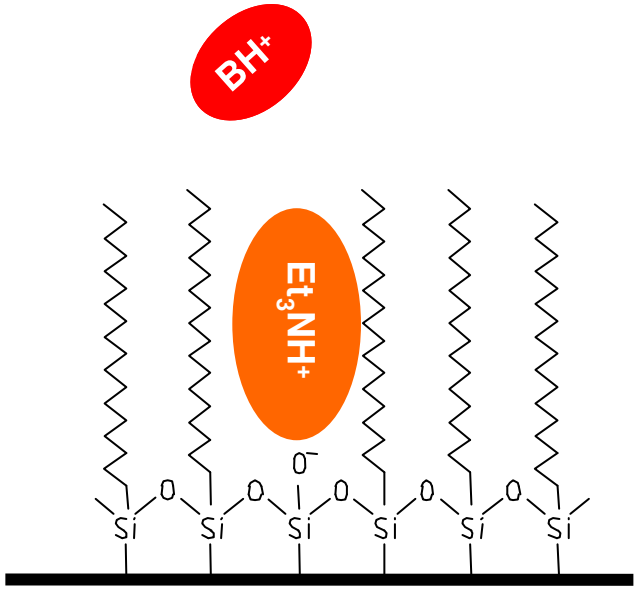
$[A]$: amine molar concentration



Silanophilic contribution to retention

$$k_0 - k = k_2 - \frac{k_2}{1 + K_A[A]} = \frac{k_2 K_A [A]}{1 + K_A [A]} \quad (4.19) \rightarrow \quad \frac{[A]}{k_0 - k} = \frac{1}{k_2 K_A} + \frac{[A]}{k_2} \quad (4.20)$$

$$\frac{[A]}{k_0 - k} = \frac{1}{k_2 K_A} + \frac{[A]}{k_2} \qquad (4.20)$$

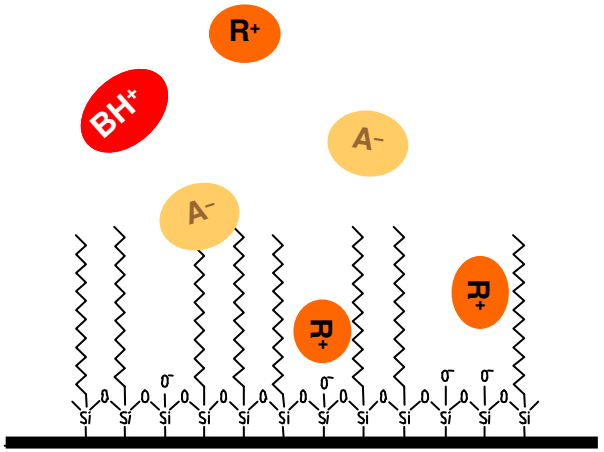


The binding constant K_A :

- evaluates the ability of the amine to block silanol sites
- can be obtained by regressing the left term versus the concentration of amine

Horváth model applied to ionic liquids

$$\frac{[A]}{k_0 - k} = \frac{1}{k_2 K_A} + \frac{[A]}{k_2} \qquad (4.20)$$



- K_A was also applied to measure the ability of the cation of an ionic liquid to block silanol sites.
- The retention is affected by the attraction of the cationic solute to the adsorbed anion on the stationary phase.
- A third contribution should be considered corresponding to the extra interaction:

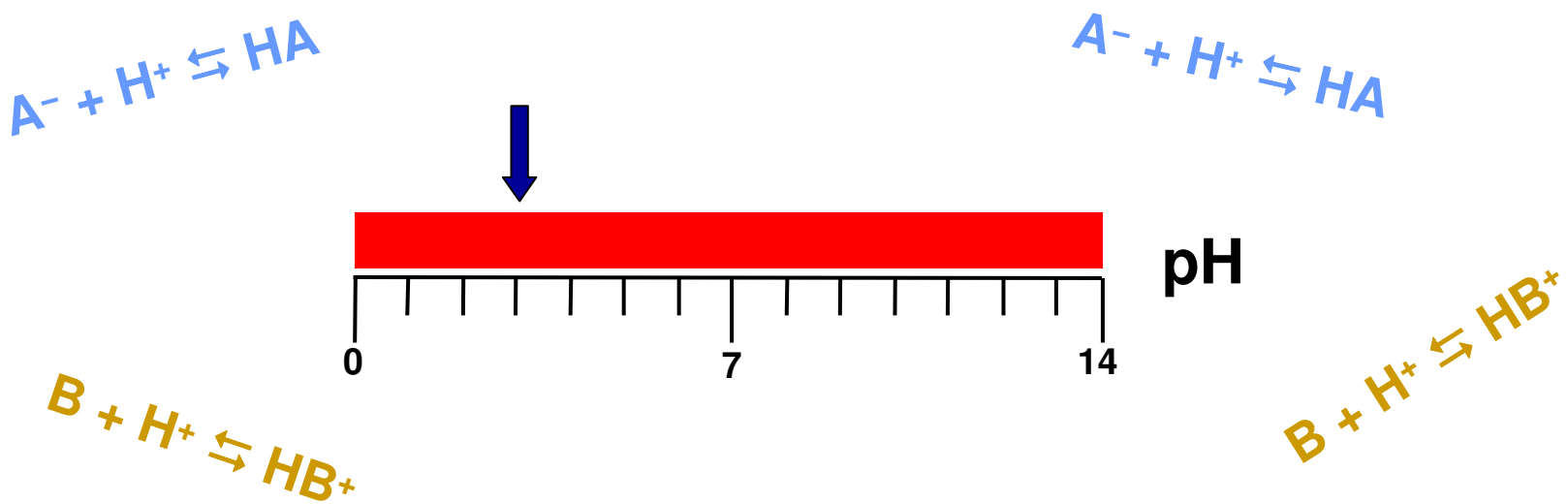
$$k_0 = k_1 + k_2 + k_3$$

The Horváth equation can only be applied to ionic liquids, where the absorption of the anion is negligible.

4.3. pH as experimental factor

4.3.1. Hydro-organic mobile phases

For ionisable solutes, pH is an additional experimental factor with a large influence on retention and selectivity. When the analysed mixture contains one or more compounds with acid-base behaviour, pH tuning can offer unique opportunities to reach resolution. However, changes in retention with pH are particularly hard to model. Not surprisingly, a widely extended practice consists of fixing the pH at a convenient value. With this practice, the benefits of this experimental factor on selectivity are lost.



RPLC retention of ionisable compounds

Weighted mean of the retention of the basic and acidic species

$$k = k_A \delta_A + k_{HA} \delta_{HA} = k_A \frac{1}{1+Kh} + k_{HA} \frac{Kh}{1+Kh} = \frac{k_A + k_{HA} Kh}{1+Kh} \qquad (4.21)$$

k : the measured retention factor

k_A : retention factor for basic species

k_{HA} : retention factor for acidic species

δ_A : molar fraction for basic species

δ_{HA} : molar fraction for acidic species

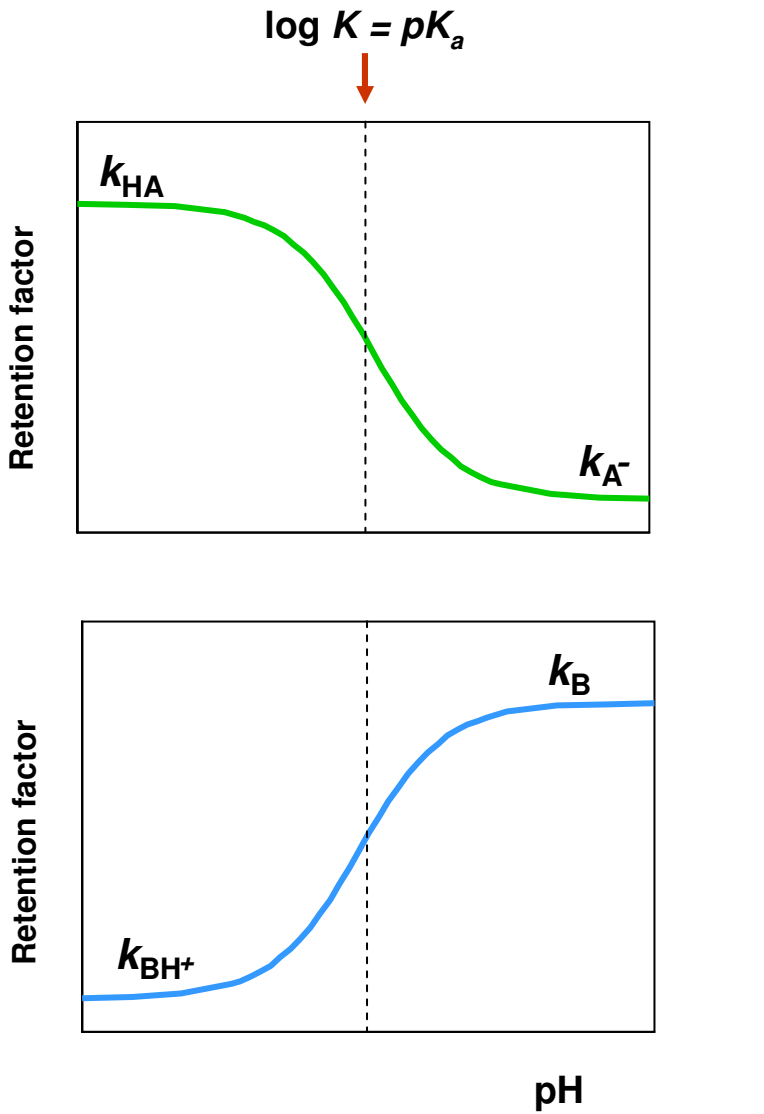
K ($= K_a^{-1}$) : apparent protonation constant

(takes into account all interactions of both acid-base species with the stationary phase and mobile phase components)

Remember that …

- Since the intrinsic retention for each acid-base species is different, a sudden change in retention will happen at pH values close to the logarithm of the apparent protonation constant (or pK_a).

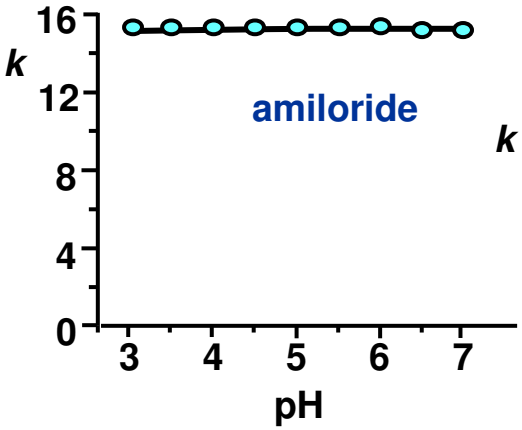
- Depending on the charge of the acidic and basic species, the retention may decrease, increase, or remain constant with pH.

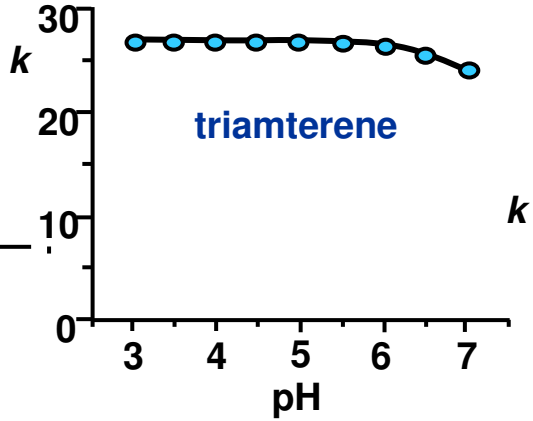


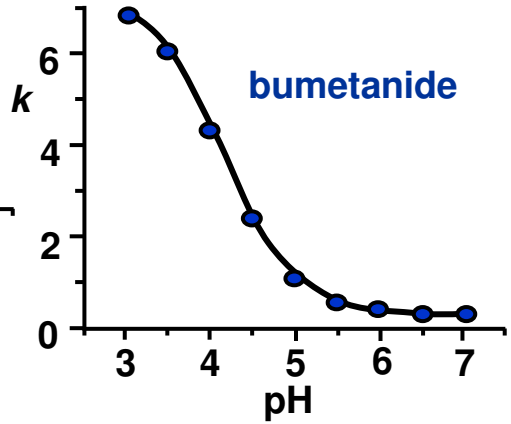
Incomplete equilibria of ionisable compounds

- The protonation process covers several pH units.
- Therefore, when the working pH range of the stationary phase is narrow (e.g. 3–7 for conventional RPLC columns), the change in retention will be only fully sampled for some compounds.
- For this reason, the retention plots of compounds with acid-base behaviour often show partial equilibria, giving rise to different patterns.

As a result of the incomplete information, fitting of the experimental data can be hard, since the retention of the acidic and basic species should be extrapolated.



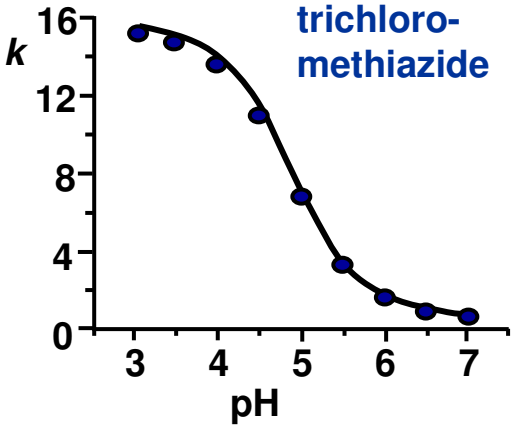




Diuretics

0.05 M sodium dodecyl sulphate

4% propanol



Fixed organic solvent content

$$k = k_A \delta_A + k_{HA} \delta_{HA} = k_A \frac{1}{1+Kh} + k_{HA} \frac{Kh}{1+Kh} = \frac{k_A + k_{HA} Kh}{1+Kh} \qquad (4.21)$$

Dividing by k_{HA} ($f = k_A/k_{HA}$) yields a more convenient fitting behaviour:

$$k = \frac{\frac{k_A}{k_{HA}} + Kh}{1+Kh} k_{HA} = \frac{f + Kh + fKh - fKh}{1+Kh} k_{HA} = \left(f + \frac{Kh}{1+Kh}(1-f) \right) k_{HA} \qquad (4.22)$$

$$\log k = \log k_{HA} + \log \left[f + \frac{Kh}{1+Kh}(1-f) \right] = \log k_{HA} + \log \left(f + \frac{10^{\log K} h}{1+10^{\log K} h}(1-f) \right) \qquad (4.23)$$

Simultaneous effect of organic solvent and pH

To account for the simultaneous effect of organic solvent and pH, the equations that describe the retention with regard to both factors are combined.

$$k = \frac{k_A + k_{HA}\, K h}{1 + K h}$$

$$\log k = \log k_w + S\varphi + T\varphi^2 \qquad (4.25)$$

- The change in the protonation constant with changes in the concentration of organic solvent should also be considered.

$$\log K = \log K_w + Q_1\varphi + Q_2\varphi^2 \qquad (4.26)$$

$$\log K = \log K_w + Q_1\varphi \longleftarrow \text{for simplicity}$$

k_w : retention in pure water K_w : protonation constant in pure water

- As a consequence of this relationship, the retention surfaces are twisted and shifted.

→ Both factors (organic modifier and pH) interact with each other.

This complicates the treatment notably.

$$\log k = \log k_{HA} + \log\left[f + \frac{10^{(\log K_w + m\varphi)}\, h}{1 + 10^{(\log K_w + m\varphi)}\, h}(1-f)\right] = \quad (4.24)$$

$$(\log k_{w,HA} - S\varphi) + \log\left[f + \frac{10^{(\log K_w + m\varphi)}\, h}{1 + 10^{(\log K_w + m\varphi)}\, h}(1-f)\right]$$

$$k = \frac{k_{w,A}\, 10^{(S_A\varphi + T_A\varphi^2)}}{1 + K_w\, h\, 10^{(Q_1\varphi + Q_2\varphi^2)}} + \frac{k_{w,HA}\, K_w\, h\, 10^{[(S_{HA}+Q_1)\varphi + (T_{HA}+Q_2)\varphi^2]}}{1 + K_w\, h\, 10^{(Q_1\varphi + Q_2\varphi^2)}} \quad (4.27)$$

quadratic relationships

$$k = \frac{10^{(R_A + S_A\varphi + T_A\varphi^2)} + 10^{(R_{HA} + S_{HA}\varphi + T_{HA}\varphi^2)}\, h}{1 + 10^{(Q_0 + Q_1\varphi + Q_2\varphi^2)}\, h} \quad (4.28)$$

easier fitting

***S*, *T*, Q_1, Q_2** : coefficients that describe the influence of the organic modifier on the retention factor and protonation constant

Polarity model

$$k = \frac{k_A + k_{HA}\, K h}{1 + K h}$$

$$\log k = c_0 + c_1 P_M^N$$

$$\log k = c_0 + c_1 P_M^N + \log\left(f + \frac{10^{(\log K_w + m\varphi)}\, h}{1 + 10^{(\log K_w + m\varphi)}\, h} (1 - f) \right) \qquad (4.29)$$

c_0 and c_1 correspond to the acidic species

4.3.2. Micellar mobile phases

$$k = \frac{k_A + k_{HA}\, K h}{1 + K h}$$

$$k = \frac{K_{AS} \dfrac{1}{1 + K_{AD}\,\varphi}}{1 + K_{AM} \dfrac{1 + K_{MD}\,\varphi}{1 + K_{AD}\,\varphi}\,[M]}$$

$$k = \frac{\left(K_{AS} \dfrac{1}{1 + K_{AD}\,\varphi}\right) + \left(K_{HAS} \dfrac{1}{1 + K_{HAD}\,\varphi}\right) K h}{\left(1 + K_{AM} \dfrac{1 + K_{MD}\,\varphi}{1 + K_{AD}\,\varphi}\,[M]\right) + \left(1 + K_{HAM} \dfrac{1 + K_{HMD}\,\varphi}{1 + K_{HAD}\,\varphi}\,[M]\right) K h} \qquad (4.30)$$

Polynomial models: acceptable accuracy in narrow pH ranges

$$\frac{1}{k} = a_0 + a_1[M] + a_2\varphi + a_3\,\mathrm{pH} + a_{12}[M]\varphi + a_{13}[M]\,\mathrm{pH} + a_{23}\varphi\,\mathrm{pH} + a_{123}[M]\varphi\,\mathrm{pH} + a_{33}\,\mathrm{pH}^2 \qquad (4.31)$$

4.4. Quantitative structure-retention relationships (QSRR)

The ideal optimisation is that one requiring no experiment. Accordingly, considerable effort has been made to estimate retention without any other extra knowledge that several experimental solute descriptors or solute chemical structures.

General solvation model (Abraham model)

A property for a series of solutes at fixed solvent conditions can be predicted from several solute molecular descriptors:

$$SP = eE + sS + aA + bB + vV + c \qquad (4.32)$$

SP **: solute property (log *k* or polarity parameter p_S)**

E, S, A, B, V **: solute descriptors**

e, s, a, b, v **: characterise differences between the mobile phase and stationary phase in HPLC**

E **: excess molar refraction**

S **: dipolarity/polarizability**

A **: hydrogen-bond acidity**

B **: hydrogen-bond basicity**

V **: McGowan volume**

e **: capability of the environment to interact with solute π- and *n*-electron pairs**

s **: dipolarity/polarizability**

a**: hydrogen-bond basicity**

b**: hydrogen-bond acidity**

v **: hydrophobic contribution**

polar contributions

Correlations between retention properties and octanol-water partition coefficients

$$SP = c_0 + c_1 \log P_{o/w}$$ SP = log *k or* p_s

- However, conventional RPLC columns still have silanol groups accessible to solutes and solvent, even when heavily coated and end-capped.
- Silanols are weakly acidic and may affect the retention: ionised silanols can retain protonated bases by cation-exchange, and neutral silanols can hydrogen bond proton-acceptor solutes.

→ Therefore, the conversion of log $P_{o/w}$ into log *k* or p_s values needs a hydrogen-bond acidity term or any other term accounting the hydrogen-bond donor interactions:

$$SP = c_1 + c_2 \log P_{o/w} + c_3 A \qquad (4.33)$$

- Comprehensive databases listing $P_{o/w}$ values are available. Also, $P_{o/w}$ values and hydrogen-bond acidity can be predicted from molecular structures.

Use of molecular descriptors predicted from molecular structures

- Software such as CODESSA or Dragon are able to calculate multiple molecular descriptors.
- Correlations between retention data and the molecular descriptors can be obtained.
- A step is required where the number of descriptors is reduced so that only a small number of uncorrelated descriptors is kept.

Example: in a study including a set of 233 compounds showing a wide variation of chemical characteristics, the solute polarity parameter p_s was best described by a model containing, besides log $P_{o/w}$, three more descriptors:

$$p_s = c + p_{o/w} \log P_{o/w} + p_{o/d}(\mathrm{pol/d^2}) + hl(\mathrm{HOMO-LUMO}) + hd(\mathrm{HDCA2}) \qquad (4.34)$$

$\mathrm{pol/d^2}$: related to the electrostatic nature

HOMO-LUMO : energy-related descriptor

HDCA2 : accounted for hydrogen-bond interactions

Unfortunately, current QSRR predictions are not yet sufficiently accurate for optimisation purposes. Reaching the high accuracy level needed in optimisation still requires the assistance of experimental data. However, this approach is interesting for prospective purposes, such as the selection of the separation system.