Journal of Chromatography A, 1577 (2018) 38–46

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Structure assisted impurity profiling for rapid method development in liquid chromatography

Sebastian Schmidtsdorff[a,b,*], Alexander H. Schmidt[a], Maria Kristina Parr[b]

[a] Chromicent GmbH, Johann-Hittorf-Str. 8, 12489, Berlin, Germany
[b] Freie Universität Berlin, Königin-Luise-Str. 2+4, 14195, Berlin, Germany

ARTICLE INFO

Article history:
Received 31 July 2018
Received in revised form
20 September 2018
Accepted 22 September 2018
Available online 24 September 2018

Keywords:
Distribution coefficient (logD)
Method development
Quality-by-design (QbD)
High-performance liquid chromatography (HPLC)
Chromatographic modeling
Extended knowledge space

ABSTRACT

The use of trial-and-error principles is a frequently used technique in method development. This may lead to the fact that analytical methods are used routinely without developers and users having gained extensive and well-founded knowledge about the robustness of their analytical methods and the influence of critical key parameters. This very often leads to unnecessary problems for analysts. A simple way in reverse phase chromatography to simulate the effects of pH value changes on the separation and retention of substances is the pH-dependent calculation of the logD value. With this tool, model substances were used to show how the time requirement for method screening can be considerably reduced *in silico* and, in addition, extended knowledge about the separation mechanics can be generated. Based on this knowledge, a new method for the purity analysis of carbamazepine was developed within a very short period of time, which improves the performance of the official Ph.Eur. monograph by far. Furthermore, the extremely high robustness of the new method was demonstrated. Using the logD based approach, Quality-by-Design is applied in method development and kept pace with the increasing requirements of regulatory authorities in the pharmaceutical industry.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

Some chromatographic methods commonly used in pharmaceutical analysis lack robustness, which can lead to a shift in retention times of individual substances in routine use and therefore to the loss of critical peak resolution (Rs). This may lead to the fact that already developed methods have to be discarded or a laborious and time-consuming new method development has to be performed. This occurs mainly when method development is done by the öne factor at a timeṗrinciple (OFAT), which does not collect comprehensive scientific knowledge on the behavior of substances under changing chromatographic conditions in routine use. Thus, it is not possible to estimate the effects of slightly changed conditions. In order to avoid time-consuming and cost-intensive changes to existing methods, which have to be revalidated, a profound knowledge about the influencing parameters of chromatographic separations should be collected as early as possible during the development stage using the Quality-by-Design (QbD) approach in method development, especially with regard to the life cycle of analytical methods [1]. This profound data can be used to generate a D̈esign Space[̈2] within the method performs robustly, even if key parameters are changed. Particularly in regulated analytical environments, as in the pharmaceutical industry, the QbD concept is therefore gaining attention [3–5]. This is already demanded and promoted even more with the guideline of the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) Q8 [2]. If such collected knowledge about the effects of chromatographic changes on the separation of substances is included in the analytical method validation and in the regulatory drug application, it is easier to make slight adjustments to critical parameters of such analytical methods, if necessary, as this is not classified as a change or variation and therefore does not require authorization in accordance with the international registration authorities [6,7].

One of the tools, which can be used to estimate the robustness of an existing method or for development of a new analytical method, is the prediction of retention times by calculating the distribution coefficient (logD). This indicates the logarithmic ratio of the concentrations of all species of a component in the distribution equilibrium in an organic, non-water-miscible phase and the concentrations of all species of the same component in the aqueous phase. Since this value depends on the degree of ionization of a substance and is therefore pKa dependent, the logD value is

* Corresponding author at: Chromicent GmbH, Johann-Hittorf-Str. 8, 12489 Berlin, Germany.
E-mail address: sebastian.schmidtsdorff@chromicent.de (S. Schmidtsdorff).

https://doi.org/10.1016/j.chroma.2018.09.044
0021-9673/© 2018 Elsevier B.V. All rights reserved.

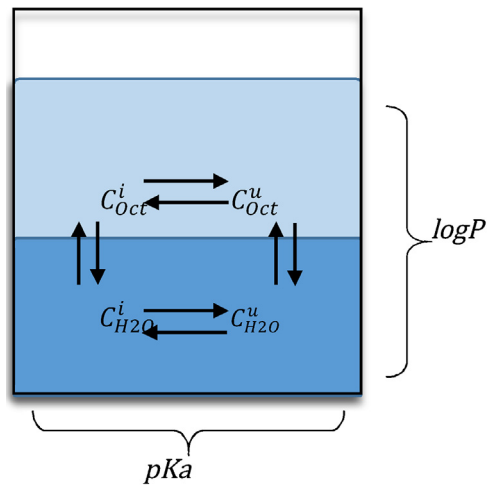


Fig. 1. Distribution of a species in a octanol-water system ("i" stands for ionized and "u" for unionized), logP only express the distribution of the unionized substance.

always given for a certain pH or as a function of pH. LogD values are determined as the distribution coefficient for an octanol/water system and can be set equal to the partition coefficient (logP) for non-ionizable substances (see also Fig. 1). The logD can also be calculated by using structure based mathematical equations [8–11] or determined experimentally [12–14].

For adsorption and desorption effects on a reversed phase column in high-performance liquid chromatography (HPLC), this model maybe used after minor modification, since it is also a system that has a non-polar (stationary) phase, which is immiscible with the hydrophilic, polar (mobile) phase with analytes distributed between them. The extent of this distribution depends on the degree of ionization and thus on the pH value of the mobile phase and the general lipophilicity (expressed by the logP) of the individual substance, which are both key parameters for the logD (see Fig. 1 for more details).

The aim of this work is to show that a prediction of retention time shifts can be achieved by the pH dependent calculation of logD values. This will be demonstrated by the purity test of the official pharmacopoeial monograph of carbamazepine and its known impurities (Ph. Eur. 9.0) and integrated in a new method development to illustrate how important such screening knowledge is for method development and risk-assessment of established analytical methods. On this basis, it will be shown that protracted screening experiments can be avoided if structure-based data are used. Based on the generated knowledge from the screening phase, the method is then further optimized and developed in a broad-based method optimization experiment according to the QbD concept so that a fully visual "Design Space" can be generated. In a final step, this is subjected to a robustness test in order to prove and demonstrate the reliability of the final new method, which is based on the process of screening, optimization and robustness testing.

A great work has been done in the past to predict structure-based retention times based on their physicochemical properties of the analytes without having to carry out HPLC experiments. Examples for these techniques were called quantitative structure-retention relationships (*QSRR*) [15–19]. However, this never led to a groundbreaking success that could replace the screening phase, which shows that a whole series of other factors, which cannot all be included *in silico*, influence the separation, which has made screening and optimization runs indispensable up until now. For this reason, we deliberately decided against this approach in our work and did not use structure-based screening to completely replace practical chromatographic experiments.

However, trial-and-error approach according to the OFAT principle is still widespread, time-consuming and ignores the advancement of science and technology. Therefore, with this work we would like to contribute and provide an approach to save time and money that can be easily implemented and could be a helpful starting point for a better method development and understanding.

2. Materials and methods

2.1. Chemicals and reagents

Reference substances of the analytes Carbamazepine (EDQM reference standard), Carbamazepine impurity A (10,11-dihydrocarbamazepine), Carbamazepine impurity D (iminostilbene), Carbamazepine impurity E (iminodibenzyl) and Carbamazepine impurity F (iminostilbene N-carbonyl chloride) were supplied by Sigma–Aldrich Chemie GmbH (Taufkirchen, Germany); Carbamazepine impurity B (9-methylacridine) was purchased from Chemos (Regenstauf, Germany); Carbamazepine impurity C (N-carbamoyl carbamazepine) and Carbamazepine impurity G (10-bromocarbamazepine) were sponsored by LGC GmbH (Luckenwalde, Germany).

Solvents and reagents were at least in analytical grade and purchased from VWR (Darmstadt, Germany). Fresh ultrapure water was obtained from a Milli-Q Gradient A10 equipped with a 0.22 μm Millipak Durapore membrane cartridge (Merck KGaA, Darmstadt, Germany).

2.2. Instrumentation

Chromatographic analysis was performed using a Waters Alliance HPLC system (Eschborn, Germany) equipped with a Waters 2695 separation module with degasser, temperature-controlled sample compartment and column heater and a Waters 2996 photodiode-array detector. For system control, data acquisition and data processing the Empower 3 software (Waters GmbH, Eschborn, Germany) was used.

The pH values of the mobile phases were measured on a Mettler MP225 pH-Meter (Mettler-Toledo GmbH, Gießen, Germany).

Instrumentation was operated fully qualified according to the 4Q model of the USP <1058> [20].

2.3. Chromatographic conditions

The HPLC was equipped with a Nucleosil 100-10 CN (dimensions 250 × 4.6 mm, particle size 10 μm – Macherey Nagel GmbH & Co. KG, Düren, Germany), which is explicitly mentioned in the official commentary to the Ph.Eur. monograph for carbamazepine [21].

As mobile phase 0.2 ml of anhydrous formic acid and 0.5 ml of triethylamine were added to 1000 ml of a solution of tetrahydrofuran, methanol, water (3:12:85, V/V/V) and used in isocratic elution. The flow rate was 2.0 ml/min with a run time of 8 times the retention time (t_R) of carbamazepine (t_R about 10 min). The injection volume was 20 μl and the detection wavelength 230 nm as mentioned in the monograph. The pH value of the mobile phase as mentioned in the official monograph was measured as pH = 4.0 at the aqueous content, after addition of formic acid and triethylamine; methanol and tetrahydrofuran were added afterwards [22].

For the subsequent method development a XSelect CSH C18 column (150 x 4.6 mm, 3.5 μm – Waters) was used, due to its high resistant for severe pH values and temperatures.

2.4. Screening procedure

The pH dependent logD values for carbamazepine and its seven impurites (illustrated in Fig. 2) were determined from pH 0 to 14 by

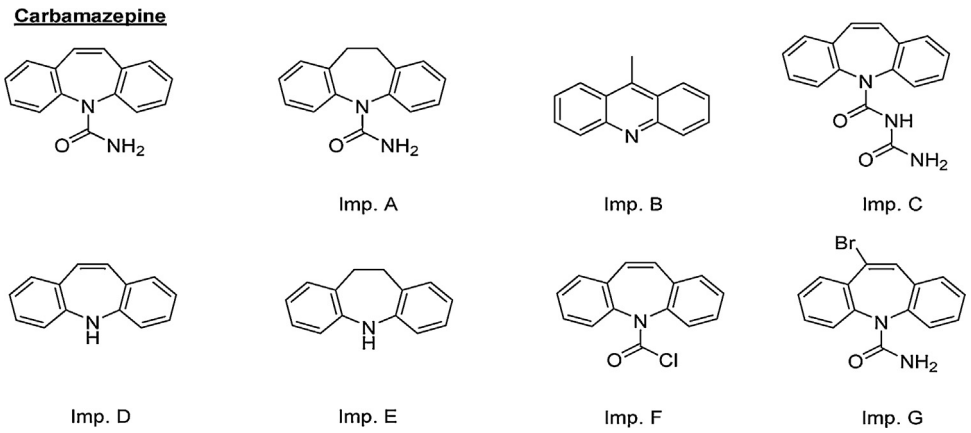


Fig. 2. Structure of carbamazepine and its impurities A–G.

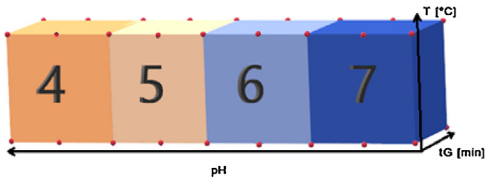


Fig. 3. Three-dimensional screening model with temperature (T, from 40 to 60 °C), gradient time (tG, from 10% to 90%B in 10 to 20 min) and pH (from 4.0 to 7.0 ± 0.5) for HPLC method development for extended knowledge space generation.

the calculated logP and pKa values [23] according to the following calculation formulas [24]:

Since the concentrations of ionized species in the octanol phase, graphically displayed in Fig. 1, are negligibly low, the value $C^{ionized}_{Octanol}$ (see equation 1.1) can be set as zero. The calculation can therefore be simplified, which results in two different equations for the calculation of logD values (see equation 1.2 and 1.3).

2.5. Optimization and robustness testing procedure

Based on the screening phase, the next step in method development is the optimization, based on the previous knowledge. The goal of this optimization step is defined by the analytical target profile (ATP) [25]. In our model system the ATP was to develop a new method, which is significantly faster (run time ≤ 30 min), has a better separation performance (Rs ≥ 2.0) and a high, visually-displayed chromatographic robustness. An additional request is to transfer and develop the method on a C18 column.

To avoid time-consuming OFAT development, the logD knowledge gathered so far was implemented in a multifactorial optimization model, to abbreviate experimental time and to generate a 3D "Design Space" for all parameters (pH of mobile phase, temperature, gradient time). For this purpose DryLab®4 chromatography modeling software package (Molnar-Institute, Berlin, Germany) was used to optimize the separation of the mixture of carbamazepine and its seven impurities for the newly developed method.

A generic, linear gradient with a flow rate of 1.2 ml/min and a fast and a slow gradient time (tG) of 10 min and 20 min with a linear increase from 10%B to 90%B with a subsequent equilibration step to the start conditions for 10 min was created on a C18 column. The aqueous eluents (A) were adjusted from 3.5 to 7.5 in 0.5 pH steps with 10% phosphoric acid or 4 mM K_2HPO_4 solution, the non-aqueous eluent was acetonitrile (B). Temperature (T) was tested at 40 °C and 60 °C. An illustration is available in Fig. 3. Due to the fact that the logD chart displays no significant change below pH 3.5 and above pH 7.5, which should also apply to the retention times, testing was limited to these corner points.

The resulting four cubes with pH 4, 5, 6 and 7, each ± 0.5 were built and put together from these 4 × 12 basic runs (pH, tG, T). Thus a sufficient resolution (Rs ≥ 2.0) can be graphically mapped in a wide pH range from 3.5 to 7.5 – termed "extended knowledge space" as described by Rácz and Kormány [26]. Bigger steps in pH than ± 0.5 are not recommended, due to the fact that there is no linear relationship between pH and retention time. After plotting each cube, the most robust area within the four cubes was chosen and the method was optimized and robustness testing was carried out with DryLab®4.

3. Results and discussion

3.1. Screening: relationship between logD and retention time in HPLC

In order to show that method development sreening can be performed with logD calculations, we calculated a logD chart (Fig. 4) for carbamazepine and its seven impurities. As shown, they have very widely dispersed logD values. All values are positive, so that all substances are rather lipophilic. At the working point of the monograph (pH = 4.0) there is a difference of three between the most hydrophilic substance (smallest logD) and the most lipophilic substance (highest logD), which indicates that their lipophilicity differs by factor 1000. This explains the very long run time of the official isocratic HPLC method (about 100 to 120 min) in the Ph.Eur. monograph.

It can be clearly observed in the logD chart that within the stable region of the cyanopropyl column (pH 2 - 8), only impurity B undergoes a significant change in its distribution coefficient and is highly sensitive to minimal pH deviations, due to its basic pyridine-analogous structural element (see Fig. 2). The calculated pKa value of 6.8 for impurity B (9-methylacridine) seems reasonable, since *p*-methyl substituted pyridine analogues show an increase in pKs value of about 0.5–1.0 [27] and acridine has an experimentally determined pKs value of about 5.6 [28,29]. All other substances do not show any changes in their logD in the relevant pH range (pH 2 – 8). At higher pH values, impurity B tends to become as lipophilic as the other substances.

To study the effect of a pH change on the retention time and to observe whether these shifts follow a pattern, correlating to the logD diagram, the pH was gradually increased to pH = 5.0 by increasing the amount of triethylamine in the eluent. It is evident that impurity B is retarded to a greater extent, as it is increasingly less protonated and thus becomes more lipophilic. From pH = 4.5

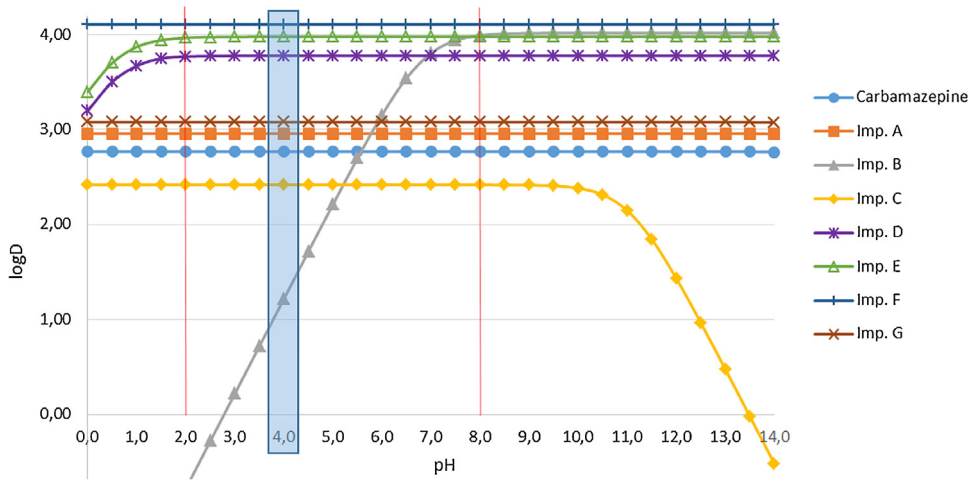


Fig. 4. logD chart for Carbamazepine and its impurities A – G with column stability range (red lines) and the measured pH value of the eluent in the official monograph (blue marked area) – calculated according to equations.1.1–1.3.

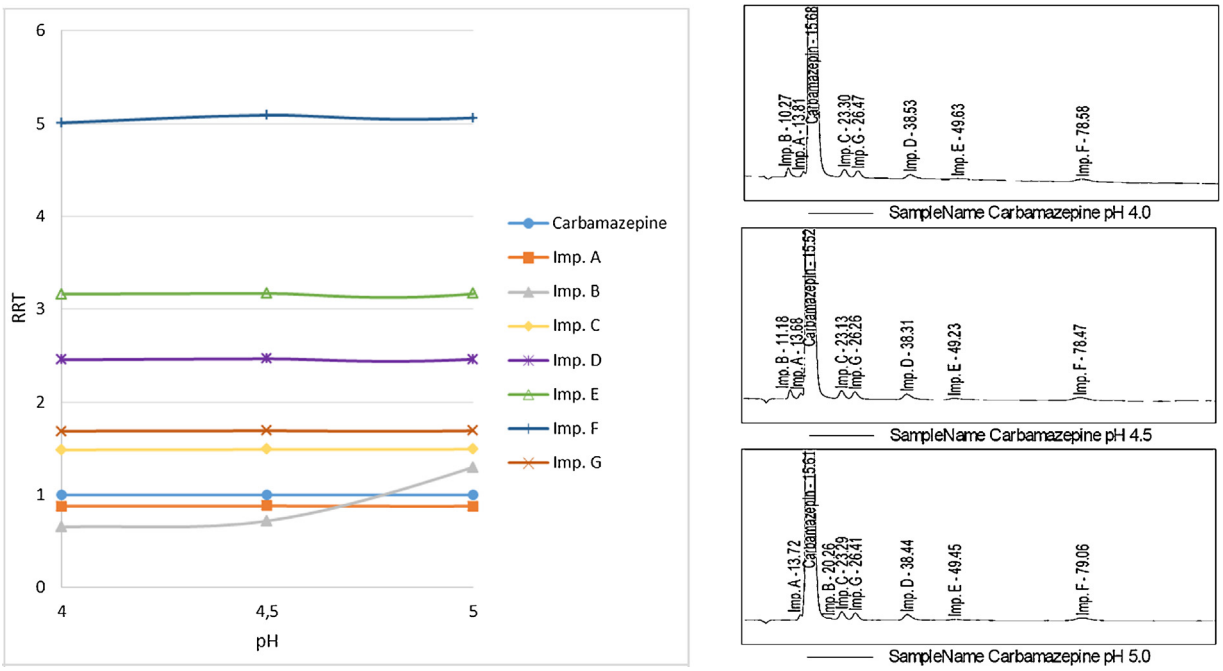


Fig. 5. Relative retention (RRT) chart and corresponding chromatograms from pH 4.0 to 5.0 adjusted with the agents of the mobile phase from the official monograph of carbamazepine [22]. The top chromatogram shows the impurity profile with retention times according to the original testing procedure of the Ph.Eur.

upwards this results in a coelution with impurity A as displayed in Fig. 5. What also can be seen is that impurity A and carbamazepine have a very poor resolution – similar to the logD chart.

The HPLC method in the monograph requires an additional system suitability test (SST), where the resolution between carbamazepine and impurity A must be at least 1.7, but the reference solution to be tested is diluted 1000 times more than the test solution for purity analysis [22]. Therefore, despite a Rs $\geq$ 1.7, partial co-elution may occur, which would lead to an incorrect overestimation of impurity A and consequently to a result outside the pharmacopoeial specification. Additionally, as column may alter during shelf life, the separation performance will decrease and therefore, despite compliance with SST, an even stronger coelution in the chromatogram of the test solution will occur, which is unacceptable.

In a subsequent study the pH of the mobile phase was increased above 5.0 und decreased below 4.0 to investigate the complete peak movement of impurity B. Therefore the buffer from the official monograph was replaced by suitable buffers, to study peak movement over a broad pH range. For this purpose, the formic acid/triethylamine buffer was substituted by a 15 mM formiate (buffer range: 3.8 ± 1.2), acetate (buffer range: 4.8 ± 1.2) and phosphate buffer (buffer range: 7.2 ± 1.2), so that a pH range of 3.0 to 7.5 can be covered. Chromatograms at pH levels more acidic than pH = 3.0 and more alkaline than pH = 7.5 were not obtained in order not to damage the stability of the cyanopropyl column.

Fig. 6 shows the relationship between the retention time (expressed as the relative retention time (RRT) for each substance with reference to carbamazepine) and the pH value of the mobile phase. Compared with the corresponding chart of the logD values (Fig. 4), an almost congruent image with a clear correlation between the RRT and the logD is obtained (Fig. 7).

As shown by the results from the logD screening, impurity B is very susceptible to retention time shifts. However, below a pH value

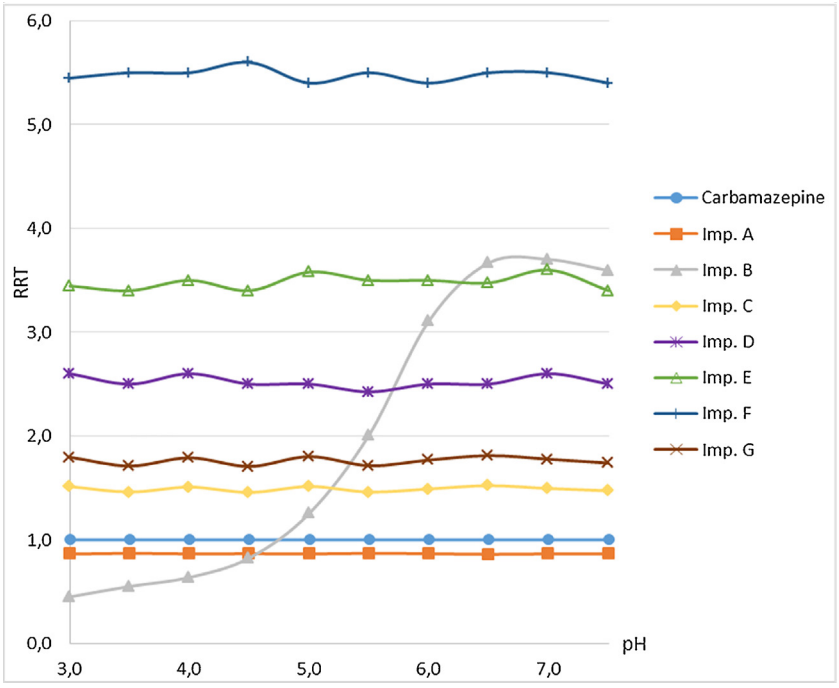


Fig. 6. RRT chart for carbamazepine and related compounds with alternative buffers compared to the logD plot.

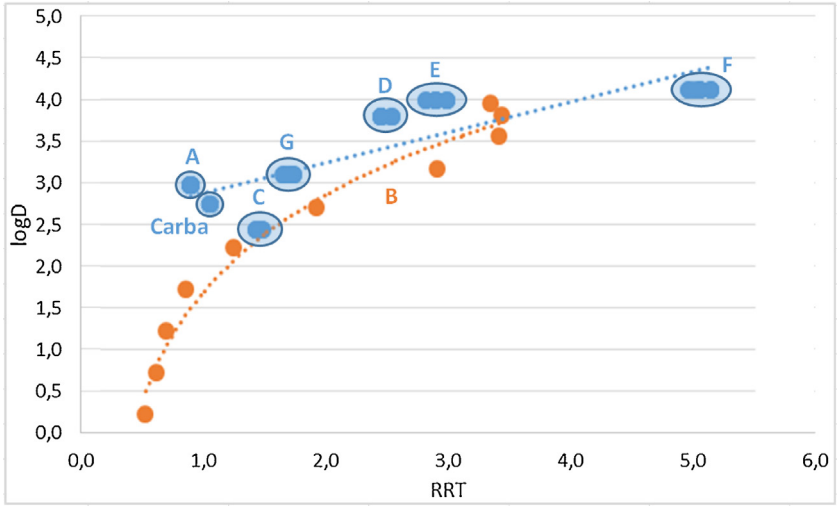


Fig. 7. Relationship between RRT and logD of carbamazepine and its impurities A–G for pH 3.0–7.5 on cyanopropyl column.

of 4.5 and above 6.5, no significant peak movement was observed. In the area in between, a critical coelution of impurity B with the early eluting substances (carbamazepine, impurity A, C, G, D and E) appears. This experimental knowledge is congruent to the data from the logD chart, which indicates that *in silico* screening by logD calculation is a viable way to accelerate and simplify screening.

3.2. Optimization: software assisted QbD method development and construction of "Design Space"

The knowledge from the screening phase demonstrated that critical pH ranges for the separation of acid/base-sensitive substances can be predicted using logD calculations. On this basis and with the pre-defined ATP a new gradient method was developed in a three-dimensional approach [30] using a conventional reversed phase column (C18) with a significantly faster ($t_{total} \leq 30$ min) and more effective separation performance (Rs $\geq$ 2.0), as compared to the cyanopropyl column and the isocratic method from the pharmacopoeial monograph. A gradient method is considered to be reasonable, since logD values (see Fig. 4) are distributed over a wide range. Isocratic elution leads to extremely long analysis runs and a strong peak broadening for the late-eluting compounds, which significantly impairs their detection and quantification limit. As an example, the chromatogram obtained applying the original pharmacopoeial method (Fig. 9) shows a baseline peak width of over 6 min for impurity F. In general, baseline separation is achieved at Rs $\geq$ 1.5, but only if the peaks are ideally symmetric and at the same height, which is not the case with the purity monograph of carbamazepine. Therefore, we choose a higher Rs value for the ATP. A C18 column was chosen to replace the column from the monograph, because cyano columns are known, to be chemically and long-term instable, due to their short ligands [31,32]. This effect can also be observed in Fig. 6. The integer pH values were screened first (from acidic to basic) and then the intermediate values in increments of

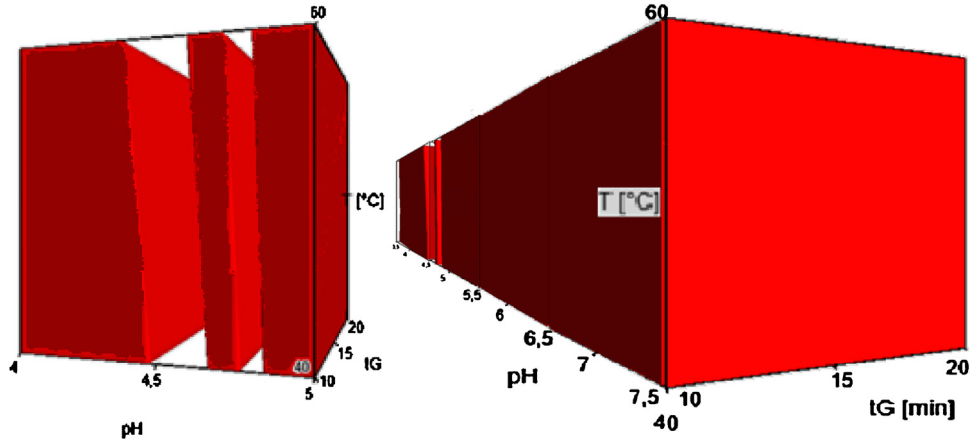
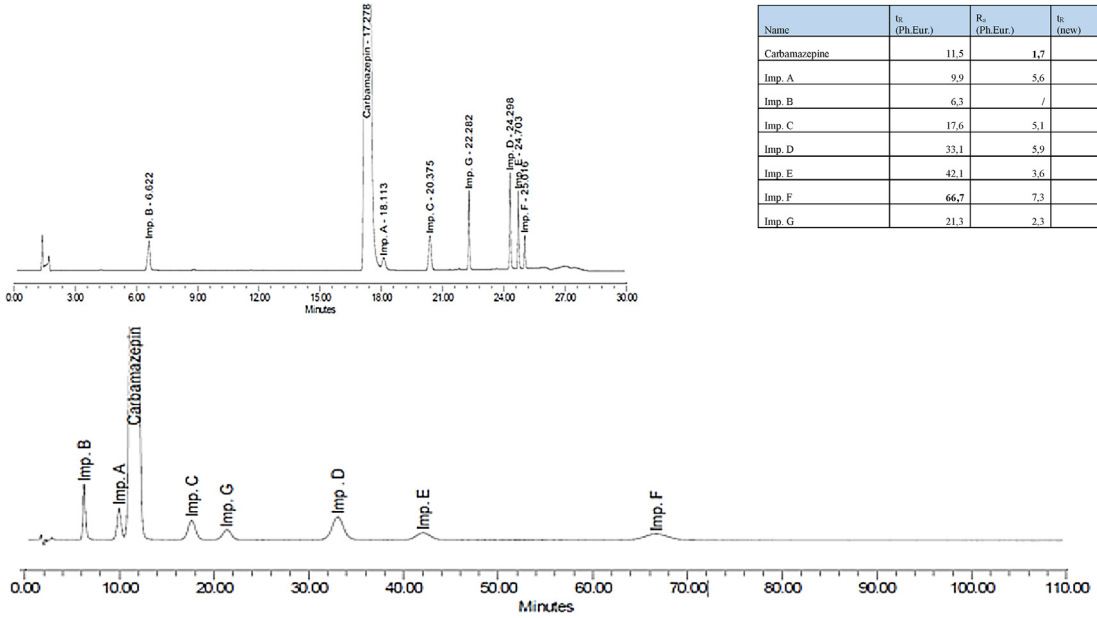


Fig. 8. Four DryLab® resolution cubes (red equals Rs ≥ 2) concatenated to an extended knowledge space (right) with a section in the critical range from pH 4.0 to 5.0 (left). Note that impurity A with a poor resolution to carbamazepine is not displayed here and optimized in a later step.



Name	t_R (Ph.Eur.)	R_s (Ph.Eur.)	t_R (new)	R_s (new)
Carbamazepine	11,5	**1,7**	17,3	45,8
Imp. A	9,9	5,6	18,1	**2,9**
Imp. B	6,3	/	6,6	/
Imp. C	17,6	5,1	20,4	8,1
Imp. D	33,1	5,9	24,3	18,3
Imp. E	42,1	3,6	24,7	3,6
Imp. F	**66,7**	7,3	**25**	2,9
Imp. G	21,3	2,3	22,3	11,4

Fig. 9. Comparison of the new developed carbamazepine method (above) with the original chromatogram obtained from the Ph.Eur. monograph (below) with the corresponding retention times and the resolution of each substance.

0.5. For the non-integer pH values, retention is already significantly shorter, which is caused by the instability of the column.

Since especially for reversed phase chromatography it is well understood how different parameters relatively affect the behavior of retention times, it is sufficient to measure predetermined chromatographic parameters (we have chosen temperature, pH value and tG for our experiments). These measurements may be used in a three-dimensional model to predict very precisely how changes in these parameters within this model affect the separation. Thereby a calculation of predicted chromatograms within this cube is very precisely possible. This makes the method optimization much easier, since only a few corner point experiments have to be carried out. Nevertheless, with the underlying physical-chromatographic theories, a complete knowledge space about the peak movement within the screened edges of the 3D model, also called ̈cubes, can be generated. In contrast to this, hundreds of experiments according to the OFAT principle would be necessary.

The software DryLab®4 implements this data in a resolution map (cube), by calculating the resolution of all peaks for each of the by far over a thousand possible experiments within the 3D model of temperature, pH and tG. It then displays the most critical (smallest) peak resolution in different colors, so that it is visually evident under which conditions all peaks are completely separated or not. The ICH Q8(R2) calls this the "Design Space", as a "multidimensional combination and interaction of input variables [. . .] and process parameters that have been demonstrated to provide assurance of quality" [2].

The experimental data from the screening phase have already shown that method optimization works best in acidic or slightly basic eluents, since small changes in the pH value have no significant influence on the resolution. Nevertheless, we have decided not only to measure in a very narrow and most suitable pH window, but to measure the full pH range. This data was then used to generate an extended knowledge space that goes far beyond the operating working point and can demonstrate again the congruence to the logD chart in 3D model on the C18 column. Fig. 8 illustrates the results for the temperature-pH-tG screening, where all red areas have a Rs of at least 2.0 for all peaks. The two notches at around pH 4.5–4.7 and 4.8–4.9 mark the pH region, where impurity B is coeluting with other impurities and carbamazepine.

T	57	±	4	[°C]
pH	3.8	±	0.2	
Flow Rate	1.2	±	0.1	[ml/min]
Start %B	15.0	±	2	%B
Step 1 Time	18	±	0.25	[min]
Step 1 %B	27.0	±	2	%B
tG	24	±	0.25	[min]
End %B	80.0	±	2	%B
Dwell Volume	1.10	±	1.10	[ml]

Required Resolution	2.0
Successful Experiments	19683
Success Rate	100%
No. of Factors	9
No. of Levels	3
No. of Experiments	19683

Fig. 10. Method robustness calculation parameters.

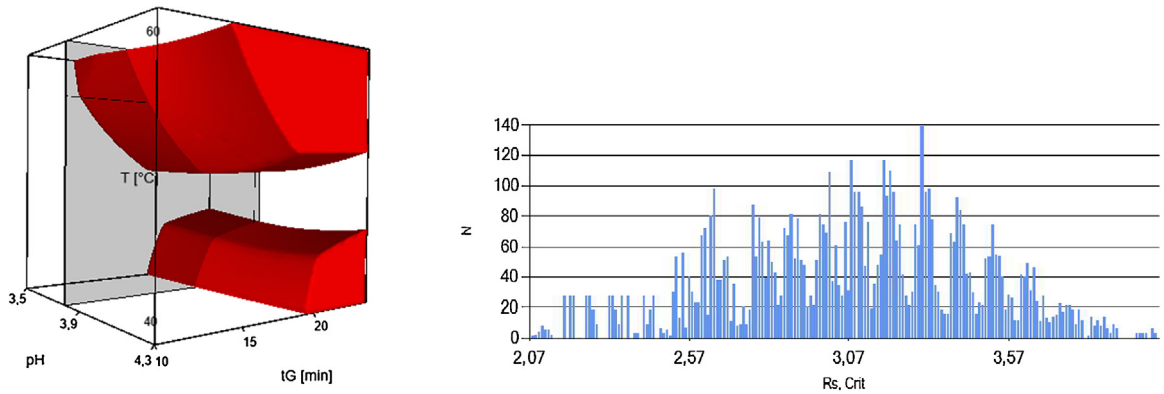


Fig. 11. Final "Design space" for step gradient method around the working point with Rs ≥ 2.0 and robustness test displayed by frequency distribution of smallest peak resolution (critical resolution) per experiment according to Fig. 10 – note that all in silico experiments succeeded the ATP specifications.

The results show that temperature and tG have almost no effect on the separation and only shorten the analysis time. The resolution cubes additionally show a much smaller critical pH range than it would have been expected from the logD calculation. In addition, this critical range starts at lower absolute pH values (in the more acidic environment), what could be explained by the fact, that organic compounds in the mobile phase greatly decrease the hydrogen ion activity and therefore increase the pH for acids [33]. Therefore, an acidic eluent becomes more neutral by addition of acetonitrile or methanol. However, the results are congruent with the pH screening on the cyanopropyl column, which suggests that the real pH value on the column is higher due to the increased temperatures and the addition of organic coeluents.

As a basis for further method development a pH of 3.8 and a column temperature of 57 °C were chosen. Based on the extended knowledge space, displayed by Fig. 8, higher pH values seem to be also suitable for a robust separation. But an advantage that can be derived from the logD screening is that in neutral to basic pH ranges impuritiy B elutes much later from the column, which would lead to an elution together with the impurities E–G. As a result, the gradient slope (tG) of the new method would have to be significantly slowed down in order to avoid coelution. This in turn would have extended the runtime, since the logD difference of these mentioned impurities is minimal in these pH regions. Therefore, a decision in advance towards significantly more acidic pH values of the eluent can be drawn, in order to save experimental screening time.

In a final step the separation of carbamazepine and impurity A (10-11-dihydrocarbamazepine) had to be optimized, due to their coelution in the previous experiments, which can be explained by their almost identical logD values. For this purpose we used DryLab®4 and the knowledge gathered so far and implemented a gradient step *in silico* in the "Design Space" with a lower slope at the beginning and increased the slope after the elution of impurity C. The final verification run confirmed the successful adaptation of the method; all substances are separated (see Fig. 9 for the obtained chromatogram in comparison to the original pharmacopoeial method) and the correlation between the *in silico* predicted and the experimental verified retention times has a coefficient of correlation of 0.99996.

The final gradient method, which completely fulfills the ATP, is: HPLC-grade water (adjusted to pH 3.8 with 10% phosphoric acid) as eluent A and HPLC-grade acetonitrile as eluent B, starting at 15% B and slowly increased linearly to 27% B within 18 min, followed by a rapid increase to 80% B in further 6 min and a reequilibration to 15% B for 6 min, resulting in a total run time of 30 min on a XSelect CSH C18 (150 x 4.6 mm; 3.5 μm) at 57 °C column temperature and a flow rate of 1.2 ml/min. The detection wavelength was maintained at 230 nm.

3.3. Robustness testing

In the two previous chapters it was shown that the method screening according to the logD approach and the subsequent optimization according to the QbD concept was successful. Now the resulting method will be tested for its robustness in order to confirm that even under minor parameter changes the analytical method is still working robust in order to have a high degree of confidence in routine use.

Since it has already been shown that there is a very high degree of correlation between the actual experimentally determined retention times and the *in silico* model, the model was also used to verify robustness [5]. As shown in Fig. 10, all essential factors were changed multifactorially (T, tG for each gradient step, pH, %B for each gradient step, Flow Rate, Dwell Volume) and for each of these theoretical working points the critical resolution was calculated (peak pair with smallest Rs value). For 9 factors with 3 levels, this results in 19.683 *in silico* experiments with a 100% success rate of Rs > 2.0. Fig. 11 shows the new calculated "Design Space" in red around the working point and the plot of the robustness testing,

indicating that the method works highly robust and fulfilling the ATP.

It would also have been possible to select a working point at a lower column temperature (second red design space within the cube), but the model clearly shows that the desired resolution of all peaks can only be achieved to a much smaller degree. A robustness test with the same parameters as described would not have been possible with a success rate of 100%. This illustrates that a method development cannot end when a complete separation has been achieved, but must also pass a robustness test so that it can be implemented in routine use, on other devices and in other laboratories.

Another advantage of the model is that the robustness tests can be used to determine which peak pair is the most critical among the changing chromatographic parameters. This can be used to develop a suitable control strategy (e.g by defining a SST) [5,34] as defined by the ICH Q8: "A planned set of controls, derived from current product and process understanding that ensures process performance and product quality" [2]. This control strategy can also be used especially with regard to the life cycle of the analytical method to monitor column and to monitor column and equipment suitability permanently and intervene in time before critical method parameters (e.g. peak resolution) are no longer achieved [1].

4. Conclusion

The work shows that the calculation of pH dependent logD values and their use for method development is a simple, complimentary and quick help in the early development stage of analytical methods and their robustness screening. It can give an approximate picture of the influence of the pH on the separation performance and is a good starting point on which structure-based work can be carried out using a rational method development strategy. Baczek, et al. [35] have demonstrated that pure QSRR based prediction of retention time is not accurate but provides information about the general separation mechanics and becomes more precise by the use of chromatographic modeling softwares. The ICH Guideline Q8(R2) describes this process as "a systematic approach to development that begins with predefined objectives and emphasizes product and process understanding and process control, based on sound science and quality risk management" - also known as QbD [2]. It was demonstrated that a set of 36 basic runs is sufficient to generate a complete "Design Space" over a wide pH range, where RRT were congruent with our logD calculations, so that an extended knowledge space about critical regions and factors for separation could be obtained quickly. This knowledge was optimized *in silico* to separate all substances according to the ATP specifications. Thus only a verification run with the final conditions had to be carried out to show performance and robustness. With the help of the logD calculation, the development work would have been significantly accelerated, since a screening over the entire pH range could have been early recognized as unnecessary. Our study indicates that scientific, structure-based preliminary work should be used for method development, since it is an easy and timesaving way with QbD principles. This can only be one possible approach among many others and the significance should not be overestimated in preliminary stages without conclusive, verifiable experiments. In addition, it is not always possible to include all substances in the method development, since some impurities are structurally not identified and a prediction in this case is not possible. Likewise the logD approach probably only works very limitedly for columns with a deviating interaction mechanism, which do not correspond to a classical reversed phase columns (e.g. amide, diol, HILIC, ion exchange). However, it is a good tool if substances cannot be purchased and can only be generated from stress tests, since it is possible to visualize possible effects of pH changes on retention for known impurities, especially, when generated impurities are not stable over a long period of time.

In our case study, we were able to show that the use of scientific tools in advance may considerably shorten development work, generate a deep understanding of the effects of individual critical parameters and thus achieve an improved method. The new method is almost four times more time-efficient (110 min vs. 30 min), solvent-saving, less harmful to health in handling (THF and methanol have been removed) and significantly more efficient than the Ph.Eur. monograph of carbamazepine (both in terms of money and separating performance).

Additionally we were able to generate an extended knowledge space over the full range of pH, temperature and tG, which goes far beyond the selected working point of the method and can represent a fundamental knowledge of performance. With this knowledge, a fully visual image of the critical influencing factors for the separation of all substances can be mapped and a suitable control strategy for routine application was derived. All this was only possible because knowledge enables an integrated risk assessment (risks are identified, understood and quantified), which can be used in the later life cycle of an analytical method to evaluate the effectiveness of control strategies and, if necessary, to implement improvements and thereby further minimize potential risks.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

LGC GmbH is acknowledged for sponsoring some of the used reference standards. Furthermore, we would like to thank Macherey-Nagel GmbH & Co. KG for providing the column.

References

[1] M.K. Parr, A.H. Schmidt, Life cycle management of analytical methods, J. Pharm. Biomed. Anal. 147 (2018) 506–517, http://dx.doi.org/10.1016/j.jpba.2017.06.020.
[2] International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, ICH Harmonized Tripartite Guideline Q8(R2) Pharmaceutical Development, 2009.
[3] A.H. Schmidt, C. Wess, A Quality-by-Design with Design-of-Experiments approach for development of an UPLC purity method for carbamazepine, J. Liq. Chromatogr. Rel. Technol. 37 (2014) 2653–2666.
[4] A.H. Schmidt, M. Stanic, I. Molnar, In silico robustness testing of a compendial HPLC purity method by using of a multidimensional design space build by chromatography modeling-case study pramipexole, J. Pharm. Biomed. Anal. 91 (2014) 97–107, http://dx.doi.org/10.1016/j.jpba.2013.12.023.
[5] A.H. Schmidt, I. Molnar, Using an innovative Quality-by-Design approach for development of a stability indicating UHPLC method for ebastine in the API and pharmaceutical formulations, J. Pharm. Biomed. Anal. 78-79 (2013) 65–74, http://dx.doi.org/10.1016/j.jpba.2013.01.032.
[6] FDA Guidance for Industry, Changes to an Approved NDA or ANDA, 2004.
[7] Commission regulation (EC) No 1234/2008, in: Concerning the Examination of Variations to the Terms of Marketing Authorisations for Medicinal Products for Human Use and Veterinary Medicinal Products, 2008.
[8] A. Tsantili-Kakoulidou, I. Panderi, F. Csizmadia, F. Darvas, Prediction of distribution coefficient from structure. 2. Validation of Prolog D, an expert system, J. Pharm. Sci. 86 (1997) 1173–1179, http://dx.doi.org/10.1021/js9601804.
[9] L. Xing, R.C. Glen, Novel methods for the prediction of logP, pK(a), and logD, J. Chem. Inf. Comput. Sci. 42 (2002) 796–805.
[10] V.N. Viswanadhan, A.K. Ghose, G.R. Revankar, R.K. Robins, Atomic physicochemical parameters for three dimensional structure directed quantitative structure-activity relationships. 4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally occurring nucleoside antibiotics, J. Chem. Inf. Model. 29 (1989) 163–172, http://dx.doi.org/10.1021/ci00063a006.
[11] G. Klopman, J.-Y. Li, S. Wang, M. Dimayuga, Computer automated log P calculations based on an extended group contribution approach, J. Chem. Inf. Model. 34 (1994) 752–781, http://dx.doi.org/10.1021/ci00020a009.

[12] N. Gulyaeva, A. Zaslavsky, P. Lechner, et al., Relative hydrophobicity and lipophilicity of drugs measured by aqueous two-phase partitioning, octanol-buffer partitioning and HPLC. A simple model for predicting blood-brain distribution, Eur. J. Med. Chem. 38 (2003) 391–396.
[13] Z. Chen, S.G. Weber, High-throughput method for lipophilicity measurement, Anal. Chem. 79 (2007) 1043–1049, http://dx.doi.org/10.1021/ac061649a.
[14] P.C. Chiang, Y. Hu, Simultaneous determination of LogD, LogP, and pK(a) of drugs by using a reverse phase HPLC coupled with a 96-well plate auto injector, Comb. Chem. High Throughput Screen. 12 (2009) 250–257.
[15] R. Kaliszan, Correlation between the retention indices and the connectivity indices of alcohols and methyl esters with complex cyclic structure, Chromatographia 10 (1977) 529–531, http://dx.doi.org/10.1007/bf02262911.
[16] T. Bączek, R. Kaliszan, Predictive approaches to gradient retention based on analyte structural descriptors from calculation chemistry, J. Chromatogr. A 987 (2003) 29–37, http://dx.doi.org/10.1016/s0021-9673(02)01701-6.
[17] T.H. Miller, A. Musenga, D.A. Cowan, L.P. Barron, Prediction of chromatographic retention time in high-resolution anti-doping screening data using artificial neural networks, Anal. Chem. 85 (2013) 10330–10337, http://dx.doi.org/10.1021/ac4024878.
[18] S. Schefzick, C. Kibbey, M.P. Bradley, Prediction of HPLC conditions using QSPR techniques: an effective tool to improve combinatorial library design, J. Comb. Chem. 6 (2004) 916–927, http://dx.doi.org/10.1021/cc049914y.
[19] E. Tyteca, M. Talebi, R. Amos, et al., Towards a chromatographic similarity index to establish localized quantitative structure-retention models for retention prediction: Use of retention factor ratio, J. Chromatogr. A 1486 (2017) 50–58, http://dx.doi.org/10.1016/j.chroma.2016.09.062.
[20] United States Pharmacopeia and National Formulary (USP 40-NF 35), Vol 2, United States Pharmacopeial Convention, Rockville, MD, 2017.
[21] Carbamazepin, Arzneibuchkommentar - Wissenschaftliche Erläuterungen zum Arzneibuch, in: Wissenschaftliche Verlagsgesellschaft, Govi-Verlag, Pharmazeutischer Verlag GmbH Eschborn, Stuttgart, 2015.
[22] Carbamazepine, European Directorate for the Quality of Medicines & HealthCare (EDQM), 9th ed., European Pharmacopoeia, 2016.
[23] pKa and logP values calculated by https://chemicalize.com/.
[24] R.A. Scherrer, S.M. Howard, Use of distribution coefficients in quantitative structure-activity relations, J. Med. Chem. 20 (1977) 53–58, http://dx.doi.org/10.1021/jm00211a010.
[25] K.L. Barnett, P.L. McGregor, G.P. Martin, et al., Stimuli article: analytical target profile (ATP): structure and application throughout the analytical lifecycle, Pharmacopeial Forum 42 (2016).
[26] N. Rácz, R. Kormány, Retention modeling in an extended knowledge space, Chromatographia 81 (2018) 585–594, http://dx.doi.org/10.1007/s10337-017-3466-0.
[27] E.F. Scriven, R. Murugan, pyridine and pyridine derivatives, in: Kirk-Othmer Encyclopedia of Chemical Technology, 2005.
[28] A. Albert, The chemical and biological properties of acridines, Sci. Prog. 37 (1933-) (1949) 418–434.
[29] J.R. Lakowicz, Principles of fluorescence spectroscopy, Springer, New York, NY, 2006.
[30] I. Molnár, A.H. Schmidt, H.-J. Rieger, J. Fekete, R. Kormany, UHPLC method development and modelling in the framework of quality by design, Column 10 (2014).
[31] J. Kirkland, J. Glajch, R. Farlee, Synthesis and characterization of highly stable bonded phases for high-performance liquid chromatography column packings, Anal. Chem. 61 (1989) 2–11.
[32] J.E. O'Gara, B.A. Alden, C.A. Gendreau, P.C. Iraneta, T.H. Walter, Dependence of cyano bonded phase hydrolytic stability on ligand structure and solution pH, J. Chromatogr. A 893 (2000) 245–251, http://dx.doi.org/10.1016/S0021-9673(00)00696-8.
[33] M. Roses, Determination of the pH of binary mobile phases for reversed-phase liquid chromatography, J. Chromatogr. A 1037 (2004) 283–298, http://dx.doi.org/10.1016/j.chroma.2003.12.063.
[34] E. Kovacs, J. Ermer, P.L. McGregor, et al., Stimuli article: analytical control strategy, Pharmacopeial Forum 42 (2016).
[35] T. Baczek, R. Kaliszan, H.A. Claessens, M.A. van Straten, Computer-assisted optimization of reversed-phase HPLC isocratic separations of neutral compounds, LC-GC Europe, 2001, pp. 14.