Journal of Chromatography A, 1217 (2010) 7222–7230

Contents lists available at ScienceDirect

Journal of Chromatography A

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

Selectivity optimization in green chromatography by gradient stationary phase optimized selectivity liquid chromatography

Kai Chen[a], Frédéric Lynen[a], Maarten De Beer[a], Laure Hitzel[b], Paul Ferguson[c], Melissa Hanna-Brown[d], Pat Sandra[a,*]

[a] Pfizer Analytical Research Centre, Ghent University, Krijgslaan 281–S4 bis, B-9000 Ghent, Belgium
[b] Chemical Technologies & Analytical Science, Pfizer Global R & D, Ramsgate Road, Sandwich, Kent CT13 9NJ, UK
[c] Research Analytics, Pfizer Global R & D, Ramsgate Road, Sandwich, Kent CT13 9NJ, UK
[d] Analytical R&D, Pfizer Global R & D, Ramsgate Road, Sandwich, Kent CT13 9NJ, UK

ARTICLE INFO

Article history:
Received 29 March 2010
Received in revised form 31 August 2010
Accepted 8 September 2010
Available online 17 September 2010

Keywords:
Liquid chromatography
Green chromatography
Selectivity
Stationary phase optimization

ABSTRACT

Stationary phase optimized selectivity liquid chromatography (SOSLC) is a promising technique to optimize the selectivity of a given separation by using a combination of different stationary phases. Previous work has shown that SOSLC offers excellent possibilities for method development, especially after the recent modification towards linear gradient SOSLC. The present work is aimed at developing and extending the SOSLC approach towards selectivity optimization and method development for green chromatography. Contrary to current LC practices, a green mobile phase (water/ethanol/formic acid) is hereby preselected and the composition of the stationary phase is optimized under a given gradient profile to obtain baseline resolution of all target solutes in the shortest possible analysis time. With the algorithm adapted to the high viscosity property of ethanol, the principle is illustrated with a fast, full baseline resolution for a randomly selected mixture composed of sulphonamides, xanthine alkaloids and steroids.

1. Introduction

Selectivity optimization is arguably the most important aspect of HPLC method development as it influences the resolution of a separation in a much larger way compared to the retention factor and the column efficiency. A large number of parameters are available to optimize the selectivity of a given separation problem. Classically the column dimensions and the type of stationary phase are preselected followed by optimization of the mobile phase composition such as the selection of organic modifier, of the buffer pH and of instrumental parameters such as the gradient profile and the column temperature. Due to the large number of possible combinations of these parameters, the optimization process is, in most cases, performed partially on experimental evidence and partially on personal experience. This combination of somewhat arbitrary choices and of incomplete experimental investigations often results in the selection of conditions which do not necessarily correspond to the best possible selectivity [1–5]. In silico method developing softwares have been proposed and commercialized in the last two decades, such as DryLab® [6], ChromSword® [7] and Osiris® [8]. These software tools are already well established and successfully applied in, e.g. the pharmaceutical industry.

Recently, with the shortage and high price of acetonitrile, the concept of green chromatography has regained attention [9–13]. The aim is hereby to minimize or eliminate the usage of environmentally hazardous organic solvents. In green chromatography, conventional organic modifiers in reversed-phase LC (RPLC) such as acetonitrile (moderate toxicity [14–16]) or additives such as trifluoroacetic acid (which is highly ecotoxic and is slow to biodegrade [17]), are substituted with environmentally friendly alternatives such as ethanol and formic acid. Ethanol is categorised as a 'green' solvent due to its low toxicity, potential synthesis from renewable feedstocks but most importantly because of its low lifecycle impact on the environment (that is its low ecological impact from synthesis through use and ease of recycling or disposal [14,16,18]). Similarly, formic acid is classed as a green solvent principally due to its rapid biodegradation to benign by-products (carbon dioxide and water) [17]. The development of a green chromatographic method therefore introduces the conceptual change to first choose the (green) mobile phase constituents followed by optimization of the remaining parameters which mainly comprises the choice of stationary phase.

A novel approach, stationary phase optimized selectivity liquid chromatography (SOSLC), commercialized as phase optimized liquid chromatography (POPLC) has recently been introduced as a promising tool for optimizing a separation [19–22]. With SOSLC the mobile phase is preselected after which the stationary phase is optimized. The approach is therefore ideally suited for green

* Corresponding author. Tel.: +32 56 204031; fax: +32 56 204859.
E-mail address: pat.sandra@richrom.com (P. Sandra).

0021-9673/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2010.09.029

chromatography. The procedure involves a number of basic measurements to obtain the retention factor (k) of each solute on each stationary phases after which an optimal serial connection of various stationary phase segments can be predicted resulting in the most favourable separation in the shortest possible analysis time. The optimization process and the corresponding algorithm are based on the PRISMA model, in which the lengths of column segments are the only variables to be optimized [20,23]. The algorithm explores all possible segment combinations within the restriction of a maximum allowable predicted analysis time and column length, calculates the selectivity for the critical pair in each possible combination and ranks it in decreasing order. The combination, which locates at the top of the list, is finally predicted as the optimal segment combination generating the best selectivity for the separation of a given mixture.

A major limitation of the PRISMA model is that it is limited to isocratic analysis. The analysis of a mixture of compounds covering a larger range of hydrophobicity therefore often results in excessive analysis time, loss of sensitivity and still incomplete separation when complex mixtures are involved. To overcome these limitations, a multiple step gradient method was developed [24]. According to their relative polarity or hydrophobicity, the compounds in the mixture are therefore, in a first step, classified into a few groups which are obtained by analysing the mixture under a linear gradient on a conventional RPLC column. The compound groups are thereby arbitrarily generated according to the actual clustering of the eluted peaks and the corresponding retention times in the gradient analysis. Each group is then handled with the classic isocratic optimization and a common column segment combination is obtained. A multiple step gradient profile consisting of a sequence of isocratic elutions is finally used for the optimized combination of segments. In the multiple step gradient optimizations, the retention time prediction is still based on isocratic measurements, which leads to certain deviations from the actual step-gradient analysis. Secondly, the order of segments is not taken into consideration in this approach, which also influences the retention times to some extent in gradient analysis.

To solve these problems, a linear gradient algorithm for SOSLC was recently proposed by our group [25]. Retention models of the compounds on the stationary phases are thereby first built as a function of the organic modifier concentration. The gradient elution is then considered as a sequence of small isocratic stages, for which the migrated distance and time of each analyte band can be calculated. The accumulated migration time of all the small isocratic stages is finally used as the predicted retention time in the gradient elution. The algorithm can as well be used in the isocratic, step-wise and linear gradient run mode.

In this contribution, the features of gradient SOSLC are demonstrated for green chromatography. Green mobile phase components and a fixed gradient are thereby pre-selected and the optimal column compositions (and order) are predicted, from a set of 8,037,725 unique column combinations, leading to baseline separation of the analytes. The concept is demonstrated with a mixture of 14 pharmaceutical compounds. The MS windows compatible software has been adapted to allow multi-linear gradients and is available for free [26].

2. Experimental

2.1. Chemicals and reagents

All chemicals and sample solutes were obtained from Sigma–Aldrich (Bornem, Belgium) except ethanol and formic acid which originated from Biosolve (Valkenswaard, The Netherlands). All stock solutions were prepared in acetonitrile. The mixture consisted of 14 pharmaceutical compounds and the individual pharmaceutical samples were prepared by diluting the stock solutions with water to a final concentration (Table 1). All the samples included uracil as unretained marker at a concentration of 5 μg mL^{-1}.

Table 1
The numbering and concentration of 14 compounds in the sample mixture.

Numbering	Compound name	Concentration in the mixture sample (μg mL^{-1})
1	Theobromine	800
2	Theophylline	50
3	Caffeine	100
4	Sulfadiazine	200
5	Sulfamerazine	50
6	Sulfamethazine	30
7	Sulfamethizole	30
8	Sulfamethoxazole	75
9	Estriol	1000
10	Sulfadimethoxine	75
11	Sulfaquinoxaline	200
12	Prednisone	75
13	Prednisolone	75
14	Cortisone	150

2.2. Instrumentation

All experiments were performed on an Agilent 1200 HPLC system (Agilent Technologies, Waldbronn, Germany) and with a POPLC® Basic Kit 250-5 (Bischoff Chromatography, Leonberg, Germany). The kit consists of five stationary phases: ProntoSIL C18 EPS 2, ProntoSIL C18 SH 2, ProntoSIL C30, ProntoSIL CN 2 and ProntoSIL Phenyl 2. Each stationary phase has a set of column segments with lengths of 10, 20, 40 (2×), 60 and 80 mm, which ensures the possibility of a segment combination of a single stationary phase from 1 to 25 cm. The column segments of all the five stationary phases have a particle size of 5 μm and the internal diameter is 3 mm. Chemstation software (Agilent Technologies) was used for data collection and peak integration. For stationary phase optimization, POPLC® Optimizer v 1.04.03 (Bischoff Chromatography) was used under isocratic conditions and the software developed by our group was used for the predictions under gradient conditions [26].

2.3. Chromatographic conditions

After preliminary tests, the column temperature was set at 50 °C and the flow rate at 0.5 mL min^{-1}. The wavelength of the VWD detector was set at 254 nm and the injected volumes were 1 μL. The mobile phases were composed of water with 0.1% formic acid and ethanol. For isocratic optimization, the mobile phase composition was 85% (v/v) water and 15% (v/v) ethanol. For gradient optimization, the basic measurements were performed at 8 isocratic levels, i.e. 10%, 15%, 20%, 25%, 30%, 35%, 40% and 45% (v/v) ethanol and the column length was 10 cm for all the five stationary phases. The mobile phase was post-mixed by a binary pump. A linear gradient condition was then employed for gradient optimization from 10% to 50% (v/v) ethanol in 30 min on the optimized segment combination. A detailed list of the experiments is included in Table 2.

3. Theory and algorithm

3.1. Isocratic optimization

The migration time of a certain compound on a given stationary phase segment can be described as the upper limit of the integral in Eq. (1). The retention time on the combined column can therefore

Table 2
List of experiments. Common conditions: the column temperature was set at 50 °C and the flow rate at 0.5 mL min^{-1}. The wavelength of the VWD detector was set at 254 nm and the injected volumes were 1 μL. The mobile phases were composed of water with 0.1% formic acid (solvent A) and ethanol (solvent B).

Numbering	Type	Mobile phase	Stationary phase	Sample
1	Basic measurements	Isocratic, 90% A + 10% B	10 cm single segment (five stationary phases respectively)	Single compound (uracil and 14 compounds respectively)
2	Basic measurements	Isocratic, 85% A + 15% B	10 cm single segment (five stationary phases respectively)	Single compound (uracil and 14 compounds respectively)
3	Basic measurements	Isocratic, 80% A + 20% B	10 cm single segment (five stationary phases respectively)	Single compound (uracil and 14 compounds respectively)
4	Basic measurements	Isocratic, 75% A + 25% B	10 cm single segment (five stationary phases respectively)	Single compound (uracil and 14 compounds respectively)
5	Basic measurements	Isocratic, 70% A + 30% B	10 cm single segment (five stationary phases respectively)	Single compound (uracil and 14 compounds respectively)
6	Basic measurements	Isocratic, 65% A + 35% B	10 cm single segment (five stationary phases respectively)	Single compound (uracil and 14 compounds respectively)
7	Basic measurements	Isocratic, 60% A + 40% B	10 cm single segment (five stationary phases respectively)	Single compound (uracil and 14 compounds respectively)
8	Basic measurements	Isocratic, 55% A + 45% B	10 cm single segment (five stationary phases respectively)	Single compound (uracil and 14 compounds respectively)
9	Isocratic optimization	Isocratic, 85% A + 15% B	1 cm C30 + 4 cm C18 + 12 cm C18EPS	Sample mixture
10	Gradient optimization, Rank 1	Gradient from 10% B to 50% B in 30 min	6 cm C18 + 12 cm C18EPS	Sample mixture
11	Gradient optimization, Rank 20	Gradient from 10% B to 50% B in 30 min	5 cm C18 + 12 cm C18EPS + 1 cm CN	Sample mixture
12	Gradient optimization, Rank 50	Gradient from 10% B to 50% B in 30 min	8 cm C18EPS + 5 cm C30 + 5 cm Phenyl	Sample mixture

be calculated by Eq. (2):

$$l_i = \int_{t=0}^{Rt_i} v_i\,dt = \int_{t=0}^{Rt_i} \frac{u}{(1+k_i)}dt \tag{1}$$

$$Rt = \sum_{i=1}^{5} Rt_i + t_{m,EC} \tag{2}$$

where l_i is the length of the ith segment, v_i is the linear velocity of an analyte band, k_i is the retention factor of the analyte on the ith segment and u is the linear velocity of the mobile phase determined by the elution time of an unretained marker (uracil). $t_{m,EC}$ is the extra-column void time measured by the elution time of uracil through the system whereby the column is replaced by a union. Rt_i is the retention time on the ith segment and Rt is the retention time on the combined column including the system void time.

Under isocratic conditions, the retention factor k_i is a constant and Eq. (1) can be rewritten to Eqs. (3) and (4).

$$l_i = \frac{u}{(1+k_i)}\int_{t=0}^{Rt_i} dt = \frac{u}{(1+k_i)}\cdot Rt_i \tag{3}$$

$$Rt_i = \frac{(1+k_i)}{u}\cdot l_i \tag{4}$$

Because the total length of the combined column L is the sum of the lengths of segments l_i (Eq. (5)), Eq. (6) can be obtained by replacing Rt_i in Eq. (2) with Eq. (4).

$$L = \sum_{i=1}^{5} l_i \tag{5}$$

$$Rt = \sum_{i=1}^{5}\frac{(1+k_i)}{u}l_i + t_{m,EC} = \frac{L}{u} + \frac{1}{u}\cdot\sum_{i=1}^{5} k_i\cdot l_i + t_{m,EC} \tag{6}$$

Assuming that the combined column has an experimental retention factor k_c, it can be defined and rewritten as Eq. (7). Eq. (8) is obtained by combining Eq. (6) and Eq. (7).

$$Rt = \frac{L}{u}(1+k_c) + t_{m,EC} = \frac{L}{u} + \frac{L}{u}\cdot k_c + t_{m,EC} \tag{7}$$

$$k_c = \frac{\sum_{i=1}^{5} k_i\cdot l_i}{L} \tag{8}$$

Equation (8) is also known as the PRISMA model valid for isocratic conditions [20].

3.2. Gradient optimization

Under gradient conditions, however, the retention factor k is no longer a constant. For a single stationary phase, the retention factor k_i at a given time point is mainly influenced by the composition of the mobile phase which is around the position of the analyte band. To investigate the relationship between the retention factor and the composition of mobile phase, the classical LSSM retention model [27] assumes a linear relationship between the logarithm of the retention factor, ln(k) and the volume fraction of the organic modifier φ in the mobile phase, as shown in Eq. (9):

$$\ln(k) = b\varphi + c \tag{9}$$

where b and c are the coefficients which can be obtained by a linear regression.

The LSS model is widely used and has adequate accuracy in retention time prediction under gradient conditions, especially when a relative small range of φ is involved [28,29]. However, the relationship between ln(k) and φ is no longer linear if the whole range of φ is investigated as shown in Fig. 1. In this work the second order polynomial expression developed by Schoenmakers et al. [30,31] is therefore used to obtain an improved relationship between (k) and φ (Eq. (10)).

$$\ln(k) = a\varphi^2 + b\varphi + c \tag{10}$$

where a, b and c are the coefficients obtained from the quadratic regression model based on the basic isocratic measurements with at least 3 different levels of the volume fraction of organic modifier. In practice, the number as well as the set of φ values can be different for each compound. For the sake of time saving, the basic measurements for the compounds with large retention will be

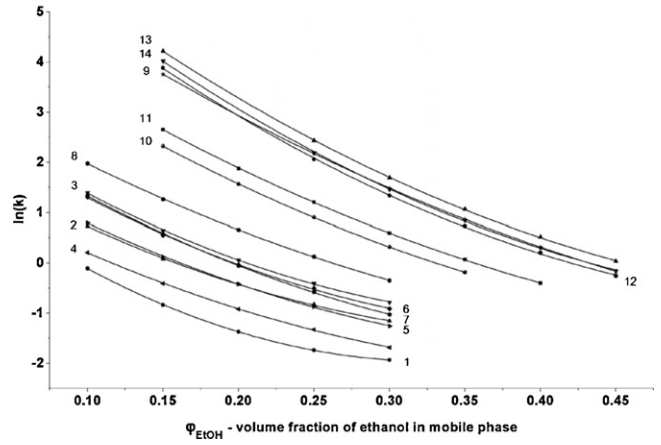

Fig. 1. Quadratic regression of $\ln(k)$ vs. φ_{EtOH} on 10 cm C18 SH2 column. At least 5 isocratic levels are used in the basic measurements for each compound. For compound identification, please refer to Table 1.

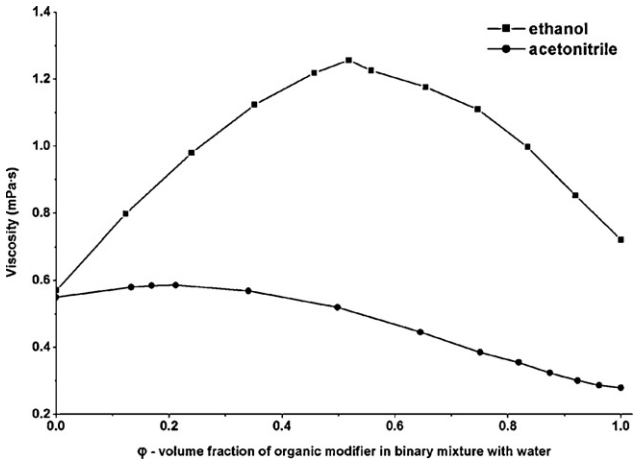

Fig. 2. Graphical representation of viscosities of ethanol and acetonitrile vs. volume fraction in the binary mixture with water at 50 °C. Data from [42] and [43].

investigated with relative large φ values. On the other hand, relative small φ values are set in the basic measurements for the compounds with small retention, so that more accurate values of the retention factor can be obtained. In this study, basic measurements of at least 5 isocratic levels are investigated for each compound for the enhancement of the regression model's reliability (Fig. 1).

In gradient conditions, however, the instant volume fraction of modifier is dependant on both the time elapsed and the distance migrated by the analyte band. It is possible to obtain an integrated solution for Eq. (1) to predict gradient retention as was, e.g. shown by Schoenmakers et al. for single phases [32]. Neue and Kuss recently proposed alternatives to Eq. (10) to predict retention times more accurately, especially when predictions are made making use of extrapolations outside the range of measurements [33,34]. However, in this work Eq. (10) was used for the calculations of k as broad ranges of φ were used in the initial measurements. Because in the gradient SOSLC procedure the calculation of the retention time of each analyte is required for all 8,037,725 unique column combinations, and if whished for a variety of gradient profiles, a numerical solution was strongly favored for the sake of simplicity. Note that the POPLC® Basic Kit allows only 142,505 possible column combinations because, under isocratic conditions, the segment order is of no influence to the predicted retention times. This number is significantly larger when gradients are applied. Although numerical approaches are less elegant than integrated solutions they have been used quite extensively before for the prediction of retention time in chromatography [35–40].

The basic numerical integration algorithm is described as follows. Assume that t_e is the time elapsed after the analyte band enters the front of the first segment, d_m is the migration distance of the analyte band from the front of the first segment and i is the numbering of the segment where the analyte is passing by.

$$\varphi = f\left(t_e - \frac{d_m}{u} - t_{m,EC} - t_{dwell}\right) \quad (11)$$

$$\ln(k) = a_i\varphi^2 + b_i\varphi + c_i \quad (12)$$

$$d_{m,updated} = d_m + u \cdot \frac{1}{1+k} \cdot \Delta t \quad (13)$$

$$t_{e,updated} = t_e + \Delta t \quad (14)$$

where $f(t)$ is the gradient function of volume fraction of organic modifier dependant on time t. Note that the function $f(t)$ can be defined as for multi-linear gradient or any other gradient profile. If t is less than 0, $f(t)$ equals to $f(0)$, the φ value at the starting point of the gradient profile. t_{dwell} is the dwell time which is needed for the front of mobile phase to migrate from the solvent mixer to the inlet of the column. The coefficients a_i, b_i and c_i are the coefficients of the analyte's retention model on ith segment in the combined column. Δt, in general set to 1 s, is the constant and small time span for which the gradient elution is considered as an isocratic stage. Smaller Δt can be used if more accurate results are required despite of the longer computing time. As the iterative variable, $d_{m,\,updated}$ is the updated migration distance of the analyte band after the time increment of Δt in an iteration and then substitutes for d_m in the next iteration. Similarly, $t_{e,\,updated}$ is the updated elapsed time with an addition of Δt to t_e for each iteration. The iteration process of Eqs. (11)–(14) starts with the variables t_e and d_m set to 0. The segment numbering i will be increased by 1 if d_m reaches the next segment. Once d_m reaches the total length of the combined column, the iteration process stops and the retention time Rt will be the addition of t_e and $t_{m,EC}$.

Given a proposed gradient profile and a restriction on the maximum analysis time, the retention times of all the compounds in a mixture can be predicted for all the possible column segment combinations. For each segment combination, the retention time difference is calculated for the most critical pair of adjacent peaks and employed as the ranking score of this combination. Finally, all the possible column combinations are sorted according to the ranking scores in descending order. The segment combination at the top of the list can then be selected and employed as the optimal column.

It should be noted that the retention time difference of the peaks of the critical pair instead of the selectivity factor α is employed in this work as the ranking factor. Because the resolution of a separation under a linear gradient condition is defined by the retention time difference and a nearly constant average peak width of adjacent eluted peaks, the difference of retention times is regarded here as a more suitable evaluation objective for the separation than the selectivity factor α, especially for the peaks with very small retention factors. The possibility of over-fitting is eliminated by a restriction on the maximum analysis time.

A software package based on the above-mentioned algorithm was developed and available free of charge from Bischoff Chromatography, Germany [26].

4. Results and discussion

Compared to acetonitrile, ethanol leads to a much larger viscosity in a binary solvent mixture with water, especially in the range of volume fractions from 30% to 50% (v/v) (Fig. 2) [41].

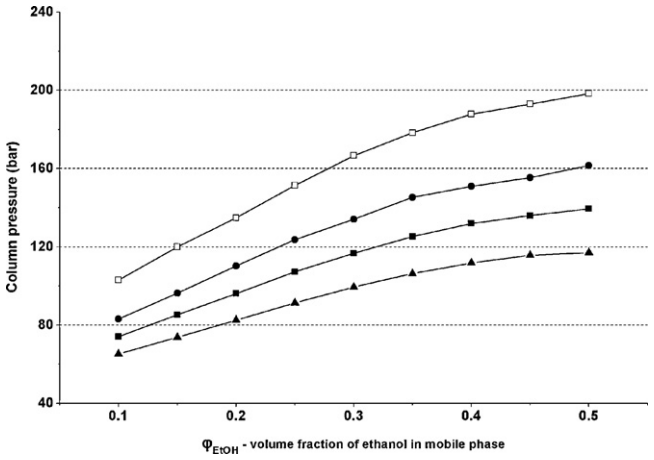

Fig. 3. Column pressure drop of a 10 cm C18 segment combination at different flow rates and column temperatures. Blank square (□): flow rate of 0.5 mL min^{-1} and column temperature at 30 °C; solid circle (●): flow rate of 0.55 mL min^{-1} and column temperature at 50 °C; solid square (■): flow rate of 0.5 mL min^{-1} and column temperature at 50 °C; solid triangle (▲): flow rate of 0.45 mL min^{-1} and column temperature at 50 °C.

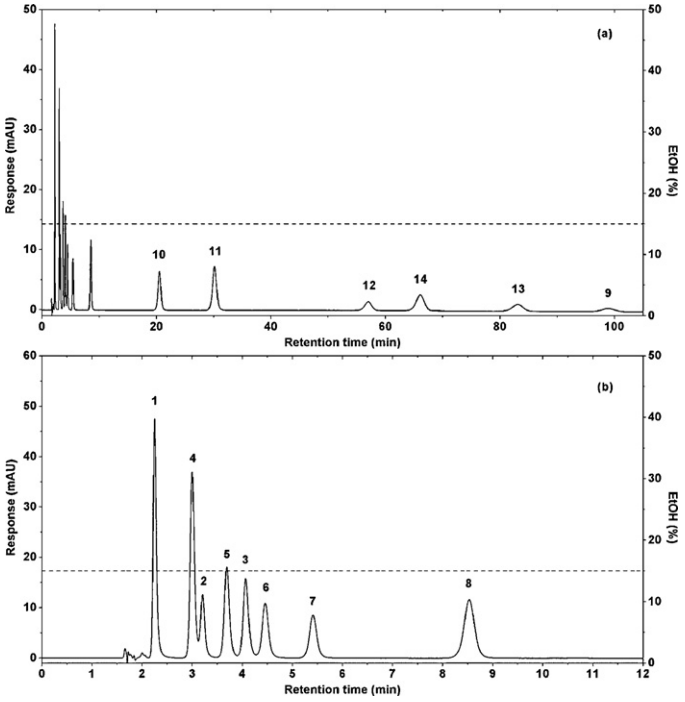

Fig. 4. Chromatogram of the isocratic optimization. Optimized segment combination: 1 cm C30, 4 cm C18 and 12 cm C18EPS. (a) Full chromatogram; (b) zoomed section from 0 to 12 min. For peak identification, please refer to Table 1.

In preliminary work the limitations of the maximum allowable column pressure drop, the suitable mobile phase flow rates and column temperatures were investigated. A series of measurements was done at different flow rate and temperature by using a 10 cm long segment combination of C18. The results are shown in Fig. 3. The maximum length of segment combination in this work was 25 cm for each stationary phase and the maximum allowable column pressure was 400 bar. In other words, the maximum column pressure is about 160 bar for a 10 cm long segment combination. The data in Fig. 3 shows that the optimal chromatographic condition consists of the combination of a column temperature of 50 °C with a flow rate of 0.5 mL min^{-1}. Both the combinations of 30 °C with 0.5 mL min^{-1} and 50 °C with 0.55 mL min^{-1} resulted in a column pressure exceeding 160 bar at higher ethanol fractions. The combination of 50 °C with 0.45 mL min^{-1} was not employed as it was leading to longer analysis times.

An isocratic optimization with 15% (v/v) ethanol was first of all performed with the POPLC Optimizer of Bischoff Chromatography. The extra-column void time $t_{m,EC}$ was determined to be 0.045 min for uracil. The optimized combination of segments was thereby 1 cm C30 coupled with 4 cm C18 and 12 cm C18EPS. The corresponding experimental chromatogram is shown in Fig. 4. The total analysis time was about 110 min and the last eluting compound estriol (peak 9) shows a low detector response because of the large retention time. In addition, the selectivity of the critical pair, sulfadiazine (peak 4) and theophylline (peak 2), was unsatisfactory. This result shows that isocratic SOSLC is considerably limited in both analysis time and selectivity, especially when analyzing mixtures of compounds covering a large range of hydrophobicities [22].

Before the gradient SOSLC algorithm could be applied, the basic measurements for the retention models needed to be performed. The length of the segment combination for each stationary phase was thereby 10 cm. For each compound on each stationary phase, a series of isocratic analyses was performed on at least 5 levels of volume fraction of ethanol (see also Table 2). Note that compared to acetonitrile and methanol, ethanol has a higher eluotropic strength and therefore the scouting runs will be shorter. Additionally, for the same reason, less ethanol is likely to be required which is another "green" feature. In total 70 quadratic models (14 compounds on 5 stationary phases) were built based on the retention times obtained in the basic measurements. The retention models constructed for the C18 phase are shown in Fig. 1. The average regression correlation coefficient R^2 of the 70 models was 0.9998. This suggests that the experimental data fits well on the models and that reliable data prediction can be obtained for the gradient SOSLC application. The dwell time t_{dwell} in Eq. (11) was then determined to be 1.50 min at the flow rate of 0.5 mL min^{-1} and the dwell volume was therefore estimated to be 0.75 mL. A linear gradient, in which the volume fraction of ethanol increases from 10% to 50% (v/v) in 30 min, was used as a fixed gradient for the stationary phase optimization by applying the in-house developed software and the retention models. The combination of a column of 6 cm C18 and 12 cm C18EPS was proposed by the software as the optimal result. The predicted and the experimental separation are shown in Fig. 5a and c, respectively. The corresponding retention times are given in Table 3.

It is clear from these results that the predicted top ranked column combination is indeed resulting into baseline separation of all analytes and that the expected elution order and selectivities of

Table 3
Comparison of the predicted retention times with the uncalibrated algorithm for gradient analysis and the experimental retention time on the optimal segment combination: 6 cm C18 and 12 cm C18EPS. Gradient: 10–50% v/v ethanol in 30 min. For peak identification, see Table 1.

Numbering	Predicted retention time (min)	Observed retention time (min)	Relative deviation (%)
1	3.224	3.212	0.37
2	4.897	5.075	−3.51
3	6.292	6.628	−5.07
4	4.381	4.458	−1.73
5	5.588	5.804	−3.72
6	6.816	7.126	−4.35
7	7.728	8.210	−5.87
8	10.269	10.902	−5.81
9	22.382	23.308	−3.97
10	14.805	15.694	−5.66
11	16.728	17.662	−5.29
12	19.192	20.005	−4.06
13	21.045	21.904	−3.92
14	19.802	20.626	−3.99

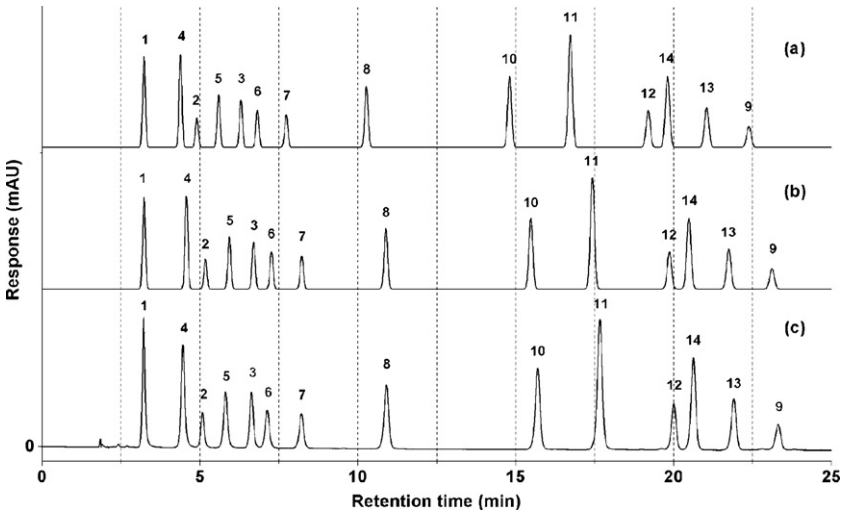

Fig. 5. Comparison of predicted and experimental chromatograms on the optimal segment combination: 6 cm C18 and 12 cm C18EPS. (a) Predicted chromatogram with the algorithm as such; (b) predicted chromatogram with the empirical calibration procedure; (c) experimental chromatogram.

the various peaks are correctly reflected in the confirmatory experiments. It can, however, also be seen that the predicted retention times for most compounds underestimate the experimental results. The maximal discrepancy in retention time comprises about 6% in this example (peak No. 7 in Fig. 5). Other, multi-linear, gradient profiles were also investigated which are given in Table 4. The corresponding predictions and experimental confirmations are shown in Table 5. The errors in prediction thereby run up to 19% in gradient 2 in Table 4, whereby a steep gradient is used in the beginning of the chromatogram. In all cases, however, the predicted column combination presented a solution leading to baseline separation of all peaks in the experimental chromatograms. The reason for this is that the software is ranking the signals according to the maximal retention time difference for every signal. It appears that the resulting over-resolution of the peaks in the chromatogram is such that it is difficult to undo by the observed discrepancies between prediction and experiment. Additionally as in all cases the shifts are quite systematic (all peaks shift either to the left or to the right

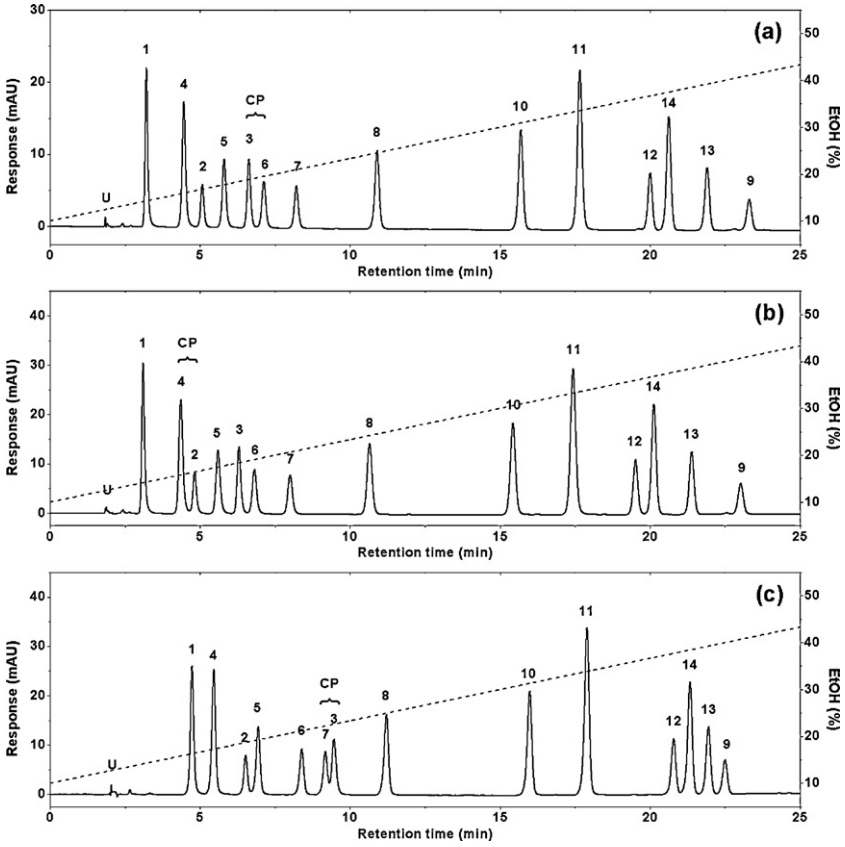

Fig. 6. Experimental chromatograms on the segment combinations obtained in gradient optimization. (a) Chromatogram of segment combination of rank 1; (b) chromatogram of segment combination of rank 20; (c) chromatogram of segment combination of rank 50. CP means critical pair. For detail description, please refer to Table 6. For peak identification, please refer to Table 1.

Table 4
Investigated multi-linear gradient profiles.

No.	Gradient profile				
1	Time (min)	0	7	29	64
	φ_{EtOH} (v/v)	10%	10%	29%	32%
2	Time (min)	0	1	8	26
	φ_{EtOH} (v/v)	0%	11%	28%	30%
3	Time (min)	0	6	7	86
	φ_{EtOH} (v/v)	4%	17%	29%	34%

in a particular region of the chromatogram). Effective overlap of 2 peaks during the experiments contrary to a prediction is not taking place with the analyzed samples.

The reliability of the top ranked predictions that they do deliver good separation of all the analytes is illustrated in Table 6 and Fig. 6. The chromatograms obtained with the first, twentieth and fiftieth ranked suitable column combination is thereby shown and it can be seen that only in the latter case the resolution between a critical pair starts to be affected. In Fig. 6a and b all 14 compounds are baseline separated in 25 min. Compared with the results obtained by isocratic optimization (Fig. 4), the analysis time was considerably shortened and the selectivity of the critical pair was improved in the gradient optimization. This demonstrates that the use of SOSLC with green mobile phases can achieve comparable separations to what can be obtained by using acetonitrile.

In order to improve the accuracy of the retention time predictions a survey of plausible causes was performed. HPLC hardware issues due to inaccurate flow rates, mixing or compressibility settings could be excluded suggesting inaccuracies in the numerical approach. The most probable cause of the observed discrepancies is thereby the void time measurements. Even very small errors in void time measurement can thereby eventually lead to the observed dis-

Table 5
Comparison of the predicted retention times obtained from the uncalibrated and calibrated algorithms with the observed retention times. For the gradient profile details, see Table 4. Compound identification, see Table 1.

No. gradient profile	Numbering of compounds	Observed *Rt* (min)	Without calibration		With calibration	
			Predicted *Rt* (min)	Relative deviation (%)	Predicted *Rt* (min)	Relative deviation (%)
1	1	3.890	4.000	2.83	4.000	2.83
	4	4.748	5.052	6.40	5.052	6.40
	2	5.721	6.036	5.51	6.036	5.51
	5	6.328	6.861	8.42	6.861	8.42
	6	8.410	9.292	10.49	9.292	10.49
	3	9.411	10.195	8.33	10.195	8.33
	7	10.194	11.040	8.30	11.078	8.67
	8	12.848	12.393	−3.54	13.293	3.46
	10	15.118	14.339	−5.15	15.127	0.06
	11	16.514	15.727	−4.77	16.424	−0.54
	12	18.746	18.461	−1.52	18.899	0.82
	14	19.466	19.288	−0.91	19.696	1.18
	13	20.812	20.641	−0.82	21.037	1.08
	9	22.934	22.473	-2.01	22.936	0.01
	Average	–	–	4.93	–	4.13
2	1	5.780	4.683	−18.98	5.622	−2.73
	4	6.445	5.336	−17.21	6.250	−3.03
	2	6.808	5.692	−16.39	6.533	−4.04
	5	7.231	6.146	−15.00	6.994	−3.28
	3	7.613	6.522	−14.33	7.297	−4.15
	6	7.942	6.895	−13.18	7.716	−2.85
	7	8.596	7.531	−12.39	8.347	−2.90
	8	10.163	9.111	−10.35	9.985	−1.75
	10	12.780	11.705	−8.41	12.385	−3.09
	11	14.436	13.294	−7.91	13.904	−3.69
	12	16.474	15.412	−6.45	15.937	−3.26
	14	17.349	16.279	−6.17	16.792	−3.21
	13	19.503	18.379	−5.76	18.877	−3.21
	9	23.370	21.954	−6.06	22.453	−3.92
	Average	–	–	11.33	–	3.22
3	1	5.965	5.156	−13.56	5.732	−3.91
	4	7.044	6.240	−11.41	6.814	−3.27
	2	7.722	6.828	−11.58	7.584	−1.79
	5	8.406	7.586	−9.75	8.315	−1.08
	3	9.026	8.172	−9.46	9.069	0.48
	6	9.550	8.768	−8.19	9.581	0.32
	7	10.427	9.601	−7.92	10.579	1.46
	8	11.457	10.550	−7.92	11.834	3.29
	10	13.493	12.546	−7.02	13.114	−2.81
	11	15.010	14.058	−6.34	14.349	−4.40
	12	16.463	15.755	−4.30	16.216	−1.50
	14	17.190	16.543	−3.76	17.013	−1.03
	13	19.096	18.532	−2.95	18.976	−0.63
	9	22.937	22.472	−2.03	21.812	−4.90
	Average	–	–	7.59	–	2.20

Table 6
The predicted segment combinations in gradient optimization and the corresponding critical pair, predicted and observed retention time differences. ΔRt is the retention time difference between the peaks of the critical pair. For peak identification, please refer to Table 1.

Rank	Segment combination	Critical pair	Predicted ΔRt (min)	Observed ΔRt (min)
1	6 cm C18 + 12 cm C18EPS	No. 3, No. 6	0.564	0.498
20	5 cm C18 + 12 cm C18EPS + 1 cm CN	No. 4, No. 2	0.459	0.464
50	8 cm C18EPS + 5 cm C30 + 5 cm phenyl	No. 7, No. 3	0.252	0.287

crepancies between prediction and experiment. In this study the time counter (Δt) was set at 1 s. Therefore a late eluting peak such as peak 9 in Fig. 5a, went through 1343 iterative cycles whereby a small error can be continuously accumulated. The fact that for a constant linear gradient the errors mostly increase with increasing retention time sustains this hypothesis. This can also explain the deviations observed in the multi-linear gradients. In gradient 1 in Table 5, the discrepancies between prediction and experiment in the first 7 min. do not exceed the shifts observed when using the POPLC kit in the conventional isocratic way [44]. The gradient applied from then on (including dwell time) leads again to negative residuals which are similar to what was observed in Fig. 5 and Table 3. Comparable effects are visible in gradients 2 and 3 in Table 5. It appears that the steep gradient in the beginning of gradient 2 is also leading to high discrepancies. It can thereby be argued if a time counter (Δt) of 1 s is sufficiently small for a fast gradient as applied in that section of the chromatogram.

However, no single void marker is truly unretained over a broad range of organic modifier fractions, especially when various stationary phases are used. Uracil is somewhat retained under RPLC conditions at low (and very high) modifier fractions and depicts a minimum around 50% modifier and is accepted as a suitable void marker for C18 and C18EPS stationary phases [45]. Although it could be expected that salts such as KI and $NaNO_2$, which can be detected at low UV wavelengths, are a better choice compared to small, polar organic molecules, it appears that salts are also retained due to various ion exchange and exclusion phenomena with residual silanol functions on the stationary phase [46]. The retention of KI (100 mM, 5 μL injection) vs. uracil is shown in Fig. 7 for the ethanol fractions used in the present study on a C18EPS segment. As it is clear that uracil is preferable for most conditions compared to the salt, the former was used as t0 marker in this study. In the proposed algorithm, quadratic models where therefore built to estimate and predict the retention of uracil for each stationary phase. As the average regression coefficient R^2 for uracil's quadratic models on the five stationary phases was 0.990, which is worse than that of the models for the compounds, it might explain the observed deviation between experiment and prediction.

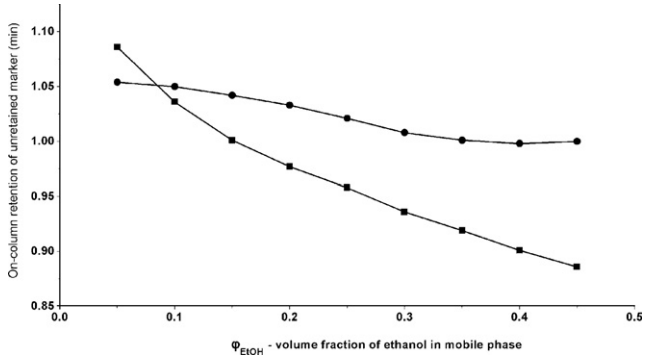

Fig. 7. On-column retention times of uracil and KI used as unretained markers on 10 cm C18EPS segment column at a flow rate of 0.5 mL min^{-1} and a column temperature at 50 °C. Solid square (■): uracil; solid circle (●): KI.

In order to improve the accuracy of retention time prediction, an empirical calibration is proposed here which acts on the migrated distance of analyte band in the numerical integration. A calibration coefficient is therefore introduced in Eq. (13) leading to

$$d_{m,updated} = d_m + \frac{1}{1+p\cdot s}\cdot u \cdot \frac{1}{1+k}\cdot \Delta t \quad (15)$$

where s is the concurrent slope of the gradient curve (the instant increment of the volume fraction φ per second) and p is the empirical constant coefficient with the unit of time in seconds.

The empirical constant p is obtained by minimizing the prediction deviation of two selected compounds' retention times under the given gradient (10–50% (v/v) in 30 min). In this study, for example, the experimental retention times of sulfadiazine (No. 4) and estriol (No. 9) under the linear gradient profile were compared with those predicted and the empirical constant p was gradually tuned to be 8 s.

To verify the calibration, the predicted chromatogram of the sample mixture from the calibrated algorithm is shown in Fig. 5b and overlaid with the predicted chromatogram from the uncalibrated algorithm (Fig. 5a) and the experimental chromatogram (Fig. 5c). A significant improvement of prediction accuracy can be observed. For multiple-linear gradients shown in Table 4, the prediction deviation after calibration is shown in Table 5 which also demonstrates the improvements in prediction which can be obtained by making use of this empirical calibration procedure.

Compared to established predictive software packages such as DryLab®, gradient SOSLC method currently requires a considerable number of basic measurements. The use of LC–MS to track all peaks simultaneously should, however, solve this problem as this should result in 5 or 6 runs per column (for each organic modifier fraction) for each stationary phase. The columns can also be imagined in a system equipped with automatic column selection which should allow unattended SOSLC optimization. Note that SOSLC makes maximal use of the orthogonality of the selectivity from the different stationary phases. This approach has proven to offer many practical solutions for the analysis of complex samples before [47]. Although some commercial software allow the isocratic SOSLC option (making use of Eq. (8)), the gradient approach described here is not included in other predictive software packages. Therefore, gradient SOSLC can be considered as a new tool in the family of in-silico method development approaches which is offering promising prospects in the field of green chromatography.

5. Conclusion

Gradient SOSLC has been extended to green chromatography. Taking into consideration the high viscosity of ethanol, a flow rate of 0.5 mL min^{-1} and a column temperature of 50 °C were used as operation conditions. A numerical integration based on the quadratic retention models was applied to predict the analytes' retention times under gradient conditions. In addition, an empirical calibration is proposed to correct the predicted retention time shift probably caused by the inaccurate measurements of void time and accumulative error in the numerical iteration process under gradient conditions. As a result, the optimized segment combination for a

given gradient profile can lead to a fast and full baseline separation, showing competitive quality of separation to what can be obtained when using acetonitrile. Green SOSLC can therefore help in fundamentally changing the way in which HPLC methods are developed: first preselecting a green mobile phase and then optimizing the stationary phase.

Acknowledgment

The authors are grateful to the National Natural Science Foundation of China (Grant No. 20705024) for their cooperation.

References

[1] L.R. Snyder, J. Chromatogr. Sci. 16 (1978) 223.
[2] A.C.J.H. Drouen, H.A.H. Billiet, P.J. Schoenmakers, J. De Galan, Chromatographia 16 (1982) 48.
[3] P.J. Schoenmakers, T. Blaffert, J. Chromatogr. 384 (1987) 117.
[4] R.M.L. Marques, P.J. Schoenmakers, J. Chromatogr. 592 (1992) 157.
[5] P.H. Lukulay, V.L. McGuffin, Anal. Chem. 69 (1997) 2963.
[6] L.R. Snyder, J.W. Dolan, D.C. Lommen, J. Chromatogr. 485 (1989) 65.
[7] S. Galushko, V. Tanchuk, I. Shishkina, O. Pylypchenko, W.D. Beinert, in: S. Kromidas (Ed.), HPLC Made to Measure: A Practical Handbook for Optimization., WILEY–VCH Verlag GmbH & Co. KGaA, Weinheim, 2006, p. 557.
[8] S. Goga, S. Heinisch, J.L. Rocca, Chromatographia 48 (1998) 237.
[9] P.T. Anastas, Crit. Rev. Anal. Chem. 29 (1999) 167.
[10] F. Brettschneider, V. Jankowski, T. Günthner, S. Salem, M. Nierhaus, A. Schulz, W. Zidek, J. Jankowski, J. Chromatogr. B 878 (2010) 763.
[11] A.S. Pereira, F. David, G. Vanhoenacker, P. Sandra, J. Sep. Sci. 32 (2009) 2001.
[12] B.W. Wenclawiak, S. Giegold, T. Teutenberg, Anal. Lett. 41 (2008) 1097.
[13] E. Destandau, E. Lesellier, Chromatographia 68 (2008) 985.
[14] C. Jimenez-Gonzalez, A.D. Curzons, D.J.C. Constable, V.L. Cunningham, Clean Technol. Environ. Policy 7 (2005) 42.
[15] K. Alfonsi, J. Colberg, P.J. Dunn, T. Fevig, S. Jennings, T.A. Johnson, H.P. Kleine, C. Knight, M.A. Nagy, D.A. Perry, M. Stefaniak, Green Chem. 10 (2008) 31.
[16] C. Capello, U. Fischer, K. Hungerbuhler, Green Chem. 9 (2007) 917.
[17] Internal data taken from the Pfizer Green Chemistry Acid and Base Selection Guide.
[18] Environmental management – Life cycle assessment – Principles and framework, EN ISO 14040, European Committee for Standardisation, Brussels, Belgium, 1997.
[19] S. Nyiredy, Z. Szűcs, J. Szepesy, Chromatographia 63 (2006) S3.
[20] S. Nyiredy, Z. Szűcs, L. Szepesy, J. Chromatogr. A 1157 (2007) 122.
[21] M. Kuehnle, J. Rehbein, K. Holtin, B. Dietrich, M. Gradl, H. Yeman, K. Albert, J. Sep. Sci. 31 (2008) 1655.
[22] M. Zedda, J. Tuerk, T. Teutenberg, S. Peil, T.C. Schmidt, J. Chromatogr. A 1216 (2009) 8910.
[23] S. Nyiredy, B. Meier, C.A.J. Erdelmeier, O. Sticher, J. High Resolut. Chromatogr. 4 (1985) 186.
[24] M. De Beer, F. Lynen, M. Hanna-Brown, P. Sandra, Chromatographia 69 (2009) 609.
[25] M. De Beer, F. Lynen, K. Chen, P. Ferguson, M. Hanna-Brown, P. Sandra, Anal. Chem. 82 (2010) 1733.
[26] http://www.poplc.de/lgpoplcoptimizer_en.html.
[27] L.R. Snyder, J.W. Dolan, J.R. Gant, J. Chromatogr. 165 (1979) 3.
[28] J.C. Ford, J. Ko, J. Chromatogr. A 727 (1996) 1.
[29] H. Vua, V. Spicera, A. Gotfrida, O.V. Krokhin, J. Chromatogr. A 1217 (2010) 489.
[30] P.J. Schoenmakers, H.A.H. Billiet, R. Tijsen, L. De Galan, J. Chromatogr. 149 (1978) 519.
[31] P.J. Schoenmakers, H.A.H. Billiet, L. De Galan, J. Chromatogr. 185 (1979) 179.
[32] P.J. Schoenmakers, H.A.H. Billiet, R. Tijssen, L. de Galan, J. Chromatogr. 149 (1978) 519.
[33] U.D. Neue, H.J. Kuss, J. Chromatogr. 1217 (2010) 3794.
[34] U.D. Neue, Chromatographia 63 (2006) S45–S53.
[35] S.A. Tomellini, R.A. Hartwick, H.B. Woodruff, Anal. Chem. 57 (1985) 811.
[36] S.A. Tomellini, S.H. Hsu, R.A. Hartwick, Anal. Chem. 58 (1986) 904.
[37] L. Wang, L. Zhang, H. Xiao, X. Liang, Anal. Chim. Acta 476 (2003) 123.
[38] G. Hendriks, J.P. Franke, D.R.A. Uges, J. Chromatogr. A 1089 (2005) 193.
[39] P. Nikitas, A. Pappa-Louisi, A. Papageorgiou, J. Chromatogr. A 1157 (2007) 178.
[40] P. Nikitas, A. Pappa-Louisi, J. Chromatogr. A 1216 (2009) 1737.
[41] T. Teutenberg, S. Wiese, P. Wagner, J. Gmehling, J. Chromatogr. A 1216 (2009) 8470.
[42] R. Belda, J.V. Herraez, O. Diez, Phys. Chem. Liq. 42 (2004) 467.
[43] M.A. Saleh, S. Akhtar, M.S. Ahmed, Phys. Chem. Liq. 44 (2006) 551.
[44] J. Lu, M. Ji, R. Ludewig, G.K.E. Scriba, D.Y. Chen, J. Pharm. Biomed. Anal. 51 (2010) 764.
[45] D. Cabooter, F. Lynen, P. Sandra, G. Desmet, J. Chromatogr. A 1157 (2007) 131.
[46] F.Z. Oumada,.M. Roses, E. Bosch, Talanta 53 (2000) 667.
[47] S. Louw, A.S. Pereira, F. Lynen, M. Hanna-Brown, P. Sandra, J. Chromatogr. A 1208 (2008) 90.