Journal of Pharmaceutical and Biomedical Analysis 51 (2010) 46–55

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

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

Investigation of the HPLC response of NSAIDs by fractional experimental design and multivariate regression analysis. Response optimization and new retention parameters

Piera Iuliani, Giuseppe Carlucci*, Alessandro Marrone

Dipartimento di Scienze del Farmaco, Facoltà di Farmacia-Università degli Studi "G. D'Annunzio" Chieti-Pescara, Via dei Vestini, 66100 Chieti, Italy

ARTICLE INFO

Article history:
Received 27 February 2009
Received in revised form 29 July 2009
Accepted 3 August 2009
Available online 8 August 2009

Keywords:
Experimental design
MLR analysis
Response optimization
Retention parameters
RP-HPLC

ABSTRACT

The optimization of the HPLC procedures and the investigation of the predictive models are generally seen as distinct tasks although they could be both approached by very similar chemometric methods.

In this work we used chemometric approaches to both optimize the separation of a non-steroidal anti-inflammatory drugs (NSAIDs) set by isocratic reversed-phase HPLC and start-up the development of retention predictive models. The screening of responsive variables and the search for an optimal experimental domain provided also an insight into the structure–chromatographic response relationships for the considered compounds which were particularly helpful to derive new retention parameters.

1. Introduction

One of the key concept of chromatographic separation is the dependence of retention on the structural features of the analytes [1]. In an attempt to give rational interpretation to this concept, quantitative structure–retention relationships (QSRRs) started to be targeted as equations that encode the relationships between the chromatographic response parameters and descriptors of the analyte molecular structure [2]. Several approaches have been proposed to predict the solute retention on selected chromatographic conditions. Among the most diffuse methods, the linear solvation energy relationship (LSER) developed by Abraham [3,4] relates the retention parameter log k with solvent–solute interactions through the so-called solvatochromic equation. Baczek [5] and Kaliszan [5–7] have shown that predictive models of retention in RP-HPLC gradient conditions can be also targeted by t_R as the retention parameter.

On the other hand, the HPLC response optimization, gradually coming out and entering the state-of-art of liquid chromatography [4], can be also included within the efforts to elucidate of the structure–retention relationships. Indeed, response optimization is oriented to describe the HPLC performance as a function of global experimental variables such as mobile phase composition, flow rate, temperature and pH. Typically, the HPLC performances are measured through response parameters [8–10] which can be regressed against the most important experimental variables.

Both QSRR investigation and response optimization can be essentially considered attempts to understand the dependence of retention on either analyte-intrinsic or global experimental factors. Moreover, the optimization of HPLC response and the investigation of QSRRs often share the use of similar regressive methods, filtering the most responsive variables and the need of effective data sourcing [11–13]. It is thus intriguing to guess whether it is possible to exploit the results of the HPLC optimization, i.e. equations correlating response parameters to global experimental variables, to improve the investigation of structure–retention relationships.

In the present work we approached the development of predictive models for the RP-HPLC retention of non-steroidal anti-inflammatory drugs (NSAIDs) starting from the results of HPLC response optimization. In the HPLC method development, the most important aspect is to achieve adequate separation in reasonable time. With the respect to the high number of factors influencing the separation, it could be difficult and time-consuming to reach optimal separation conditions using the single variable optimization approach, when one variable is changed in time while the others are kept constant. The use of chemometric approaches allows to obtain the combination of the variables which provides the best analytical response with a limited number of experiments [1,11]. In the chemometric approach each parameter can be examined and optimized in a predefined range by conducting a series of exper-

* Corresponding author. Tel.: +39 08713554588; fax: +39 08713554911.
E-mail address: g.carlucci@unich.it (G. Carlucci).

0731-7085/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2009.08.001

iments in which the values for several parameters are changed at the same time. Two major groups of experimental design are important: screening and optimization designs [14]. In the case of detailed modeling it is often desirable at a first stage to reduce the number of factors via screening designs to a smaller number of main factors that are to be studied in detail (employing optimization designs) for which both squared and interaction terms in the model are of interest [15].

Fractional factorial design (FFD) approaches are particularly effective and versatile to reduce the number of trial experiments in the development of HPLC procedures [12,13,16–18]. These methods rely on the definition of a response function which is used to score the HPLC performances and is thus the real function to be optimized. However, there is not a unique way to estimate the HPLC separation performance and several suitable response functions can be found in the literature [8–10]. Multivariate regression (MLR) analyses are then carried out to derive analytical expressions of the selected HPLC response parameter as a function of the most responsive experimental variables. The FFD-MLR combination has been successfully used in many HPLC applications, resulting generally reliable and effective [19–21].

As such the final expression of the HPLC response derived by chemometric approaches does encode to a certain extent the structure–retention relationships of the analytes. So it should be possible to exploit the results of the HPLC analysis chemometric optimization to derive quantitative structure–retention relationships (QSRR) allowing to draw predictive models.

In this work, the RP-HPLC separation of the seven NSAIDs was optimized through the use of fractional factorial design (FFD) and central composite design (CCD) combined with multivariate regression analysis (MLR). The selected analytes, although not representing a real analytical problem, are characterized by a moderate molecular similarity (aryl acetic or propionic acids with a good structural diversity in the aromatic moiety) which is expected to be particularly appropriate to test QSRR-based predictive models.

We firstly showed the usefulness of *Y* response function, [19,21] originally designed to be sensitive to both the position and the shape of each HPLC signal. This study suggested that a retention time-dependent function *Y*, reflecting both the analyte run times (peak positions) and the accumulation of diffusive effects (peak widths), can be used both to assess the best operative conditions for the RP-HPLC separation and to correlate the observed HPLC responses with the chemical structure of the NSAIDs considered, deriving a new retention parameter no more dependent on peak shape features.

2. Experimental

2.1. Chemicals

The non-steroidal anti-inflammatory drugs ketoprofen, fenbufen, ibuprofen, furprofen, indometacyn, indoprofen, flurbiprofen, were purchased from Sigma–Aldrich (St. Louis, MO). HPLC grade acetonitrile was provided by Carlo Erba Reagenti (Milan, Italy). Potassium hydrogenphosphate and potassium dihydrogenphosphate, ortophosphoric acid, citric acid were of analytical quality and supplied by Fluka-Riedel-de Haën (Buchs, Switzerland). Water HPLC grade was obtained by passage through an Elix 3 and Milli-Q Academic water purification system (Millipore, Bedford, MA, USA).

2.2. Analytical instrumentation and methods

The HPLC analyses were performed using a Thermofinnigan (Thermofinnigan, St. Josè, CA, USA) system composed a model P2000 pump and a model UV6000 LP diode array detector. A model 7125 sample injector (Rheodyne, Cotati, CA, USA) equipped with a 20 μL loop was used. The analyses were performed on an analytical reversed-phase Phenomenex Hyperclone ODS (250 mm × 4.6 mm i.d, 5 μm particle size) column equipped with a Phenomenex security guard C_{18} (4 mm × 3.0 mm i.d.) precolumn (Phenomenex, Torrance, CA, USA). The column was thermostated in a Igloo-Cil column heater (CIL Cluzeau Info Labo, France). The chromatograms were recorded by means of a computer and were treated with the aid of the software Xcalibur 1.2 from Thermofinnigan. The pH of the mobile phase was measured with a Fischer Scientific Accumet pH meter 15 (Denver Instrument, USA). The pH meter was calibrated with buffer solutions of pH 2.00, 3.00 and 4.00.

The composition of the mobile phase, as well as its pH, were changed in accordance with the optimization design (FFD and CCD, see Tables 1a and 2a F). The temperature of the column was also included in the FFD, while in CCD it was kept constant at 35 °C. The buffers used to obtain the desired pH of mobile phase were phosphate buffer 20 mM (FFD) and citrate buffer 20 mM (CCD). All experiments were performed using isocratic eluition at a flow rate of 1 mL min^{-1} in triplicate and in random order to avoid systematic errors.

The phosphate and the citrate buffers, were filtered prior to use through a WCN 0.45 μm filter (Whatmann, Ltd, Maidstone, UK), while acetonitrile was filtered through a Millipore FA 0.5 μm filter. The mobile phases were prepared daily and were degassed using an in-line model SCM 1000 membrane degasser (Thermofinnigan). Column eluate was monitored at 230 nm (scan from 200 to 400 nm).

2.3. Preparation of stock solutions

Stock solutions of drugs (1000 μg mL^{-1}) were prepared in acetonitrile:water (50:50, v/v) and stored at 4 °C. The working solution, consisting of a mixture of the seven NSAIDs, was prepared daily from stock solutions at a final concentration of 100 μg mL^{-1} for each analyte.

Table 1
(a) Full factorial design for the investigation of Φ, pH and *T* HPLC responsivity and (b) the corresponding experimental response values (scaled by 10^{-6}).

Trial	Experimental design			Experimental set-up		
	x_1	x_2	x_3	Φ	pH	*T* (°C)
(a)						
1	0	0	0	50	2.5	25
2	0	0	0	50	2.5	25
3	1	1	1	55	3.0	35
4	1	−1	1	55	2.0	35
5	1	−1	−1	55	2.0	15
6	1	1	−1	55	3.0	15
7	−1	1	1	45	3.0	35
8	−1	−1	1	45	2.0	35
9	−1	−1	−1	45	2.0	15
10	−1	1	−1	45	3.0	15

Trial	*Y*						
	furp	indop	ketop	fenbuf	flurbip	indom	ibup
(b)							
1	2.116	3.858	1.882	1.014	1.400	0.9794	0.5474
2	2.114	3.857	1.876	1.012	1.401	0.9780	0.5469
3	3.130	6.823	2.937	1.335	2.020	2.396	1.586
4	4.028	8.730	3.383	2.298	3.412	3.158	2.076
5	3.123	6.908	2.659	1.838	2.384	2.484	1.564
6	2.492	0.5502	1.966	1.188	1.527	1.610	1.232
7	2.477	4.027	1.307	0.6103	0.6849	0.6236	0.4371
8	1.912	4.820	1.344	0.8513	0.6798	0.6727	0.4685
9	1.631	3.933	1.045	0.5906	0.6272	0.4982	0.3583
10	1.190	2.637	0.7752	0.3324	0.4041	0.4159	0.3008
SD[a]	0.001	0.001	0.004	0.001	0.001	0.001	0.0004

[a] Standard deviation (SD) calculated on trials 1 and 2 (duplicates of central point).

Table 2
(a) Central composite design for the HPLC response optimization and (b) the corresponding experimental response values (scaled by 10^{-6}).

Trial	Experimental design		Experimental set-up	
	x_1	x_2	Φ	pH
(a)				
1	−1	−1	50	2.5
2	+1	−1	60	2.5
3	−1	+1	50	3.5
4	+1	+1	60	3.5
5	−2	0	45	3.0
6	+2	0	65	3.0
7	0	−2	55	2.0
8	0	+2	55	4.0
9	0	0	55	3.0
10	0	0	55	3.0

Trial	Y						
	fur	indo	keto	fenbu	flurbi	indomet	ibu
(b)							
1	9.752	7.178	11.26	3.092	4.914	2.498	0.1705
2	7.532	8.129	10.89	4.735	8.309	5.646	0.6491
3	5.126	4.832	5.709	6.949	1.994	1.321	0.1521
4	6.268	10.00	8.884	5.122	5.264	6.352	0.8625
5	7.706	5.212	6.699	1.809	2.342	1.103	0.06495
6	7.631	11.85	12.71	6.660	11.42	9.342	1.201
7	10.34	9.117	14.43	5.283	9.526	6.156	0.5826
8	5.230	7.905	7.317	2.777	5.010	3.322	0.5542
9	5.876	6.504	7.646	3.066	4.979	2.882	0.3163
10	6.153	6.328	7.680	2.924	4.765	2.640	0.2906

2.4. Experimental designs and statistical analysis

2.4.1. First experimental design

The influence of the organic modifier CH_3CN (Φ, expressed as percentage), the pH of the mobile phase and the in-column temperature (T) on the HPLC response of the investigated mixture of NSAIDs was studied by using a fractional factorial design (FFD) [16–18]. In this way, the number of experiments was kept low based on the assumption that interaction effects between three or more parameters are small compared to main and two-variables interaction effects. Thus, it is possible to select a fraction of the full factorial design and omit several combinations of parameters from experimental plan [14]. The number of experiments in FFD is given by $2^{k-p}+C$, were k is the number of variables, C the number of replicates at center point and p the whole number that indicates how fractionated the experimental design will be. When p is zero, the experimental design is full. In our study, a total of 10 experiments (2^3 of the factorial design plus two replications of the central point to consider the experimental error) were carried out in randomized run order. By using this design, the three experimental variables were tested at two different levels: Φ at 45 and 55 percent, pH at 2.0 and 3.0 and T at 15, and 35 °C, plus two replications of the central point 50, 2.5 and 25 respectively (Table 1a). The selected response variable was defined as

$$Y = \frac{A}{t_R w_h}$$

where t_R is the relative retention time, A is the peak area and w_h is the peak width at half height. The model proposed for the HPLC response variable (Y) at this stage of the analysis was

$$Y = a_0 + a_1\Phi + a_2 pH + a_3 T + a_{12}\Phi \cdot pH + a_{13}\Phi \cdot T + a_{23} pH \cdot T$$

where a_0 is the intercept, a_1, a_2 and a_3 are the linear coefficients, a_{12}, a_{13} and a_{23} the interaction coefficients.

The parameters of the model were estimated by multiple linear regression (MLR) analysis. Initially, the whole seven-parameters model was fitted to gain preliminary information of the statistical relevance of the considered variables. In all but ketoprofen case (see below), the seven parameter models showed low statistical significance, the regression coefficients being affected by $p > 0.1$. Subsequently, a mixed stepwise-greedy algorithm (MS-GA) was used to filter out non-relevant variables and improve the statistical significance of regression models. The MS-GA decomposed the fitting problem into two tasks in which the linear terms by one side and the cross-terms by the other were separately regressed. In this case we should consider:

$$Y = a_0 + a_1\Phi + a_2 pH + a_3 T \quad \text{and}$$

$$Y = a_0 + a_{12}\Phi \cdot pH + a_{13}\Phi \cdot T + a_{23} pH \cdot T$$

Being essentially a double evaluation of three-variables models, we could fit the two complete models, the six one-variable models or all the possible two-variables models, in total 14 fitting procedures per task. After the separate evaluation of the linear and the cross-products models (the stepwise phase), the best model was obtained by combining the most performing partial model in terms of both correlation and statistic relevance (the greedy phase). To better estimate the responsivity of each term and to also reduce the effects due to the different physical units in the regressed model, the same models were regressed after conversion of the experimental variable values in the corresponding coded values through the general equations:

$$c_i = \frac{x_i - \overline{x_i}}{(1/2)(\max(x_i) - \min(x_i))} \quad i\text{-th independent variable;}$$

$$c_{ij} = \frac{x_{ij} - \overline{x_{ij}}}{(1/2)(\max(x_{ij}) - \min(x_{ij}))}$$

product of i-th and j-th independent variables ($i \neq j$);

where the difference between the actual value of each experimental variables to the corresponding mean value was divided by the half of the variable range.

The regression coefficients of the six models recalculated on the coded variables were straightforwardly used to compare the responsivity of each equation term. The linear and cross-terms not significantly different from zero ($p > 0.05$) were excluded from the model, and the mathematical model was refitted by MLR. The goodness of fit and the overall statistical significance of the model were evaluated by the coefficient of determination (R^2) and F-Fisher (F), respectively.

2.4.2. Second experimental design

From the results of the FFD, a central composite design (CCD) was built using only the variables Φ and pH, while temperature was kept fixed at the value of 35 °C [16–18]. The CCD was built from the FFD 2^k to which star design was added. The length of the arms of the star determined the number of levels and the shape of the experimental design. The CCD was completed by addition of a center point. The total number N of experiments with k factors is: $N = 2^k + 2k + c$. The first term is related to the full factorial design, the second to the star points and the third to the center point. The length of the arms of the star (α) played a major role for the appearance of the CCD. If $\alpha \neq 1$, each variable will assume five levels ($-\alpha$, -1, 0, +1, $+\alpha$) [22]. We tested the two experimental variables at five levels: the CH_3CN fraction at 45, 50, 55, 60 and 65 percent, the pH at 2.0, 2.5, 3.0, 3.5 and 4.0. A total of 10 experiments (2^2 points of the factorial design, 2×2 for the star points plus 2 repetitions of the central point to consider the experimental errors) were carried out in randomized run order (Table 2a). The HPLC response function Y was used also in this case, but the proposed model was quadratic:

$$Y = b_0 + b_1\Phi + b_2 pH + b_{12}\Phi \cdot pH + b_{11}\Phi^2 + b_{22} pH^2$$

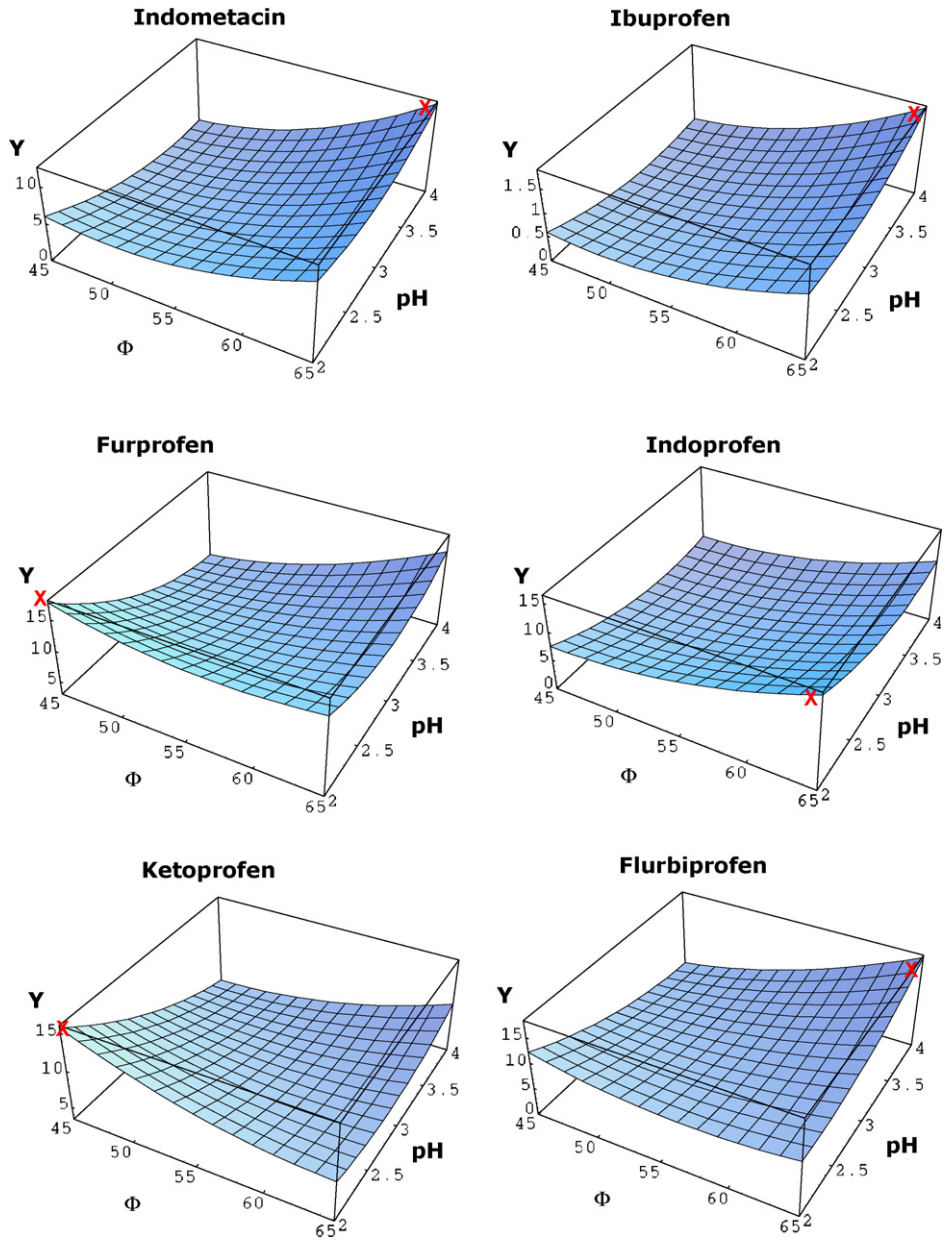

Fig. 1. Calculated HPLC response surfaces for six of the seven considered NSAID. Response values are scaled by a factor of 10^{-6}.

where b_0 is the intercept, b_1 and b_2 are the linear coefficients, b_{12} is the interaction coefficients and b_{11} and b_{22} are the quadratic coefficients. Surface plots (Fig. 1) were developed using the fitted quadratic polynomial equation and were used to locate the points of maximum HPLC response for each analyte in the considered domain [23]. The optimal conditions were obtained by maximizing the HPLC response under adequate constraints, warranting an upper limit for the analysis duration and a lower limit for the peak resolution (vide infra). The statistical analysis was performed by using essentially the same methodologies employed for the screening of variables, even though the higher rank of data matrix allowed the use of direct MLR approaches rather than a MS-GA-based one. The CCD experimental domain was eventually exploited to investigate the correlation between $1/t_R$ and A/w_h, whose product gives the Y response. This led us to derive a new response variable, only dependent on the retention time and not explicitly dependent on the peak shape features. MLR models were subsequently calculated by regressing Y' against Φ and pH, and their predictive capability tested and compared with other retention parameters, such as $\log k$ and t_R.

2.4.3. Statistical analysis

The experimental data processing required for the HPLC response calculation, the fitting procedures and the corresponding statistical analysis were performed by using the Microsoft Excel 2003 software (Microsoft Corporation, Redmond, WA, USA). The linear and multilinear regressions and the corresponding analysis of variance were performed by using the statistical analysis module implemented in Excel.

For the sake of clarity, only the R^2 and F values will be reported in the following beside the values of the regressed coefficients.

3. Results and discussion

The present investigation aims at exploiting the information about the HPLC response of a set of NSAIDs provided by fractional

design to develop retention time predictive models. Our view is that it is possible to take advantage of response optimization to improve and support the development of predictive models. This can be better explained by the following work outline:

(1) Screening of responsive variables ⇒ reducing the experimental space.
(2) Optimization of HPLC response ⇒ information about structure–retention relationships.
(3) Deriving a new retention parameter (Y') from Y ⇒ preserving linear dependence on the experimental variables.
(4) Y'-based models ⇒ prediction of HPLC retention time.

The first two phases of the work are used to provide expressions of HPLC response of single analytes in terms of experimental variables. In the next two phases, a new response parameter is derived by warranting that (i) it depends on retention time only, (ii) it preserves the same HPLC experimental variables asymptotics of Y. The above single phases will be discussed in the next.

3.1. Screening of responsive variables: first experimental design

The screening of variable responsivity was performed by considering the impact of different experimental set-ups on the quality of the HPLC response as measured by the Y function. This response parameter takes into account two factors, one responsive of the time course of the analyte and the other of the shape of the corresponding chromatographic peak; the response of the HPLC signal increases as the corresponding Y value increases.

According to previous investigation [10,21], the Y response analysis was performed by means of a central composite design (CCD) of a three dimensional experimental domain. Table 3a shows the results of the MLR analysis in which the Y response function is regressed against the three selected independent variables: the CH_3CN percentage (Φ), the pH of the mobile phase (pH), and the temperature (T). These results were obtained by employing the raw values of the three selected variables reported in Table 1a (Experimental set-up). The statistical significance of the seven-parameters MLR models, even though resulted in some case with high R^2, were generally low for the most of the analytes (lower F ratio compared to the corresponding MS-GA models) because the associated regression coefficients were often affected (data not shown) by $p > 0.1$. Ketoprofen was the only case in which the direct MLR led to a statistical significant model with seven parameters. The MS-GA procedure was thus employed to filter out non relevant terms from the models and to lead to statistically significant models in which all regression coefficients were affected by $p < 0.01$. As shown in Table 3a, the regressed models obtained through the MS-GA procedure were all improved with respect to the corresponding seven parameter models. No regressed model, either by direct MLR or by MS-GA, was eventually found for indoprofen whereas acceptable correlation was detected in the remaining six cases, R^2 being in the range 0.78–0.99. The overall statistic significance of the regression, estimated as the Fischer ratio, was quite low in only two cases, furprofen and indometacin, 16 and 18, respectively. It is likely that the low significance of the above models could be due to either neglecting higher order terms in the model, likewise quadratic terms. Nevertheless, we assumed the six models with acceptable correlation to screen the responsivity of the three variables and to analyze the impact of each of the three independent variables on the response, discarding those less responsive in the next stages of the investigation. At this purpose, the six independent variables were converted into the corresponding coded values (Section 2.4). This procedure is widely used to eliminate the bias due to the different units of raw data. This bias could affect the analysis of variables responsivity because the regression coefficients do not quote in response unit the contribution of each equation term. For instance, the large and negative values the a_0 coefficients would have induced to conclude that a large extent of response was not encoded by the selected experimental variables. In fact, a_0 can be expressed by

$$a_0 = \bar{Y} - \sum_{i=1}^{3} a_i \overline{x_i} - \sum_{i \neq j=1}^{3} a_{ij} \overline{x_{ij}}$$

and it can be easily negative if the unit scales of the dependent and independent variables are very different.

The six models regressed against the coded variables are reported in Table 3b. As a consequence of coding, the a_0 terms in these models turn to be equal to the corresponding mean of Y, while the remaining terms encode for the amount of response corresponding to the deviation from the mean of Y.

As shown in Table 3b, Φ and pH resulted to be positive responsive variables with coefficients in the ranges 0.5–0.7 and 0.8–2.4 respectively (with the exception of −0.3 for Fenbufen); the temperature was slightly less responsive with coefficients within 0.1–0.4. It is worth noting that in three out six models, and precisely in those for ketoprofen, flurbiprofen and indometacin, a_1 was estimated to be zero indicating that the HPLC response is not linearly correlated to Φ in these cases. The regression coefficients of product terms a_{23} and a_{13} were calculated to be all zero, including the model for ketoprofen, thus indicating that the cross-dependence between Φ and temperature can be neglected.

On the other hand, the coefficients a_{12} were found large and negative in all but furprofen and ibuprofen models. The latter result is not easy to interpret, but we noticed that a_{12} is included in the model when a_1 is not or, as in the case of fenbufen, a_1 is present but a_2 is negative.

These observations led us to hypothesize that a_{12}, in the models for ketoprofen, flurbiprofen and indometacin, encodes for the negative dependence of the HPLC response on increasing Φ. On the other hand, in the model for fenbufen a_{12} is small and parallels the negative response dependence on pH ($a_2 < 0$).

The analysis of response showed that both T and pH are positive responsive variables, hence higher values of Y are expected upon increasing these variables, whereas the remaining variable Φ can be positive or slightly negative responsive depending on the considered analyte. The new domain was thus obtained by shifting up the central point of the experimental design with respect to pH, namely at pH 3.0 instead of 2.5. Therefore, the dependence of the HPLC response and, more specifically, of the chromatographic retention on T is expected to result from the balance of a large number of events. We thus kept this parameter constant at its most responsive value of 35 °C in the next, to warrant by one side the higher response due to this variable and by the other to simplify the rationalization of the response-variables relationships. The fraction of organic modifier was the second variable included in the second experimental domain. According to the standard HPLC theory, the retention time of each analytes is expected to decrease upon increasing Φ; the lack of such an effect in 3/6 models calculated in the first experimental design could be due to the absence of quadratic terms. In this respect we eventually enrolled Φ as the second variable in the new domain and opted to shift up its central value from 50 to 55.

3.2. Optimization of HPLC response: second experimental design

On the basis of the results of the screening of variable responsivity, a second experimental design was built up to search for the Φ and pH values maximizing the Y function. Table 4 shows the results of the MLR analysis giving the Y response as a function of these two

Table 3
Fitted models of the HPLC response against (a) raw and (b) coded values of Φ, pH and T and their cross-products. The regression coefficients (scaled by 10^{-6}) reported in both (a) and (b) refer to the model with the highest F value (bold).

Coefficients		furp	ketop	fenbuf	flurbip	indom	ibup
(a)							
a_0		−5.504	−7.8079	−10.23	−18.96	−16.79	−5.205
a_1		0.1391	0.2119	0.2459	0.4276	0.3740	0.1223
a_2		–	1.420	2.254	4.461	3.319	–
a_3		0.03889	−0.1064	0.01432	0.02317	0.02301	–
a_{12}		–	−0.04163	−0.05563	−0.1016	−0.07522	–
a_{23}		–	0.01200	–	–	–	–
a_{13}		–	0.002160	–	–	–	–
MLR	R^2	0.886	**0.998**	0.967	0.962	0.858	0.727
	F	13	**867**	41	34	10	5
MS-GA	R^2	**0.823**	–	**0.974**	**0.940**	**0.884**	**0.784**
	F	**16**	–	**84**	**36**	**18**	**34**

Coefficients	furp	ketop	fenbuf	flurbip	indom	ibup
(b)						
a_0	2.422	1.917	1.107	1.454	1.382	0.9117
a_1	0.6953	0	0.5342	0	0	0.6117
a_2	–	0.8600	−0.2639	2.230	1.660	–
a_3	0.3889	0.3157	0.1432	0.2317	0.2301	–
a_{12}	–	−1.561	−0.1391	−3.808	−2.821	–
a_{23}	–	0	–	–	–	–
a_{13}	–	0	–	–	–	–

variables. A first effect of changing the experimental domain can be appreciated through the comparison of Tables 1b and 2b indicating an overall increase of HPLC response affecting all the considered analytes. The regression of Y in the 2D domain was performed by extending the model equations to include the quadratic terms and by using the central composite design (CCD).

The reduction of the experimental domain dimensionality from three to two together with the use of the CCD induced a higher rank of the data matrix and allowed to significantly improve the HPLC response description. The resulting models, compared to those obtained in the previous step, are much more accurate and reliable as indicated by the high degree of correlation, with R^2 values in the range 0.91–0.99. In only one case, namely fenbufen, we did not address any model for the HPLC response; notably this compound does not belong to the class of 'profen' which includes five out of the seven analytes considered in the present work. The other six successfully regressed models are all characterized by a similar profile of importance and significance of the corresponding terms, as shown by the values of the regression coefficients, spreading into similar intervals among the several analytes. It is worth noting that, according to previous consideration, both Φ and pH are positive responsive variables in the quadratic terms and negative responsive in the linear terms for all the calculated models.

Both the higher correlation and the similar responsivity pattern of the equation terms are thought to be meaningful breakthroughs of the investigation. Fig. 1 shows the estimated surface plots corresponding to the regressed expressions of the Y variable. As can be seen, an increase of Y occurs when approaching the edges of the experimental domain and, more precisely, at the extremes of the allowed variable range. Such an effect is probably explained by the fact that is asymptotic to $1/t_R$ whose value tends to increase rapidly as t_R decreases. Factors enhancing the mobile phase affinity of a compound are expected to reduce the corresponding retention time.

Table 4
Fitted models of the HPLC response against Φ and pH together with the corresponding correlation (R^2) and statistical significance (F). Regression coefficients (scaled by 10^{-6}).

	furp	ketop	fenbuf	flurbip	indom	ibup
b_0	133.5	141.8	151.9	78.89	105.7	13.17
b_{11}	0.01569	0.02095	0.01980	0.02250	0.02438	0.003276
b_{22}	1.688	2.076	3.151	2.639	1.955	0.2631
b_1	−2.755	−3.247	−2.996	−2.024	−2.836	−0.3722
b_2	−31.30	−36.13	−42.07	−17.65	−23.11	−2.830
b_{12}	0.3362	0.4217	0.3550	−0.01253	0.1883	0.02317
R^2	0.960	0.992	0.944	0.917	0.965	0.982
F	44	211	31	21	51	99

The two considered independent variables affect differently the mobile phase affinity of the analytes. Let consider separately the effect of pH and Φ onto the retention. The pH control is essentially enrolled when ionizable functional groups, such as COOH, NH_2, etc, are present in the analyzed compounds. The analyte in the corresponding ionized form would be less retained by the lipophilic stationary phase. The maximum effect of pH is thus observed in proximity of the pK_a value: for acidic compounds when pH > pK_a and for basic compounds when pH < pK_a. On the other hand, the amount of organic modifier affects the analyte retention time by tuning the polarity of the elution buffer, hence tuning the properties of the media rather than those of the analyzed molecules.

The curvatures of the response surfaces reflect essentially the retention tuning expected by pH and Φ, the pH dependence is however worth of major attention. The analyzed NSAIDs are all characterized by the presence of an acidic functional group inducing an anionic form at pH > pK_a, so that the higher amounts of carboxylate fraction would enhance the $1/t_R$ contribution to Y. In fact, the partial ionization of the carboxylic function of the NSAIDs is likely to be negligible in the RP-HPLC system considered here because: (i) the pH range is below the pK_as of the analytes considered; (ii) the pK_as themselves are expected to be tuned up by the presence of the organic modifier [21].

To explain the slightly positive slope of the surfaces along pH we are rather inclined to a pH dependence of the mobile phase polarity, which is eventually reflected in the HPLC response. The shape of the response surfaces is mainly determined by the contribution of Φ, hence by modulation of the media polarity and qualitatively correlated to the molecular structure of the analyzed NSAIDs. Indeed, these compounds present also an aryl or biaryl tail whose hydrophilicity depends on the presence and arrangement of heteroatomic functional groups which are involved in the polar interactions with the mobile phase. The strength of the hydrophobic interactions involved in the stationary phase affin-

ity is expected to be proportional to the size of the aromatic tails of these molecules. These moieties are then responsible for the different sensitivity of the NSAIDs considered to the partitioning chromatographic system and they probably induce a higher sensitivity of the analyte retention to Φ rather than pH, related to the hydrophilicity of the NSAIDs. More hydrophilic compounds should be characterized by a higher affinity for the mobile phase which is further enhanced by decreasing Φ. On the other hand, less hydrophilic (more lipophilic) compounds should be characterized by a lower affinity for the mobile phase, further enhanced by increasing Φ. Accordingly, the furprofen and ketoprofen response maxima are at $\Phi = 45$, showing that these compounds are probably more hydrophilic because they are less retained at a low percentage of the organic modifier: these two NSAID are characterized by a greater affinity for the mobile phase. On the other hand, the maximum response for the remaining analytes was detected at $\Phi = 65$, thus indicating that they increase their affinity for the mobile buffer at higher CH_3CN/H_2O ratios, may be because these NSAIDs are less hydrophilic than the formers. Eventually, we stated that all maxima located in all but one case (indoprofen) at pH 4.0 even though a minor curvature of the response surfaces is observed for these NSAIDs along this variable, indicating that HPLC response is weakly dependent on pH.

It may be that the pH responsivity in the considered experimental domain encodes some modulation of the mobile phase polarity, expected to give rise to only slight effects onto the response. The analysis of the response surfaces shows that the maximum location qualitatively correlates with the molecular structure of the analyte giving rise to two groups analytes: (i) alpha-aromatic acetic acids with meta-aryl tails and (ii) alpha-aromatic acetic acids with para-aryl tails. Ketoprofen and furprofen form the (i) group and are very similar molecules, differing by only a bioisosteric CH=CH/O substitution in the corresponding distal rings: both their response maxima are located at the (45; 2.0) point. The other analytes form the (ii) group and are all characterized by maximum response at 65% of CH_3CN, although they show a higher structural variability; indometacin can not be included in either of two groups. It seems remarkable as the response surface analysis could easily provide an insight about the different hydrophilicity of the analyzed molecules which are instead structurally very similar. The Y function seems thus to be a valuable tool for characterizing the structure/response relationships observed in the RP-HPLC analysis of the NSAIDs, because the topology of response surface is sensitive to the polarity of the analytes.

Until now, the response optimization led to two possible combination of Φ and pH, one optimizing the HPLC response for the more hydrophilic (group i) members, the other optimizing the response for the less hydrophilic ones (group ii). Fig. 2b shows the chromatograms corresponding to Φ and pH values maximizing Y for both the (i) and (ii) groups of analytes. Even though corresponding to maximum response, the obtained profiles are affected by severe peak tailing and poor resolution. Such negative features of the HPLC profiles are may be due to the higher contribution of $1/t_R$ in determining the value of Y with respect to peak shape. The major weight of $1/t_R$ together with the lack of any description of the peak overlapping in Y invariably lead to neglect the resolution in the HPLC optimization. Therefore, a single optimizing point rather than two possible maxima of response would be much more significant because the (Φ, pH) pair of values corresponding to global optimal apparatus set-up would be assessed. The optimal HPLC conditions have thus been found by imposing adequate constrains concerning the chromatographic profile to the calculated response surfaces:

$$r_{ij} = \frac{2 \cdot |t_{R,i} - t_{R,j}|}{w_i + w_j} \geq 1.5 \quad \text{and} \quad \max(t_R) \leq 15\,\text{min}$$

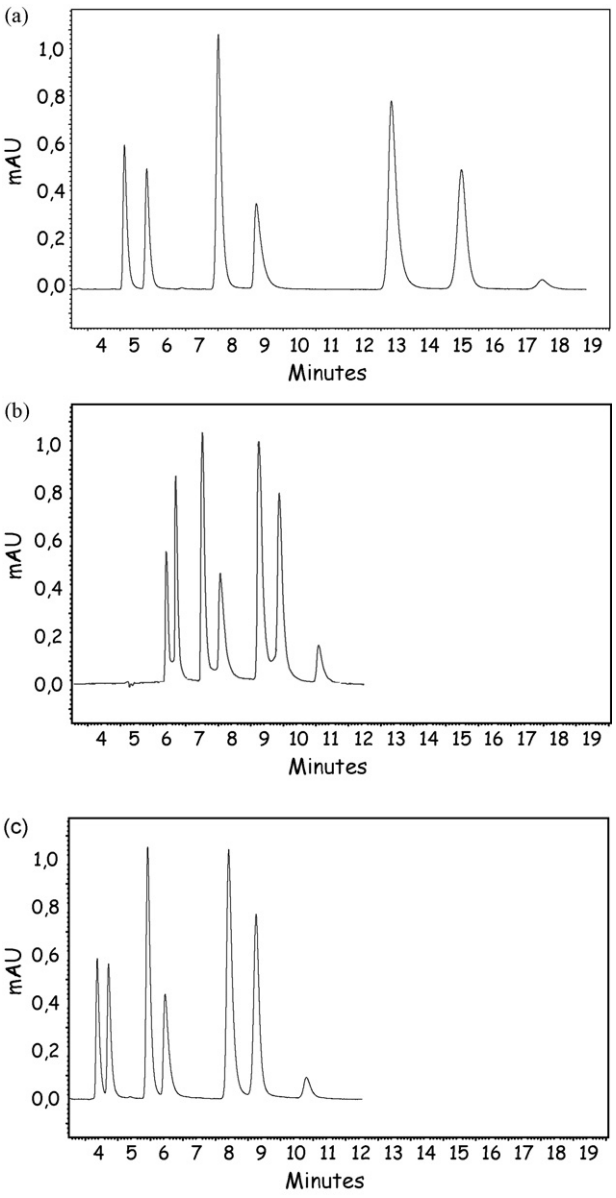

Fig. 2. Chromatograms obtained at (a) the c-CCD central point, (b) the (65.0; 4.0) point of the fc-CCD and (c) the optimal HPLC conditions provided by the maximizing Y under the constraints of adequate resolution and time analysis. For all chromatograms the eluition order was: furprofen, indoprofen, ketoprofen, fenbufen, flurbiprofen, indometacin, ibuprofen, respectively.

where r_{ij} is the pairwise resolution calculated for each ij pair of neighboring peaks, while $\max(t_R)$ is the maximum detected value of t_R in the HPLC chromatogram and is a measure of the HPLC analysis duration. Obviously, these constrains were not imposed analytically to the response surfaces but they were systematically imposed by testing them at each point of the fc-CCD. The overall maximum Y response was then searched among the experimental points satisfying the above constrains.

To locate the best HPLC conditions, the total amount of response at each point satisfying the imposed constrains was calculated. The point corresponding to the maximum total response was detected at $\Phi = 55$ and pH 2.0. Interestingly, the optimal Φ value occurs at the middle of the considered experimental range for this variable thus indicating that a certain compromise between retention and resolution should be achieved to find an overall optimal value. Besides, the increase of peak tailing observed at pH 4.0 could be a possible

explanation for the optimal pH value of 2.0, which is the least. Fig. 2c reports the chromatogram corresponding to the set of NSAID at the optimal experimental set-up. This HPLC profile is notably improved if compared to both that obtained at the center of CCD (Fig. 2a) and those at the conditions maximizing Y (Fig. 2b).

3.3. New HPLC response parameters

The analysis of the HPLC response expressed by the function Y shows that Φ and pH are the most responsive variables that work during the partitioning process of the analytes as tuning parameters of the physicochemical context. The optimization stage provides also model equations to predict the HPLC response of each analyte as a function of the experimental condition (Φ and pH). Since Y is a measure of the peak quality in terms of both position (i.e. retention time, t_R) and shape (i.e. peak area, A, and half-height width, w_h), the model equations derived in the previous stage could be used to score the quality of the HPLC signal of a certain compound in a set, but not to predict its elution ranking.

Typically, the prediction of the peak position is carried out by assuming $\log k$ as the targeted retention parameter and the use of $\log k$ to target HPLC predictive models is frequently reported in the literature [3,4], the neat t_R is also targeted in some other example [5–7].

Here, we used the information about the response gained during the optimization stage to drive the development of predictive models for the NSAIDs considered by introducing a new retention parameter.

The analysis of the response surfaces allowed to optimize the HPLC response and provided an insight into the structure–response relationships for the considered set of analytes. In particular, we observed that the fraction of organic modifier is the most responsive variable, while the responsivity of pH, ranging at low values in the considered domain, is likely to be limited to only a slight modulation of the mobile phase polarity. Moreover, the main contribution to the Y response comes from the retention term, $1/t_R$, which is responsible for the asymptotics of response variable at the edges of the (Φ, pH) domain.

The less retained analytes, furprofen and ketoprofen, being the most hydrophilic molecules of the set, were characterized by a maximum Y value at $\Phi = 45$ corresponding to the most hydrophilic mobile phase composition. On the other hand, the remaining NSAID are less retained at $\Phi = 65$ indicating that they are represented by less polar molecules. The Y response function was thus able to approximately underline the different structure–retention relationships for the analytes considered. In principle, it is reasonable to point out that: (i) it is possible to derive retention parameters expressed in terms of global variables, such as Φ and pH, and analyte-intrinsic variables, encoding the different molecular features of the analyzed NSAIDs; (ii) retention parameters can be derived from Y by separating the most relevant retention time dependence from that of the peak shape; a new function characterized by the same t_R asymptotics of Y can then be extrapolated. The latter point can be outlined as follows:

$$Y = Y(t_R, w_h, A)$$
$$Y' = Y'(t_R)$$
$$Y' \sim Y$$

Firstly, the correlation between $1/w_h$ and $1/t_R$ was investigated at each point of the second fractional design, i.e. by exploiting the CCD data. The results of this analysis are reported in Table 5.

A significant linear correlation between $1/w_h$ and $1/t_R$ was observed at each experimental point. A certain degree of correlation between $1/t_R$ and $1/w_h$ is not surprising, if the expression for the number of theoretical plates in the standard theory of chromatography is considered:

$$N = 5.54 \times \left(\frac{t_R}{w_h}\right)^2$$

Table 5
Correlation between $1/w_h$ and $1/t_R$ for each point of the fractional design (rows 1–9) and for all points (row 10).

Φ	pH	Slope	Intercept	R^2	F
50	2.5	34.61	1.067	0.950	95
60	2.5	23.09	2.236	0.812	22
50	3.5	26.78	1.105	0.929	65
60	3.5	21.92	1.913	0.841	26
45	3	36.89	0.6340	0.979	233
65	3	20.59	2.585	0.752	15
55	2	29.05	1.921	0.949	93
55	4	23.21	1.284	0.840	26
55	3	23.46	1.709	0.880	37
45–65	2–4	25.32	1.630	0.917	321

Less expected is that this linear correlation holds even though the number of theoretical plates N, representing the theoretical extracting units each compound is subjected to during the HPLC course, should be different for each compound. In our view, the observed linear dependence between $1/w_h$ and $1/t_R$ indicates that N assumes approximately similar values in the analyzed set, as a consequence of the high structural similarity therein. This also suggests that under such circumstances the dependence of retention upon the global variables, Φ and pH, is similarly encoded for all the analyzed compounds. Therefore, the contribution of the intrinsic variables to the HPLC retention (i.e. the contribution of structural variables and/or descriptors) should be more easily decoupled from that of the global experimental variables.

In Table 5 the results of regression extended to all the experimental points are also reported. As shown, even if at lower extent, the linear correlation between $1/w_h$ and $1/t_R$ is good in the whole experimental domain allowing, as a first approximation, to assume the all-points regression equation as a general model expressing $1/w_h$ as a function of $1/t_R$. So that, by substituting the all-points regressed equation in the Y definition we obtained:

$$Y = \frac{1}{t_R} \cdot \frac{A}{w_h} = \frac{A}{t_R}\left(m\frac{1}{t_R} + q\right) \quad (1)$$

Eq. (1) discloses a homogeneous quadratic dependence of Y upon $1/t_R$ which led us to assume the very simple hypothesis:

$$Y' = \frac{1}{t_R^2} \quad (2)$$

The prime notation was used to mark that this function is characterized by the same $1/t_R$ asymptotics of Y/A.

3.4. Prediction of HPLC retention

The second stage of the approach was to regress Y' versus the global variables Φ and pH.

Table 6 reports the models obtained by regressing the values of the Y' function calculated by using the CCD data. The resulting description of the HPLC response is notably improved if compared with that obtained at the response optimization stage (see above). On overall, both R^2 and F increase, with R^2 in the range 0.93–0.98 and F in the range of 106–488, indicating a higher accuracy and reliability of the regressed models. Remarkable is also the fact that by using Y' the HPLC response can be expressed through a very simple linear expression of the Φ variable only:

$$\frac{1}{t_R^2} = c_1\Phi - c_0 \quad (3)$$

Table 6
Fitted models of the Y' response against the Φ together with the corresponding correlation (R^2) and statistical significance (F). Coefficients c_1 and c_0 are scaled by a factor of 10^3.

Coefficients	furp	indop	ketop	fenbuf	flurbip	indom	ibup
c_0	2.855	3.056	2.529	2.441	1.781	1.619	1.180
c_1	−96.95	−114.2	−105.4	−107.0	−81.20	−75.37	−54.75
R^2	0.971	0.981	0.984	0.967	0.958	0.930	0.942
F	268	410	488	231	183	107	128

The absence of any term explicitly dependent on pH suggests that the time response of the analyzed NSAIDs is controlled essentially by Φ whereas the pH contribution is approximately constant in the considered experimental domain.

It is confirmed that pH is quite less responsive than Φ, probably because the pK_as of the analytes increase due to the presence of the organic modifier, thus making ionization unlikely in the considered experimental domain [21].

The Y'-derived model of the retention consists of two components corresponding to the $c_1\Phi$ and the c_0 terms of the regressed equations. The chemical information concerning the diverse molecular structure of each analyte is all encoded in the c_1 and c_0 coefficients. The first term is represented by the product between the global variable Φ and the c_1 coefficient which assumes different values per analyte.

The c_1 values qualitatively correlate with the NSAID affinity for the mobile phase; the higher the c_1 values less retained is the compound. Accordingly, the c_1 spread of values allow to discern two groups of compounds: (i) NSAIDs with $c_1 > 2.0$, (ii) NSAIDs with $c_1 < 2.0$. Such grouping scheme is in essentially fair agreement with the Y response optimization results (see previous section) in which a qualitative correlation between the maximum response location and the molecular structure was observed. The only exception is represented by indoprofen whose maximum Y value is located at 65% of organic modifier, while being characterized by the highest c_1 value. The regression intercepts c_0 give rise to negative contributions to Y' assuming again different values per analyte. On the other hand, a certain degree of correlation between c_1 and c_0 might occur, as can be easily deduced from Table 6. Actually, regressing c_1 against c_0 an appreciable linear correlation ($R^2 = 0.8$, slope = 27.53, intercept = 0.02919) is detected.

If we then substitute the above expression of c_0 into the regressed equation of the response function we obtain:

$$\frac{1}{t_R^2} = c_1(\Phi - 27.53) - 0.02919 \quad (4)$$

Eq. (4) can be written in a general form:

$$\frac{1}{t_R^2} = c_1(\Phi - a) - b \quad (5)$$

The resulting equations (5) are now characterized by three parameters: one of these, c_1, is specific of the analyte considered, the others, a and b, are dependent on to the experimental apparatus set-up but not on the analyzed NSAID. The so obtained expression of the time response is very interesting because it resembles the following one:

$$\log k = (\log k)_0 + p(P_m^N - P_s^N)$$

The above expression has been proposed by others [24] to introduce a new HPLC-based polarity scale including the p values. The similarity between the two models suggested the existence of a general expression of the retention parameters, as the product between an analyte-specific polarity and a media-specific polarity. A further analogy between the two models is that both of them describe the contribution of the media polarity as the difference of two parameters, one specific of the mobile phase and the other of the stationary phase. This is clearly expressed in the $\log k$-derived model with the concept of the normalized HPLC polarity of mobile and stationary phase [24]. In the Y'-based model the same assumption can be made assuming the coefficient a in Eq. (5) as the stationary phase polarity. If this should hold, these evidences would support the hypothesis of a general expression correlating the analyte and the media polarity to the retention:

$$retention\_parameter = constant + (analyte\_polarity) \times (media\_polarity)$$

A major difference between the two descriptions of the HPLC time response concerns the media polarity expressions which change appreciably on turning from $\log k$ to Y':

$\log k$-based model

$$media\_polarity = (P_m^N - P_s^N) \propto \left(\frac{(a\Phi^2 + b\Phi + c)}{(d\Phi^2 + e\Phi + f)}\right)$$

Y'-based model

$$media\_polarity = (\Phi - a)$$

As shown, the use of Y' leads to a very simple expression of the media polarity represented by the difference between Φ, encoding the mobile phase polarity, and a, expected to encode the polarity of the stationary phase. On the other hand, the use of $\log k$ would fit the general expression of retention given above by a more complex expression of the media polarity which requires, in particular, many more parameters.

Another difference is that the analyte-specific polarity in the $\log k$-derived model is the tendency of an analyte to be retained during the eluition process (i.e. the affinity with the stationary phase). On the other hand, the analyte-specific index derived with Y', c_1, is related to the mobile phase affinity since a decrease of the retention time occurs upon increasing its value.

Obviously, further investigations should be performed to assess the generality of the proposed models, in particular by analyzing larger sets of analytes, and by testing the approach on several combinations of mobile and stationary phases.

Table 7
Linear regression analysis of the experimental values of $(1/t_R)^2$ against those calculated by using the Eq. (4). Standard error reported in min^{-2} and *min* corresponding to $SE(Y_1')$ and $SE(t_R)$, respectively.

Compound	Slope	Intercept	$SE(Y_1')$	$SE(t_R)$	R^2	F
furp	0.991	−0.0629	0.00285	0.13	0.974	295
indop	0.990	−0.0850	0.00257	0.18	0.981	416
ketop	0.995	−0.0773	0.00180	0.37	0.986	584
fenb	1.06	−0.0854	0.00364	0.42	0.950	152
flurbip	1.00	−0.0538	0.00228	0.96	0.958	185
indom	1.00	−0.0479	0.00271	1.2	0.953	107
ibup	1.00	−0.0272	0.00179	1.2	0.922	131

Eventually, the predictive reliability of the Y' response function was tested for each considered NSAID by regressing the corresponding experimental values of $(1/t_R)^2$ with those calculated from Eq. (4). As reported in Table 7, the proposed models are accurate and reliable, as shown by the high values of correlation coefficients (0.92–0.99) and low values of the standard errors. The standard errors, calculated by comparing calculated and experimental values of t_R, range between 0.13 and 1.2 min and show that the overall predictive accuracy is high.

4. Conclusions

The study of the RP-HPLC separation of a set of NSAIDs showed that response optimization and retention–structure relationships could be investigated by essentially similar chemometric approaches, based upon the combination of fractional design and multilinear regression analysis. The influence of the organic modifier CH_3CN (Φ, expressed as percentage), the pH of the mobile phase and the in-column temperature (T) on the HPLC response of the investigated mixture of NSAIDs was studied by using the Y response function. First the screening of the selected variables was performed through a full fractional design and the impact of each of the three independent variables on the response was evaluated. The HPLC response of the investigated mixture of NSAIDs was then optimized using a central composite design on the domain of the most responsive variables Φ and pH. The surface plots corresponding to the regressed expressions of the Y variable were estimated and used for predicting future responses and optimizing the response. Maximum points on these surfaces have been correlated to the molecular structure of the analytes. Two groups of analyte structures were qualitatively correlated to two distinct maximum response location: (i) alpha-aromatic acetic acids with meta-aryl tails and (ii) alpha-aromatic acetic acids with para-aryl tails. The Y function may thus serve as a probe of the structure/response relationships in the RP-HPLC analysis of these molecules. The optimal HPLC conditions were then assessed as a compromise between retention and resolution to be achieved for optimal separation. The HPLC optimization results have been subsequently exploited to derive new retention parameters. A very simple retention function, Y' (inverse of squared-t_R), makes it possible to obtain reliable predictive models and provides analyte-intrinsic parameters that could potentially be employed in further QSRR investigations.

The first term of these equations is represented by the product between the global variable Φ and the c_1 coefficient which assumes different values for each analyte. The c_1 values are qualitatively correlated with the NSAID affinity for the mobile phase, being higher the less retained is the compound.

This model is similar to a $\log k$-based model reported in the literature [24] and suggests that Y' could be used as an alternative to $\log k$ to develop new HPLC polarity indexes. The c_1 values correlate with the NSAID affinity for the polar mobile phase and allow to rank compounds with an high structural similarity on the basis of their different hydrophilicity. This study suggests that c_1 values could be potentially correlated to molecular structure descriptors of the analytes, thus paving the way to the possibility of developing QSRR models through chemometric approaches. Eventually, the predictive reliability of the Y'-based models was tested and compared to that of other retention parameters, such as $\log k$ and t_R. The model accuracy and reliability resembles that observed for the other parameters considered, suggesting that similar approaches could be adopted to investigate the HPLC time response and to develop predictive models for other classes of compounds.

References

[1] J.G. Dorsey, W.T. Cooper, B.A. Siles, J.P. Foley, H.G. Barth, Liquid chromatography: theory and methodology, Anal. Chem. 68 (1996) 515–568.
[2] R. Kaliszan, QSSR: quantitative structure–(chromatographic) retention relationships, Chem. Rev. 107 (2007) 3212–3246.
[3] J.A. Platts, M.H. Abraham, D. Butina, A. Hersey, Estimation of molecular linear free energy relationship by a group contribution approach, J. Chem. Inf. Comput. Sci. 39 (1999) 835–845.
[4] J.A. Platts, M.H. Abraham, D. Butina, A. Hersey, Estimation of molecular linear free energy relationship by a group contribution approach. 2. Prediction of partition coefficients, J. Chem. Inf. Comput. Sci. 40 (2000) 71–80.
[5] T. Baczek, R. Kaliszan, Combination of linear solvent strength model and quantitative structure–retention relationships as a comprehensive procedure of approximate prediction of retention in gradient liquid chromatography, J. Chromatogr. A 962 (2002) 41–55.
[6] R. Kaliszan, Quantitative structure–retention relationships applied to reversed-phase high-performance liquid chromatography, J. Chromatogr. A 656 (1993) 417–435.
[7] R. Kaliszan, M.A. Van Straten, M. Markuszewski, C.A. Crames, H.A. Claessens, Molecular mechanism of retention in reversed-phase high-performance liquid chromatography and classification of modern stationary phases by using quantitative structure–retention relationships, J. Chromatogr. A 855 (1999) 455–486.
[8] D. Kantoci, A study of an empirical equation for the evaluation of separation efficiency in chromatography, J. Liq. Chromatogr. Relat. Technol. 19 (1996) 853–866.
[9] D. Bylund, A. Bergens, S.P. Jacobsson, Optimization of chromatographic separations by use of a chromatographic response function, empirical modeling, and multivariate analysis, Chromatographia 44 (1997) 74–80.
[10] Y. Guillaume, E.J. Cavalli, E. Peyrin, C. Guinchard, A new approach to study benzodiazepine separation and the differences between a methanol/water and acetonitrile/water mixture on column efficiency in liquid chromatography, J. Liq. Chromatogr. Relat. Technol. 20 (1997) 1741–1756.
[11] J.G. Dorsey, W.T. Cooper, B.A. Siles, J.P. Foley, H.G. Barth, Liquid chromatography: theory and methodology, Anal. Chem. 70 (1998) 591–644.
[12] T.H. Ong, Y.S. Wong, S.O. Woo, S. Ng, Factorial designs for optimization in high-performance liquid chromatography of pharmaceuticals, Bull. Singapore Natl. Inst. Chem. 22 (1994) 19–29.
[13] P. Kaufmann, Multivariate optimisation strategy for liquid chromatography. II. Exploring and finding optimal conditions in the search area of the multidimensional solvent space, Chemom. Intell. Lab. Syst. 27 (1995) 105–114.
[14] M. Preu, D. Guyot, M. Petza, Development of a gas chromatography-mass spectrometry method for the analysis of aminoglicoside antibiotics using experimental design for the optimization of the derivatization reactions, J. Chromatogr. A 818 (1998) 95–108.
[15] R.G. Brereton, Chemometrics-data Analysis for the Laboratory and Chemical Plant, John Wiley & Sons Ltd., Chirchester, England, 2002, pp. 76–77.
[16] W.G. Lan, K.K. Chee, M.K. Wong, H.K. Lee, Y.M. Sin, Orthogonal array design as a chemometric method for the optimization of analytical procedures. Part 5. Three-level design and its application in microwave dissolution of biological samples, Analyst 120 (1995) 281–287.
[17] H.B. Wan, W.G. Lan, M.K. Wong, C.Y. Mok, Orthogonal array designs for the optimization of liquid chromatographic analysis of pesticides, Anal. Chim. Acta 289 (1994) 371–380.
[18] A.M. Andersson, A. Karlsson, M. Josefson, J. Gottfries, Evaluation of mobile phase additives in LC-systems using chemometrics, Chromatographia 38 (1994) 715–722.
[19] Y. Guillaume, C. Guinchard, A new method to study column efficiency and the separation of seven compounds in reversed-phase chromatography, J. Chromatogr. Sci. 33 (1995) 204–210.
[20] E. Marengo, M.C. Gennaro, V. Gianotti, E. Prenesti, Optimization of the separation of mono- and dichloroanilines in ion interaction high-performance liquid chromatography, J. Chromatogr. A 863 (1999) 1–11.
[21] A.B. Baranda, O. Berasaluce, R.M. Jiménez, R.M. Alonso, LC-DAD determination of calcium channel blockers by using an experimental design approach, Chromatographia 61 (2005) 447–453.
[22] M. Novak, A. Seubert, Application of experimental design for the characterization of a novel elution system for high capacity anion chromatography with suppressed conductivity detection, J. Chromatogr. A 855 (1999) 91–109.
[23] G.E.P. Box, W.G. Hunter, J. Stuart Hunter, Statistics for Experimenters—An Introduction to Design, Data Analysis and Model Building, Wiley, New York, 1978.
[24] R. Bosque, J. Sales, E. Bosch, M. Rosés, M.C. Garcia-Alvarez-Coque, J.R. Torres-Lapasiò, A QSPR study of the p solute polarity parameter to estimate retention in HPLC, J. Chem. Inf. Comput. Sci. 43 (2003) 1240–1247.