Available online at www.sciencedirect.com
SCIENCE DIRECT®
Analytica Chimica Acta 528 (2005) 63–76
ANALYTICA CHIMICA ACTA
www.elsevier.com/locate/aca

Multi-objective optimisation strategy based on desirability functions used for chromatographic separation and quantification of L-proline and organic acids in vinegar

M.J. Sáiz-Abajo, J.M. González-Sáiz, Consuelo Pizarro*

Department of Chemistry, Complejo Científico Tecnológico, University of La Rioja, c/Madre de Dios 51, 26004 Logroño, La Rioja, Spain

Received 12 January 2004; received in revised form 4 June 2004; accepted 4 June 2004
Available online 28 July 2004

Abstract

An optimisation strategy based on desirability functions together with experimental design has been used to optimise a chromatographic method applied to the separation and quantification of L-proline and seven organic acids in vinegar samples. Chromatographic problems often force to reach a compromise between different experimental variables in order to achieve the best chromatographic separation. The importance of the use of multi-objective optimisation methods lies in the ability to cope with this kind of problems. The quality of the multicriteria optimisation method was tested through the validation of the analytical parameters of the final chromatographic method developed. The versatility of this methodology allows to use it in other chromatographic applications resulting in a suitable adaptive procedure to solve new analytical problems. Furthermore, the determination of L-proline and organic acids in vinegar is useful for several industrial goals such as the correct monitoring of fermentation, for the study of nutrient needs at all times during the twofold fermentation process and for the detection of possible adulterations in the final product.

Keywords: Multi-objective optimisation; Experimental design; Desirability functions; Vinegar

1. Introduction

High-performance liquid chromatography has resulted to be a well established technique in analytical research, quality control and many other applications in analytical chemistry. Besides, the optimisation of HPLC method development has always been an important objective in many contexts.

The first attempt to solve these kind of problems was based on the existing theoretical advances in chromatography that allowed to predict the experimental response by controlling and modifying the chromatographic parameters [1]. Afterwards, the development of unclear computer-assisted methodologies were suggested as a more sophisticated method with optimisation purposes [2].

It was Cela et al. that proposed finishing with the trial-and-error methodology and the computerised obscure chemometrical tools by using evolutionary algorithms and applying them to HPLC separations [3]. This approach was based in the well-populated Pareto front and used it to allow a direct definition of chromatographer goals in the optimisation process.

Other tools such as genetic algorithms, neural networks and response surface methodology have also been used in liquid chromatography to develop and optimise the separations [4,5] as well as Derringer or desirability functions [6,7]. The latter, seem to be an effective tool for that purpose due to the wide versatility that these functions have to transform each response separately and comprising all them in an overall desirability function to be globally optimised in a final step.

In this study a multi-objective optimisation methodology is applied to HPLC separations based on experimental design and desirability functions. Experimental design is the best way to cover the experimental domain in a given number of experiments and desirability functions allows to take into account multiple responses at the same time. Often, this

* Corresponding author. Tel.: +34-941299626; fax: +34-941299621.
E-mail address: consuelo.pizarro@dq.unirioja.es (C. Pizarro).

0003-2670/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.aca.2004.06.027

requires finding optimal compromises between the experimental responses.

In chromatographic matters, responses such as resolution, symmetry or area are usual experimental parameters that should be considered when optimising a chromatographic procedure and in this research were optimised 20 experimental parameters extracted from the chromatograms. Different importance was given to each response when calculating the overall desirability depending on its importance on the chromatographical separation.

The application of the multicriteria optimisation methodology for the separation and quantification of L-proline and organic acids in vinegar, resides in the importance of these compounds in the quality and genuineness of the vinegar samples apart from the importance in the fermentative procedure. This task is important nowadays when analytical chemistry is focused on the development of new and rapid techniques for the assessment of quality in food industry.

Organic acids, such as L-tartaric, L- and D-malic, lactic, acetic, citric and succinic acid, account for a significant fraction of wines and vinegars that are made of them [8]. The level and nature of these acids provide useful information about the origin of the raw material, microbiological growth and process techniques [9]. Moreover, they can contribute significantly to the taste of the product.

Proline is a primary α-amino acid that accounts for 78% of the total amino acid content in wine [10]. It is the most abundant amino acid in vinegar and comes from the raw materials. The control of the concentration of L-proline during the fermentation is quite important because, together with L-leucine, provide between the 50 and 74% of the total amount of the nitrogen consumed during the alcoholic and the acetic fermentations. The lack of proline during the elaboration of vinegar could result in a premature stop of fermentation. It is also known that proline provides information of possible mixtures done during the production of wine vinegar. The amount of proline in the different raw materials makes possible the detection of adulteration in the final product. Wine has a big amount of L-proline whereas other materials such as cider, malt or molasses have a little amount of it. Proline isomeration has been also used to determine ageing processes in wine and vinegar that should be evaluated through the proportion of L- and D-proline in the sample [11].

Many methods of analysis of organic acids and proline have been described in the literature [8–15]. The collection of high quality chromatograms has traditionally been a difficult task because of the acid nature of these compounds, and the necessity of a previous derivatisation such as in the case of the L-proline. The importance of controlling the mobile phase pH when analysing ionizable compounds such as organic acids by reverse phase HPLC is recognized and understood. Parameters such as pH and percentage of organic modifier in the mobile phase influence the retention time and symmetry of the peaks. Therefore, buffered mobile phases are commonly used in reverse phase chromatography. Usually, the determination of amino acids requires pre-column derivatisation with *o*-phthaladehyde-3-mercaptopropionic acid (OPA) or 9-fluorenylmethyl chloroformate (FMOC) depending on whether they are primary or secondary amino acids [16,17]. This derivatisation step has the objective of improving UV or fluorimetric detection but it means an extra previous stage in the analytical process that can be avoided in this study. This work tries to obtain a chromatographic method that can supply chromatograms of high quality in terms of resolution and peak symmetry, so that they can be used, together with other chromatographic data, as spectroscopic profiles in future research works as well as for performing new optimal chromatographic separations in vinegar or other matrices.

The main aim of this study is to optimise the chromatographic analysis of L-proline and organic acids at the same time in vinegar and evaluating the variables involved in the process in order to reach the best separation method. The length of the analysis is also remarkable because vinegar samples are directly injected in the HPLC system after being previously filtered to avoid other pre-treatment processes that may lead to a lack of reproducibility and are quite time-consuming.

2. Experimental

2.1. Chemicals and reagents

L(+)-tartaric, L(−)-malic, lactic, acetic, succinic, citric acids and L-proline of analytical-reagent grade were supplied by Merck (Darmstadt, Germany). For the preparation of the mobile phase, potassium di-hydrogen phosphate of analytical-reagent grade from Panreac (Barcelona, Spain), methanol of gradient grade and 85% phosphoric acid supplied by Merck were used. Water purified using a Milli-Q system académic A10 (Millipore S.A., Molsheim, France) was used.

2.2. Equipment and software

The equipment was composed of an HP 1100 Series liquid chromatograph (Hewlett-Packard GmbH, Chemical Analysis Group Europe, Waldbronn, Germany) with a high pressure gradient pump, vacuum degasser, autosampler, thermostated column compartment, diode array detector and an HP Chemstation data processing system (Hewlett-Packard) to perform peak purity analyses.

The data were processed with Statgraphics Statistical Computer Package [18] and Nemrod-W [19] for the experimental design.

2.3. Column and mobile phase

The column used was Zorbax SB-C18, 250 mm × 4.6 mm ID with 5 μm particle size (Hewlett-Packard GmbH,

Waldbronn, Germany). The stable bond packaging is suitable for working at low pH values and prevents tailing. The operating temperature was 25 °C.

The mobile phase was 0.009 M potassium di-hydrogen phosphate (adjusted to pH 2.06 with phosphoric acid)–methanol (92:8, v/v) at a flow rate of 0.64 ml min^{-1} at 25 °C and a working pressure of 90 bar (1 bar = 10^5 Pa). Detection was performed by measuring of the UV absorption at 210 nm. A previous multi-wavelength detection study (190–280 nm) was done in order to select the optimal absorbance wavelength.

2.4. Vinegar samples and sample preparation

Sixty-three vinegar samples of different origins were collected from the industry and several supermarkets from the north of Spain. From the total amount of samples, 33 were white wine vinegar samples, 13 red wine vinegar samples, 2 alcohol vinegar samples, 3 malt vinegar samples and 11 cider vinegar samples. A mixture of white wine vinegar and alcohol vinegar was made in the laboratory to evaluate the differences between this sample and the original samples. This adulterated sample contained 40% (v/v) of alcohol vinegar and 60% (v/v) of wine vinegar which represents a realistic fraudulent vinegar sample.

All the solutions and vinegar samples were passed through a 0.7 μm glass microfibre GF/F from Whatman (Whatman International Ltd., Maidstone, England), before injection into the HPLC system.

2.5. Experimental design

Experimental design is often used to determine efficiently the set of conditions required in a process or method with desirable, often optimal, characteristics with a minimum of experiments. However, modelling is one of the main techniques applied to optimise responses [20,21]. Multi-level designs for quantitative factors are used to evaluate several experimental factors at a time defining quadratic response surfaces.

In this study four experimental factors were selected as the main variables involved in the chromatographic separation and quality of the chromatogram:

(i) Flow rate, which influences the peaks shape.
(ii) pH, which influences the retention time of the peaks and the ionisation state of the compounds that affects the sensitivity of the analysis.
(iii) Phosphate concentration in mobile phase, which influences the peaks separation.
(iv) Percentage (v/v) of organic modifier (methanol) in the mobile phase, which influences also the retention time.

Therefore, a Doehlert design was chosen. Doehlert uniform shell design is a very useful design introduced by Doehlert [22]. It describes a spherical experimental domain, but with fewer points than the central composite design and it stresses uniformity in space filling. Doehlert designs are much more efficient than central composite designs where efficiency is defined as the number of estimated *b*-coefficients divided by the number of experiments [23,24].

This design involves 21 experiments randomly performed to provide protection against the effects of lurking variables. Some replicas were made at the central point of the experimental space in order to validate the model and the final number of experiments was 25. Previous experiments and the experience in the use of this technique led us to select the most relevant parameters involved in the quality of the chromatogram. The factors were studied at different levels depending on their individual impact on the system. The selection of the extreme levels of the factors in the experimental design was done considering the information given by previous experiments, the guidance given by the manufacturer and the technical limitations of the system. The flow rate was studied at five levels, the pH and percentage of methanol were studied at seven levels and the concentration of phosphate was studied at three levels. The values corresponding to the high (+), and low (−) points for each factor that define the experimental domain and the different levels of study are shown in Table 1 whereas the experimental schedule is shown in Table 2.

The sample volume was set at 10 μL and the temperature at 25 °C. Due to the changes of mobile phase during the experimentation, a time of 120 min was set as equilibration time before performing each experiment if the mobile phase had changed.

2.6. Derringer functions or desirabilities

Multicriteria decision-making is applied when several responses have to be considered at the same time. This usually requires finding optimal compromises between the total number of responses taken into account. Derringer functions or desirabilities are the main tool of one of the most important multicriteria decision-making methods. The multicriteria problem is reduced to a single criterion problem of *D*

Table 1
Factor levels in the Doehlert experimental design

Flow rate (ml min^{-1})	0.40	0.60	0.80	1.00	1.20		
pH	1.77	1.91	2.06	2.20	2.34	2.49	2.63
Methanol percentage (v/v)	0	2	6	8	10	14	16
Potassium di-hydrogen phosphate (M)	0.003	0.010	0.017				

Table 2
Scheme of the experiments

Experiment number	Flow rate (ml min^{-1})	pH	Methanol percentage (v/v)	Phosphate concentration (M)
1	1.20	2.20	8	0.010
2	0.40	2.20	8	0.010
3	1.00	2.63	8	0.010
4	0.60	1.77	8	0.010
5	1.00	1.77	8	0.010
6	0.60	2.63	8	0.010
7	1.00	2.34	16	0.010
8	0.60	2.06	0	0.010
9	1.00	2.06	0	0.010
10	0.80	2.49	0	0.010
11	0.60	2.34	16	0.010
12	0.80	1.91	16	0.010
13	1.00	2.34	10	0.017
14	0.60	2.06	6	0.003
15	1.00	2.06	6	0.003
16	0.80	2.49	6	0.003
17	0.80	2.20	14	0.003
18	0.60	2.34	10	0.017
19	0.80	1.91	10	0.017
20	0.80	2.20	2	0.017
21	0.80	2.20	8	0.010
22	0.80	2.20	8	0.010
23	0.80	2.20	8	0.010
24	0.80	2.20	8	0.010
25	0.80	2.20	8	0.010

optimisation. Single functions are defined for each criterion to be optimised:

$$d_i = f(y_i) \tag{1}$$

where $0 \leq d_i \leq 1$. All the functions are combined into a single criterion where each single function can be weighted:

$$D_i = \sqrt[C]{d_{1i} d_{2i} \cdots d_{ci}} \tag{2}$$

$$D_i = (d_{1i}^{w1} d_{2i}^{w2} \cdots d_{ci}^{wc})^{1/\sum wk} \tag{3}$$

where $0 \leq D \leq 1$ [20,25], w_k are the weights for each one of the single k weighted functions. Through the individual functions the analyst introduces the specifications that each response must fulfil and through the weighting the relative importance given to each of them. The overall desirability function D is not derivable, so the optimisation process must use a method which is free of derivatives, such as the simplex method, a genetic algorithm or a simulated annealing algorithm which is that used in Nemrod-W [19].

3. Results and discussion

3.1. Optimisation

The response selection is a critical stage in the optimisation. The responses were directly related to the parameters that define the quality of the separation that will allow a correct identification and quantification of the several compounds in the studied matrix, in our case, vinegar.

Three types of responses were chosen:

(i) Resolution, which determines the separation between the peaks and it is directly related with the final time of the analysis.
(ii) Symmetry factor, which determines the existence of front or back tails in the peaks, which affects the afterwards quantification step.
(iii) Area values, which determines the sensitivity of the analysis.

The most common formula to measure the resolution (R_S) is

$$R_S = \frac{2(t_2 - t_1)}{w_1 + w_2} \tag{4}$$

where t_1 and t_2 are the retention times of two consecutive peaks, and w_1 and w_2 are the peak widths measured at the baseline between the tangents drawn next to the peak sides. From a practical point of view, it is much easier to measure the peak width at half the peak height. The resolution equation using the half-height method is

$$R_S = \frac{2(t_2 - t_1)}{1.7(w_{0.5,1} + w_{0.5,2})} = \frac{1.18(t_2 - t_1)}{(w_{0.5,1} + w_{0.5,2})} \tag{5}$$

where $w_{0.5,1}$ and $w_{0.5,2}$ are the peak widths measured at half height. This equation was used in the study. Seven resolution values were calculated between each pair of the eight peaks present in the chromatogram.

Symmetry was calculated using the following equation:

$$\text{Symmetry} = \frac{a_1 + a_2}{a_3 + a_4} \tag{6}$$

where a_1, a_2, a_3 and a_4 are the areas under the peak once its total area has been divided in four parts. Firstly, the peak is divided into two halves down the apex point, and then each half is again divided into other two parts down the upslope and downslope inflexion points of the peak. Letters a_1, a_2, a_3 and a_4 stand for the area values of the first, second, third and fourth parts of the divided peak from left to right side.

Area values of the peaks depend on the UV absorption and the ionisation state of the compound plays an important part in the detection process. Big peak areas may cause overlapping and small peak areas make the correct quantification very difficult.

Our aim was to find the trade-off situation where all the response factors are optimised and take the values that afford the best chromatographic quality. In general, the optimisation of an analytical procedure implies solving a conflict between resolution of peaks, symmetry and signal size.

The sequence of peaks obtained in the chromatogram is: L-proline, tartaric acid, L-malic acid, lactic acid, acetic acid, citric acid, succinic acid and D-malic acid.

Twenty response values were collected from the chromatograms obtained: seven resolution values between each

pair of consecutive peaks (R_{S1}, R_{S2}, R_{S3}, R_{S4}, R_{S5}, R_{S6}, R_{S7}), six peak area values of L-proline, L-tartaric acid, L-malic acid, lactic acid, citric acid, succinic acid; seven peak symmetry values of L-proline, L-tartaric acid, L-malic acid, lactic acid, acetic acid, citric acid and succinic acid. The peak area of acetic was not considered because it is the main acid in vinegar and always shows acceptable area values. The peak area of D-malic acid and the symmetry values were not considered either because L-malic acid is the main enantiomer in vinegar.

As it has been stated, Doehlert experimental design describes a spherical space defined by a quadratic model according to the following equation:

$$Y = b_0 + \sum_i b_i x_i + \sum_i \sum_j b_{ij} x_i x_j \quad (7)$$

A standard solution containing the organic acids and L-proline studied at the concentrations typically found in wine vinegar was injected in triplicate for each run.

From the total amount of responses, there were two that, at a first sight, could seem more relevant to the chromatographic separation than the others: $R1$ and $R4$. $R1$ (R_{S1}) is the resolution value between L-proline and L-tartaric peaks, and $R4$ (R_{S4}), the resolution value between lactic and acetic acids peaks. These two pairs of peaks are the most difficult to resolve. The optimisation of these two critical responses can manage the final optimal conditions to some extent. This is the reason why a further analysis on these two points was carried out.

3.2. Optimisation of R1 and R4

With the aim of the optimisation of these responses (resolution between the peaks of L-proline and L-tartaric acid and between lactic and acetic acid) a second-order approach was used to model the experimental data obtained. The analyses of the variance (ANOVA) allowed to accept that the fitted surface model described satisfactory $R1$ and $R4$ in the experimental domain. The models built are significant and there is not lack of fit at a significance level of 0.1% ($R^2 = 0.972$ and 0.984) for $R1$ and $R4$, respectively.

Looking at the estimates of the coefficients and the statistics of Table 3, it can be done some comments. Taking into account response $R1$, the coefficients b_2 and b_4 of the computed model are very significant together with the coefficients b_{22}, b_{23} and b_{34}. This means that second order interactions between factors are important and should be taken under consideration. Considering the coefficients of the built model for $R4$, only b_2 and b_3 are very significant, in this case, second-order interaction between factors are not significant. The study of the optimum path of the response surface for $R1$ and $R4$ of Fig. 1, demonstrated the necessity of the use of a more complex function that collected the combination of responses. $R1$ is more sensitive to variations in the percentage of methanol and concentration of phosphate in the mobile phase than to changes of the flow rate and pH of the mobile phase. So, if $R1$ must be maximised, the percentage of methanol and the concentration of phosphate must have low values, pH must have values close to the central value of the experimental domain and the flow rate must tend to less positive values. $R4$ is more sensitive to variations in pH and percentage of methanol in the mobile phase than to changes of flow rate and concentration of phosphate in the mobile phase. The percentage of methanol in the mobile phase and pH must have higher values, the concentration of phosphates must have a value close to the central value and the flow rate must tend to less positive values.

The need of a multi-objective optimisation had become clear through the examination of the evolution of the factors in the optimisation of these two responses. $R1$ and $R4$ are

Table 3
Estimates and statistics of the coefficients from the fitted response surface for $R1$ and $R4$

	$R1$			$R4$		
	Coefficient	t (exp.)	Significant %	Coefficient	t (exp.)	Significant %
b_0	0.944	24.69	<0.01***	0.870	62.10	<0.01***
b_1 (flow rate)	−0.092	−2.41	3.53*	−0.010	−0.74	48.0
b_2 (pH)	−0.409	−10.69	<0.01***	0.121	8.67	<0.01***
b_3 (% methanol)	0.040	1.04	32.6	0.320	22.83	<0.01***
b_4 (phosphate)	−0.303	−7.93	<0.01***	0.031	2.19	5.1
b_{11}	0.011	0.15	87.5	−0.010	−0.38	71.1
b_{22}	−0.606	−8.47	<0.01***	0.063	2.42	3.50*
b_{33}	0.309	4.60	0.105**	−0.041	−1.66	12.5
b_{44}	−0.095	−1.50	16.2	−0.062	−2.69	2.19*
b_{12}	0.029	0.29	77.2	−0.058	−1.60	13.9
b_{13}	−0.053	−0.48	64.5	−0.029	−0.71	50.2
b_{23}	−0.656	−5.94	0.0174***	−0.002	−0.06	95.3
b_{14}	0.294	2.56	2.72*	0.098	2.33	4.04*
b_{24}	−0.260	−2.27	4.49*	0.046	1.09	30.4
b_{34}	0.765	6.67	<0.01***	−0.030	−0.71	50.2

Signification: * 95%, ** 99% and *** 99.9%.

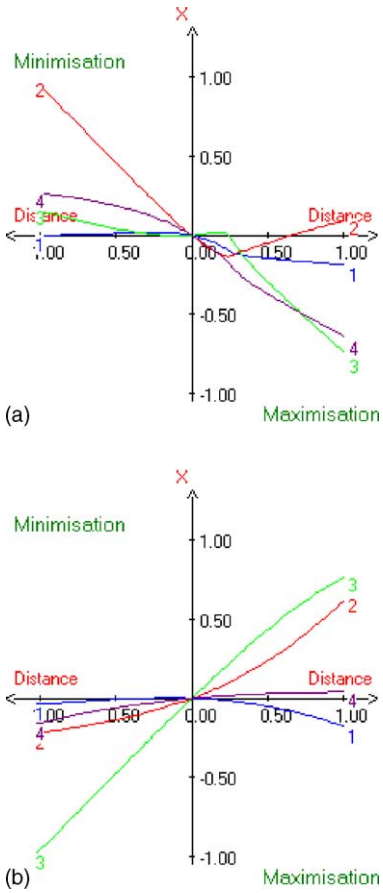


Fig. 1. Coordinates of the points of the plot of the optimum path of the response $R1$ (a) and $R4$ (b) for each factor in codified variables. (1) Flow rate; (2) pH; (3) percentage of methanol and (4) concentration of potassium di-hydrogen phosphate.

important and critical responses but, if we only considered these two responses, some other peaks appeared overlapped, the length of the chromatographic analysis increased and some other parameters as symmetry or area values had not desirable values. Thus, it is important the trade-off between all the chromatographic responses related to the quality of the analytical process. Other responses such as $R5$ (R_{S5}, acetic–citric resolution), $R6$ (R_{S6}, citric–succinic resolution) and the symmetry of some of the peaks ($R14$, symmetry of proline peak or $R18$, symmetry of acetic acid peak) are influenced by the global chromatographic conditions an cannot be left out.

3.3. Derringer functions

In our case, the objective is to achieve a good chromatographic separation with symmetrical peaks of measurable areas to be correctly quantified. Thus, different desirability functions, defined by the following specifications, were chosen:

(i) For the resolution responses, linear functions were defined from $R = 0.5$ (zero desirability), completely unsatisfactory resolution, to $R = 1.5$ (desirability $d_i = 1$). Between these two values the desirability function varied linearly. The higher resolution limit was fixed in order to not to have a too long analysis.

(ii) For the area values, triangular functions were defined; these functions took the highest desirability value ($d_i = 1$) for the mean data of the interval defined by the smallest and the largest area value for each compound. The value of d_i decreased till $d_i = 0.1$ when the values of area took the maximum and minimum values of each compound.

(iii) The symmetry responses were described as triangular functions centered on $S = 1$ (perfectly symmetrical peak) where $d_i = 1$ and 0 for both $S = 0$ and 2 symmetry values for which the peaks were considered completely asymmetrical.

The overall desirability D was obtained by combining single desirability functions according to Eq. (3). The responses $R1$ (resolution of L-proline–tartaric acid peaks) and $R4$ (resolution of lactic–acetic acid peaks) were four times weighted due to their relevance as stated before. The rest of resolution responses ($R2$, $R3$, $R5$, $R6$ and $R7$) were two times weighted and the responses of symmetry ($R14$–$R20$) and area values ($R8$–$R13$) were mono-weighted. Nemrod software was used to calculate the optimal conditions. A value of $D = 0.827$ was obtained after the modelling and optimising stages.

The optimal conditions to perform the separation and quantification of L-proline and organic acids in vinegar to make the overall desirability function take its maximum value were: 0.64 ml min^{-1} flow rate, 2.06 pH, 8% (v/v) methanol percentage and 0.009 M phosphate concentration in the mobile phase; all of them computed by Nemrod-W after the modelling stage. For these conditions, the values of the resolution values ($R1$–$R7$) ranged from 0.85 for $R4$ to 5.16 for $R3$; the area values ($R8$–$R13$) ranged from 143.2 for L-proline to 819.7 for L-tartaric and the symmetry factor ($R14$–$R20$) ranged from 0.305 for acetic acid peak to 1.275 for citric acid peak. All these values were considered acceptable for the final purposes of the study.

The different influence of the four factors on the experimental responses was evaluated going through the values of the coefficients of the model. It can be graphically plotted in the Pareto chart of Fig. 2 that shows the importance of the different variables based on the t-student experimental values that provide Nemrod. Even though single coefficients b_2, b_3 and b_4 are significant as also are the coefficients of the quadratic terms b_{22}, b_{33} and b_{44}, no consideration may be done of neither of them due to the existence of a remarkable second-order effects b_{23}, b_{24} and b_{34}. As these interactions are very significant, pH, methanol percentage and phosphate concentration in the aqueous phase are hardly correlated and its correlation effects should be studied through the evaluation of

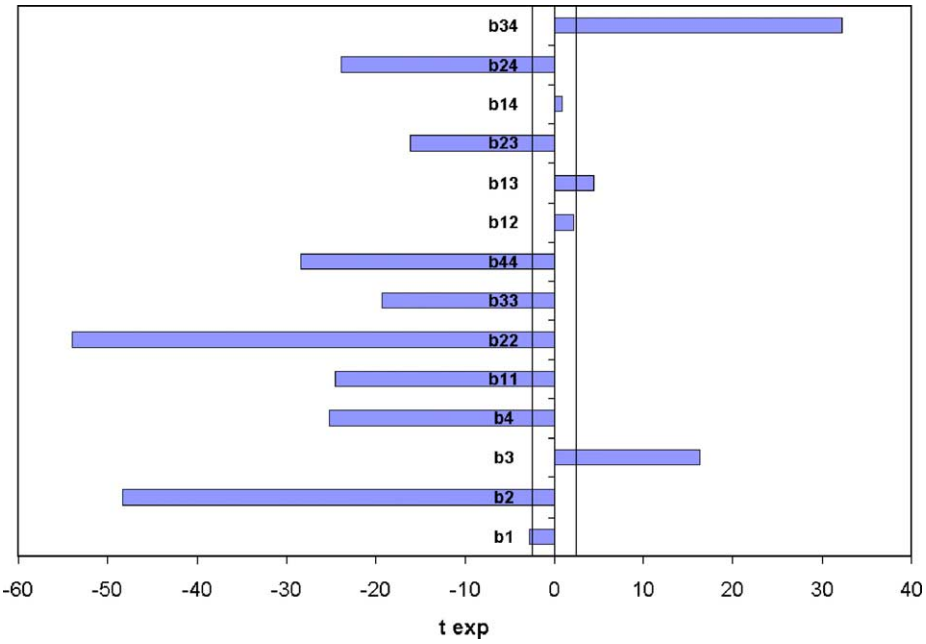


Fig. 2. Pareto chart of main effects of the Doehlert design for global desirability.

the behaviours of a multicriteria response in the experimental domain. Flow rate factor is the factor with less influence in the global multicriteria response in the experimental design due to the low value of the b_1 coefficient and the second-order interactions in which it is involved: b_{12}, b_{13} and b_{14}. The optimisation performed allows the collection of good chromatograms as can be seen in Fig. 3.

3.4. Method validation

Once the chromatographic method had been developed and optimised, it must be validated. The validation of an analytical method verifies that the characteristics of the method satisfy the requirements of the application domain. The validation study was performed on each individual compound, on standard mixtures of the eight substances or

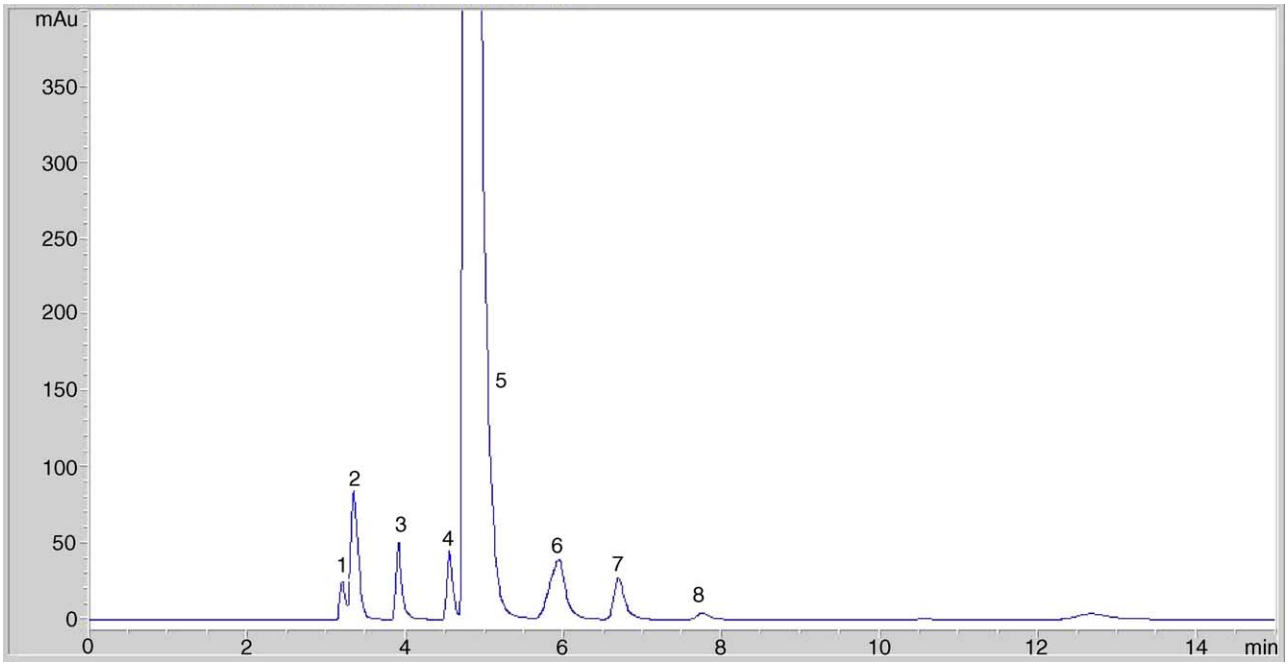


Fig. 3. Chromatogram of a standard solution of L-proline and organic acids at optimal chromatographic conditions. Peaks and concentrations(g L^{-1}): (1) L-proline, 0.504; (2) L(+)-tartaric, 0.511; (3) L-malic, 0.385; (4) lactic, 0.676; (5) acetic, 60.223; (6) citric, 0.733; (7) succinic, 0.633; (8) D-malic, 0.064.

Table 4
Characteristics of the calibration curves

Compound	Linear range ($g\,L^{-1}$)	Correlation coefficient	Linearity (LOL, %)	Slope ± S.D.	Intercept ± S.D.
L-Proline	0.105–4.040	0.9991	97.2	266.65 ± 2.84	−6.37 ± 5.72
L(+)-Tartaric acid	0.127–4.040	0.9999	96.6	1016.95 ± 3.36	16.07 ± 6.93
L-Malic acid	0.033–1.749	0.9995	94.5	632.01 ± 5.53	−4.71 ± 4.99
Lactic acid	0.104–3.110	0.9993	96.9	331.49 ± 3.14	−0.05 ± 5.13
Acetic acid	40.508–80.352	0.9998	98.3	377.65 ± 1.69	549.07 ± 105.68
Citric acid	0.065–3.600	0.9997	95.8	739.58 ± 4.17	−10.67 ± 7.42
Succinic acid	0.061–1.304	0.9985	94.7	384.71 ± 5.28	−16.20 ± 3.99
D-Malic acid	0.005–0.291	0.9907	73.6	694.51 ± 26.45	−8.54 ± 3.96

on commercial vinegar samples depending on the studied parameter.

3.4.1. Linearity, response function and calibration curves

Five concentration levels were tested in triplicate; these concentrations correspond to the expected values in commercial vinegar without any dilution. This study was performed on individual substances. The linearity range studied for each compound is shown in Table 4. An excellent linearity was obtained in all cases with correlation coefficients higher than 0.999. This was also corroborated by the 'on-line' linearity (LOL). This parameter was determined by using the following equation where RSD (m) is the relative standard deviation of the slope (expressed in %) and is the degree of dispersion of the data around the calibration line:

$$\text{LOL}\,(\%) = 100 - \text{RSD}\,(\text{m}) \tag{8}$$

The calibration parameters were validated by means of ANOVA and graphical analysis of residuals were made. Confidence intervals were also calculated for each parameter with a significance level of 5%. The correlation coefficient appears in Table 4, the lack-of-fit test was not significant which meant that the linear models described accurately the experimental data and the residuals analysis revealed that the residuals were homogeneously distributed.

3.4.2. Sensitivity

IUPAC [26] and ISO [27] define the instrumental response sensitivity as the slope of the calibration line because a method with a large slope is better able to discriminate between small differences in analyte content. It is also important that the standard deviation on the slope and the risks α and β to assert the smallest difference d that can be distinguished between two signals. To determine the smallest difference d it should be related signal to concentration using the slope m. The following equation can be proposed for the sensitivity:

$$d = (t_{1-\alpha/2} + t_{1-\beta})s\sqrt{2}\left(\frac{1}{m}\right) \tag{9}$$

where the t-values are determined for $\alpha = 0.05$ (two-sided) and $\beta = 0.05$ (one-sided) for the number of degrees of freedom with which s was determined. As in relevant precision measure (repeatability) s was determined with 10 determinations, then $t_{1-\alpha/2} = 2.26$ and $t_{1-\beta} = 1.83$ so the difference d can be calculated:

$$d = \frac{5.76s}{m} \tag{10}$$

Sometimes d value is also known as analytical sensitivity while the value of the slope of the calibration curve is the instrumental sensitivity. Table 5 shows the values of analytical and instrumental sensitivities. L-Tartaric and acetic acids are the compounds that show the fewest sensitivity values, 0.019 and 0.026 respectively.

3.4.3. Detection and quantification limits

The limits of detection (LOD) (three times the relative standard deviation of the value of the intercept divided by the value of the slope of the calibration curve) and quantification (LOQ) (10 times the limit of detection value calculated, divided by 3) obtained allow the correct determination of these compounds in real vinegar samples, taking

Table 5
Sensitivity and detection and quantification limits

Compound	Analytical sensitivity ($g\,L^{-1}$)	Instrumental sensitivity ($g\,L^{-1}$)	Detection limit (LOD, ($g\,L^{-1}$)	Quantification limit (LOQ, $g\,L^{-1}$)
L-Proline	0.061	266.65	0.088	0.294
L(+)-Tartaric acid	0.019	1016.95	0.020	0.068
L-Malic acid	0.050	632.01	0.031	0.104
Lactic acid	0.054	331.49	0.047	0.155
Acetic acid	0.026	377.65	0.840	2.798
Citric acid	0.033	739.58	0.045	0.148
Succinic acid	0.079	384.71	0.073	0.244
D-Malic acid	0.219	694.51	0.029	0.098

Table 6
Recovery study

Compound	Recovery level 1 (%, m/m)	t_{cal}	Recovery level 2 (%, m/m)	R_m[a] (%, m/m)	t_{cal}	t_{theo}
L-Proline	101.1	1.535	103.5	102.3	1.736	1.833
L(+)-Tartaric acid	100.8	1.572	102.5	101.7	1.426	1.833
L-Malic acid	98.1	1.533	102.4	100.3	1.716	1.833
Lactic acid	97.6	0.610	100.6	99.1	0.485	1.833
Acetic acid	98.4	1.227	98.9	98.7	1.065	1.833
Citric acid	99.3	0.605	98.0	98.7	1.762	1.833
Succinic acid	103.6	1.828	100.2	101.9	0.210	1.833
D-Malic acid	101.9	0.893	98.7	100.3	1.479	1.833

[a] R_m is the mean recovery percentage (m/m).

into account the concentrations found for them from [28,29]. The LOD and LOQ in Table 5 are values calculated from the calibration curve. Both parameters, the detection and quantification limits, were verified in real vinegar samples. Quantification limits were checked by verifying the repeatability and reproducibility of the individual determinations of the compounds in real vinegar samples and detection limits were checked by the observation of detection signal in the chromatogram.

3.4.4. Accuracy and recovery studies

In order to check the accuracy of this analytical method, the technique of standard additions was used. A representative sample of commercial wine vinegar was used as matrix, the concentration of the eight compounds in the sample were determined and known quantities of a global standard solution were added at two levels: 75 and 125% of the concentrations in the commercial vinegar sample. The average recovery (R_m) was calculated for each compound as the mean value obtained by two different additions of the compounds studied. Each sample was injected 10 times. The average recovery ranged from 98.7 ± 0.4 to 98.7 ± 0.9 for acetic acid and citric acid, respectively, and 102.3 ± 1.7% for L-proline. The acceptance criterion proposed was the Student's *t*-test applied to the average recovery. The null hypothesis (the recovery is close to unity and the method is accurate) was accepted at a significance level of 5%. The recovery values are collected in table in Table 6.

3.4.5. Precision: repeatability and reproducibility

A test solution of a mixture of the eight compounds studied was used for verifying the instrument precision (injection repeatability) and the intra-assay precision (within-laboratory reproducibility or intermediate injection precision).

The instrument precision (repeatability) expressed as a relative standard deviation (RSD%) or variation coefficient (CV) was evaluated from 10 successive injections and yielded values that ranged from 0.05% for acetic acid to 5.2% for D-malic acid. The intra-assay precision (within-laboratory reproducibility) was tested for 10 working days by injecting the same test solution three times a day. The values obtained ranged from 0.1% for acetic acid to 5.2% for proline.

The acceptance criterion proposed was Horwitz's variation coefficient. Horwitz's formula relates RSD to the concentration of the analyte. This relationship is widely used to predict what a reasonable variation coefficient ought to be for a given analyte concentration [30,31].

In our case, the variation coefficients obtained are lower than the Horwitz's variation coefficients proposed as reference and thus we can conclude that the method proposed has a high repeatability and reproducibility. The results are shown in Table 7.

3.4.6. Specificity and selectivity

The chromatogram of a standard solution is depicted in Fig. 3. It shows a mixture of all the acids and L-proline in an aqueous solution at concentrations close to those usually found in wine vinegar. The specificity criterion tries to demonstrate that the result of the method is not affected by the presence of interferences. These parameters were evaluated using peak purity studies that indicate if the compound elutes without any other interfering compounds. The purity analyses of the chromatographic peaks were implemented with the software HP ChemStation rev.A.06.03 by means of spectra superposition in three points through

Table 7
Repeatability and reproducibility studies

Compound	Repeatability, RSD (%)	Reproducibility, RSD (%)	Horwitz % CV
L-Proline	0.4	5.2	6.3
L(+)-Tartaric acid	0.1	0.7	6.3
L-Malic acid	0.1	0.5	6.5
Lactic acid	1.4	1.0	6.0
Acetic acid	0.05	0.1	3.2
Citric acid	0.1	1.6	5.9
Succinic acid	0.2	0.8	6.1
D-Malic acid	5.2	2.5	8.6

RSD is the residual standard deviation and CV the variation coefficient.

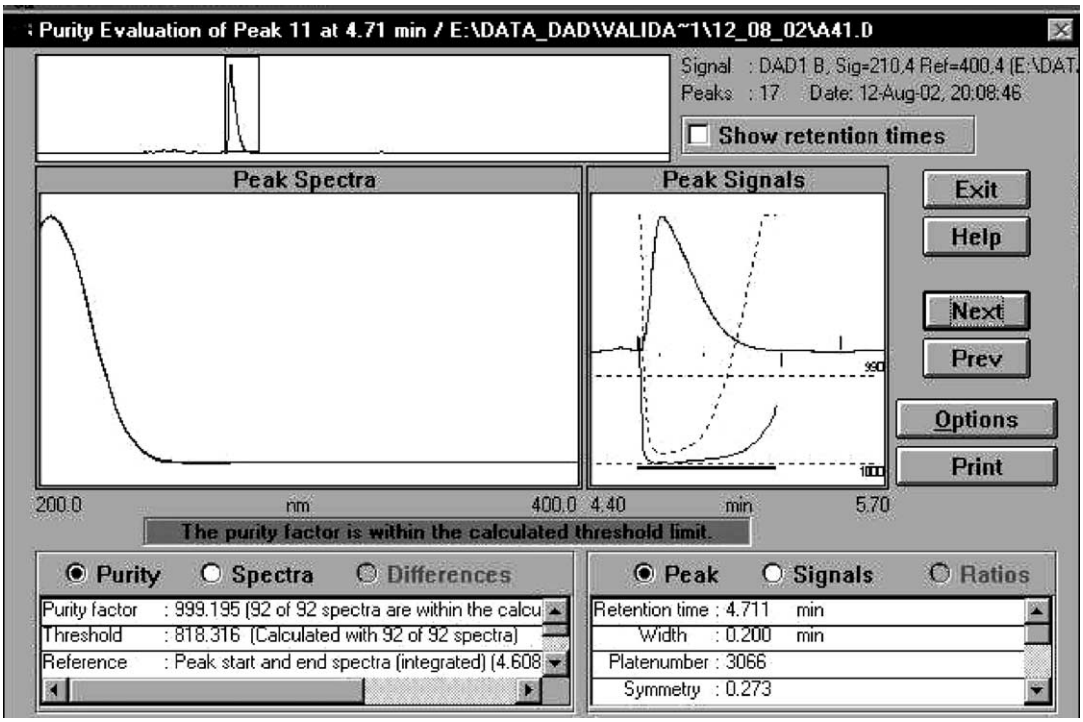


Fig. 4. Peak purity study of acetic acid for one commercial wine vinegar sample.

each chromatographic signal (the upslope, the apex and the downslope spectra) [32]. In the present case, impurities that might be found, could be some interferences due to the complexity of the vinegar matrix injected directly. For this reason, the results of the purity test for each chromatographic peak in the mixture were compared with those obtained with the individual components. Purity threshold was set at 990, so, peaks were considered pure when matched factors were superior to 99%. Peak purity test was performed over each peak to verify that it fulfilled the purity requirements. An example is shown in Fig. 4 for acetic acid peak.

3.4.7. Robustness

It was studied to detect possible critical experimental parameters that have a large effect on the results than other parameters. Controlling such parameters may lead to a better reproducibility and to avoid sources of laboratory bias.

As it was stated in the analysis of the Pareto's chart in Fig. 2, flow rate is the factor with less influence in the chromatographic system whereas the interactions between the other three factors were relevant. That conclusions were taken into account when the robustness study was done.

Experimental conditions were randomly changed: pH of eluent 2.06 ± 0.05 units, flow rate 0.64 ± 0.05 ml min^{-1}, phosphate concentration in aqueous solution 0.009 ± 0.001 M, percentage of methanol in mobile phase 8 ± 1% (v/v) and column temperature 25 ± 2 °C. Unaffected results in the peak quantification were observed so the method was considered robust.

Two analyses were done at experimental conditions I (pH 2.11, flow rate 0.69 ml min^{-1}, phosphate concentration 0.008 M, methanol 9% (v/v) and temperature 27 °C) and experimental conditions II (pH 2.01, flow rate 0.59 ml min^{-1}, phosphate concentration 0.010 M, methanol 7% (v/v) and temperature 23 °C) and the concentration of L-proline and organic acids were in both cases statistically equivalent to the result obtained with the optimal conditions. Thus, the method can be considered robust.

3.5. Determination of L- proline and organic acids in vinegar samples

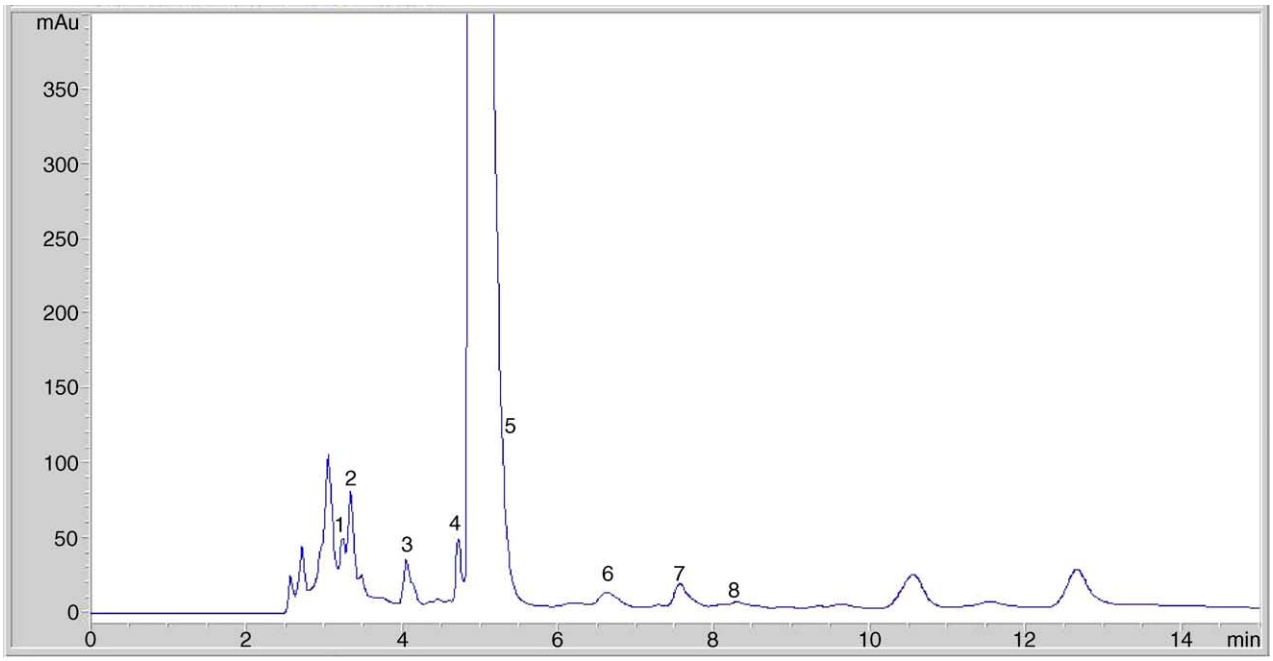
The method optimised and validated was applied to different vinegar samples directly taken from the market and the industry. Table 8 shows the maximum and minimum values registered for each type of sample and for each compound in the total amount of samples analysed. Fig. 5A shows a typical white wine vinegar whereas Fig. 5B shows the chromatogram of a typical alcohol vinegar where the only organic acid present is acetic acid. There are also traces of succinic and D-malic acids.

Synthetic vinegar mixtures of wine and alcohol vinegar were made in the laboratory and the analyses stressed the relevance of the L-proline and organic acids values to detect a potential fraudulent vinegar. The amount of these compounds in alcohol vinegar is far away from the amount in wine vinegar and a low value or even a lack of them

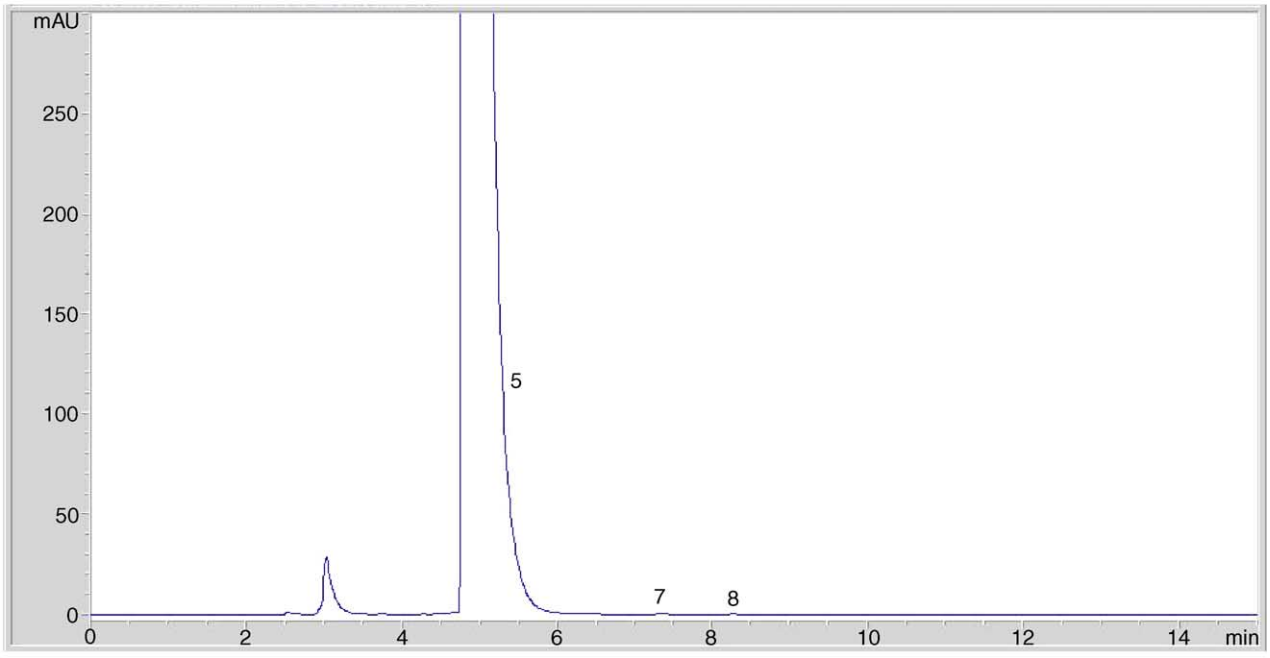
Table 8
Ranges of concentration of L-proline and organic acids in different vinegar samples

Vinegar sample	*N*	Range	Compounds							
			L-Proline	L(+)-Tartaric acid	L-Malic acid	Lactic acid	Acetic acid	Citric acid	Succinic acid	D-Malic acid
White wine vinegar	33	Maximum	22.09 ± 0.30	1.58 ± 0.03	1.03 ± 0.01	1.61 ± 0.01	62.31 ± 0.08	1.11 ± 0.01	1.62 ± 0.01	2.16 ± 0.01
		Minimum	1.01 ± 0.11	0.33 ± 0.01	0.18 ± 0.01	0.16 ± 0.01	58.20 ± 0.08	nd	0.18 ± 0.02	nq
Red wine vinegar	13	Maximum	24.87 ± 0.17	1.46 ± 0.01	1.08 ± 0.01	0.94 ± 0.03	69.30 ± 0.06	1.53 ± 0.01	1.23 ± 0.01	2.69 ± 0.02
		Minimum	1.69 ± 0.15	0.41 ± 0.03	0.18 ± 0.01	0.14 ± 0.03	57.80 ± 0.08	nq	0.27 ± 0.03	0.25 ± 0.01
Alcohol vinegar	5	Maximum	nd	nd	nq	nq	60.97 ± 0.02	nd	nq	nq
		Minimum			nd	nd	50.99 ± 0.01			
Malt vinegar	3	Maximum	7.74 ± 0.09	nd	nq	4.66 ± 0.05	48.78 ± 0.03	0.33 ± 0.03	8.51 ± 0.04	nq
		Minimum	1.02 ± 0.01		nd	nq	50.61 ± 0.08	nd	0.33 ± 0.01	nd
Cider vinegar	11	Maximum	8.85 ± 0.34	0.82 ± 0.03	2.38 ± 0.01	2.44 ± 0.01	52.87 ± 0.11	1.30 ± 0.04	1.79 ± 0.05	0.24 ± 0.01
		Minimum	nd	nd	0.10 ± 0.01	1.12 ± 0.01	48.63 ± 0.03	nq	0.51 ± 0.11	nq
Mixture I	1		0.46 ± 0.01	0.16 ± 0.01	0.18 ± 0.01	0.37 ± 0.01	60.23 ± 0.04	nq	0.35 ± 0.01	nq

Results expressed as mean ± 95% confidence interval. nd is not detectable, values under the LOD and nq is not quantifiable, values under the LOQ.



(A)



(B)

Fig. 5. Chromatograms of a white wine vinegar (A) and an alcohol vinegar (B) at optimal chromatographic conditions. Peaks: (1) L-proline, (2) L(+)-tartaric, (3) L-malic, (4) lactic, (5) acetic, (6) citric, (7) succinic; (8) D-malic.

suggests an alteration. In Table 8 can be observed a notable decrease of L-proline, L- and D-malic, lactic, citric and succinic acids in the mixture that can be quantified.

Cider vinegar has an astringent taste and aromas reminiscent of fruits. The total acidity in commercial samples must reach values of up to 5°. This kind of vinegar is characterised by the absence of tartaric acid. In some of the samples the content of L-proline was also very low. On the contrary, lactic acid content was higher than in a typical wine vinegar (Fig. 6A).

Malt vinegar is a bitter vinegar with a large amount of lactic acid, even more abundant than in cider vinegar samples. Malt vinegar samples are characterised by the lack of L-tartaric and L and D-malic acids (Fig. 6B).

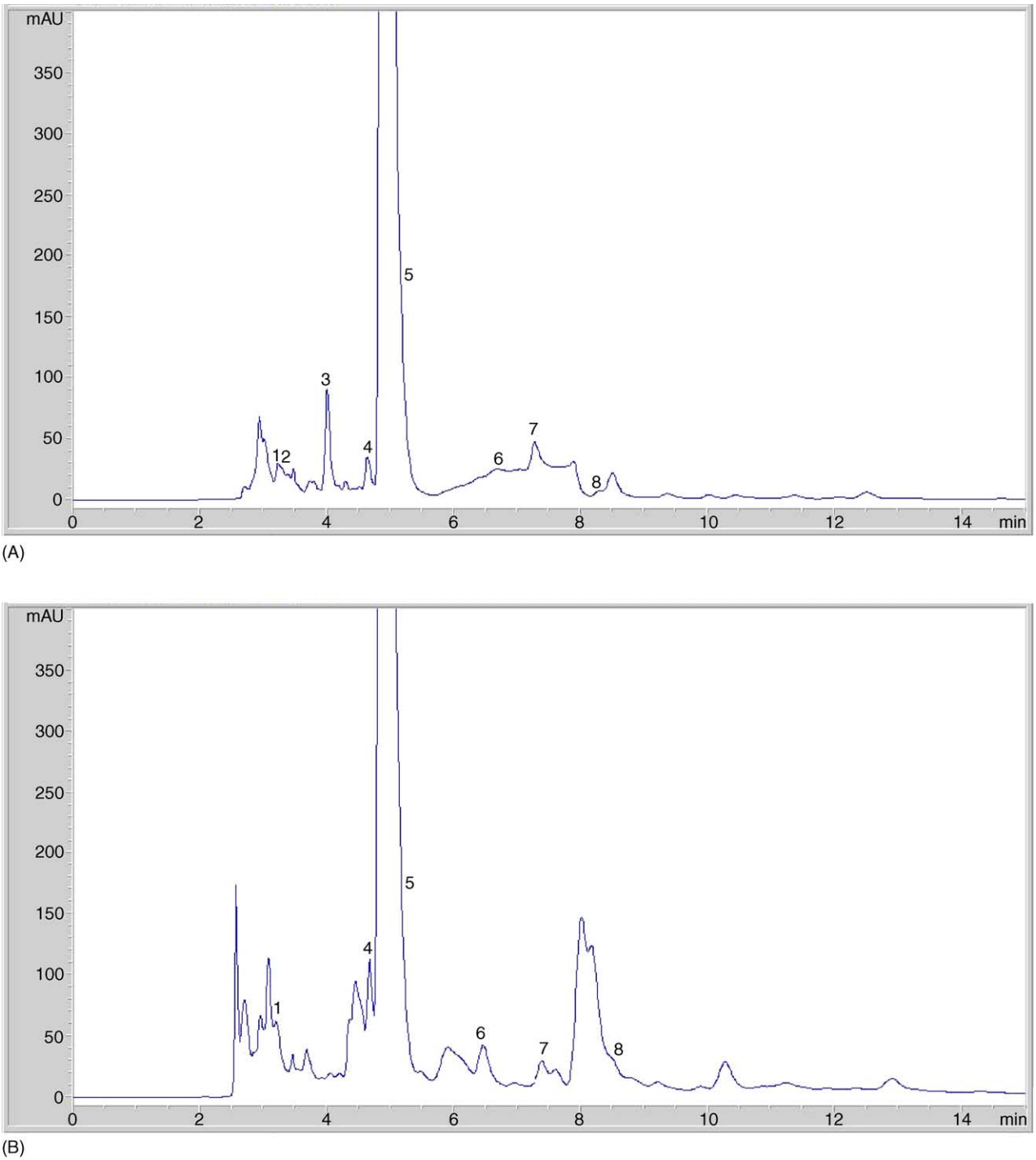


Fig. 6. Chromatogram of a cider vinegar (A) and a malt vinegar (B) at optimal chromatographic conditions. Peaks: (1) L-proline, (2) L(+)-tartaric, (3) L-malic, (4) lactic, (5) acetic, (6) citric, (7) succinic, (8) D-malic.

4. Conclusions

The multicriteria strategy proposed for optimisation of the chromatographic analysis of L-proline and organic acids in vinegar, permits the collection of quality chromatograms used for the quantification of L-proline and organic acids in vinegar. Besides, the aim of this study allows to obtain interesting data that make new ways to detect fraudulent vinegar mixtures of vinegar samples or to control the fermentation process through the study of the changes in the concentration of L-proline and organic acids. No pre-treatment is needed and the validation results are satisfactory. Therefore, the analysis is not time-consuming and suitable for the final purpose of the research.

Acknowledgements

The authors thank the Spanish Government (Ministerio de Ciencia y Tecnología, INIA CAL01-053), the Local Government of La Rioja (Consejería de Educación, Juventud y Deportes) under the research grant FPI-2001 and their financial support (ACPI-2002/07), the University of La Rioja for their financial support (API-02/16, API-03/13) and Professor Phan-Tan-Luu of the University of Marselle (France) for providing Nemrod-W.

References

[1] L.R. Snyder, J.J. Kirkland, J.L. Glach, Practical HPLC Method Development, 2nd ed., Wiley-Interscience, New York, 1997.
[2] L. Glajch, L.R. Snyder, Computer-assisted Method Development for High-performance Liquid Chromatography, Elsevier, Amsterdam, 1990.
[3] R. Cela, J.A. Martínez, C. González-Barreiro, M. Lores, Chemom. Intell. Lab. Syst. 69 (2003) 137.
[4] P. Nikitas, A. Pappa-Louisi, A. Papageorgiou, A. Zitrou, J. Chromatogr. A 942 (2002) 93.
[5] X.G. Chen, X. Li, L. Kong, J.Y. Ni, R.H. Zhao, H.F. Zou, Chemom. Intell. Lab. Syst. 67 (2003) 157.
[6] B. Bourguignon, D.L. Massart, J. Chromatogr. A 586 (1991) 11.
[7] M. Jimidar, B. Bourguignon, D.L. Massart, J. Chromatogr. A 740 (1996) 109.
[8] E. García Romero, G. Sánchez Muñoz, P.J. Martín Alvarez, M.D. Cabezudo Ibáñez, J. Chromatogr. A 655 (1993) 111.
[9] M.L. Morales, A.G. González, A.M. Troncoso, J. Chromatogr. A 822 (1998) 45.
[10] D. Heems, G. Luck, C. Fraudeau, E. Vérette, J. Chromatogr. A 798 (1998) 9.
[11] M. Calabrese, B. Stancher, A. Del Signore, Ital. Beverage Technol. 11 (28) (2002) 11.
[12] S.H. Ashoor, M.J. Knox, J. Chromatogr. A 299 (1984) 288.
[13] S.H. Ashoor, J. Welty, J. Chromatogr. A 287 (1984) 452.
[14] L. Bertolini, G.B. Castagnetti, G. Gambini, Industrie delle Bevande 23 (132) (1994) 324.
[15] D. Chabert, B. Giorgio, C. Tisse, M. Guérrère, J. Estienne, Ann. Fals. Exp. Chim. 80 (1987) 259.
[16] D. Kutlán, I. Molnár-Perl, J. Chromatog. A 987 (2003) 311.
[17] D. Heems, G. Luck, C. Fraudeau, E. Vérette, J. Chromatogr. A 798 (1998) 9.
[18] Statgraphics Plus 4.0, Professional Version, Copyright 1994–1999 by Statistical Graphics Corp.
[19] Nemrod-W. version 9901 by Didier Mathieu, Jean Nony, Roger Phan-Tan-Luu.
[20] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke, Data Handling in Science and Technology 20[a], Handbook of Chemometrics and Qualimetrics: Part A, Elsevier, Amsterdam, 1997.
[21] F. Hellal, R.Phan. Tan-Luu, A.M. Siouffi, Analusis 24 (1996) 127.
[22] D.H. Doehlert, Appl. Statist. 9 (1970) 231.
[23] E. Hund, Y. Vander Heyden, M. Haustein, D.L. Massart, J. Smeyers-Verbeke, J. Chromatogr. A 874 (2000) 167.
[24] G.A. Lewis, D Mathieu, R. Phan. Tan-Luu, Pharmaceutical Experimental Design, vol. 92. Marcel Dekker, New York, 1999, pp. 221–235.
[25] R. Todeschini, Introduzione alla chemiometria, EdiSES, Napoli, 1998.
[26] L.A. Currie, Pure Appl. Chem. 67 (1995) 1699.
[27] ISO Standard 8466-2, Water quality – Calibration and evaluation of analytical methods and estimation of performance characteristics, Part 2, Calibration strategy for non-linear second order calibration functions, 1994.
[28] C. Llaguno, C. Polo, El Vinagre de Vino, Consejo Superior de Investigaciones Científicas, CSIC, Madrid, 1991.
[29] F. Mecca, R. Andreotti, L. Veronelli, L'Aceto, Edizioni aeb brescia, Brescia, 1979.
[30] W. Horwitz, L.R. Kamps, K.W. Boyer, JAOAC 63 (1980) 1344.
[31] W. Horwitz, Anal. Chem. 54 (1982) 67A.
[32] US FDA, Proceedings of the International Conference on Harmonisation, Draft Guideline on the Validation of Analytical Procedures: Methodology, Availability, Notice, Department of Health and Human Services, 1996.