

**Scope of partial least-squares regression applied to the enantiomeric composition  
determination of ketoprofen from strongly overlapped chromatographic profiles**

Juan M. Padró <sup>1</sup>, Jaiver Osorio-Grisales <sup>1</sup>, Juan A. Arancibia <sup>2</sup>, Alejandro C. Olivieri <sup>2</sup>,  
Cecilia B. Castells <sup>1,\*</sup>

<sup>1</sup> LIDMA (Laboratorio de Investigación y Desarrollo de Métodos Analíticos) y División  
Química Analítica, Universidad Nacional de La Plata, 47 y 115 (1900) La Plata, Argentina.

<sup>2</sup> Departamento de Química Analítica, Facultad de Ciencias Bioquímicas y Farmacéuticas,  
Universidad Nacional de Rosario, Instituto de Química de Rosario (IQUIR-CONICET),  
Suipacha 531, Rosario S2002LRK, Argentina.

**Keywords**

Chiral analysis, enantioresolution, ketoprofen, liquid chromatography, multivariate  
calibration

Corresponding author: C. Castells

E-mail: castells@isis.unlp.edu.ar

Tel: 0054 221 4228328

Received: 25-Feb-2015; Revised: 27-Apr-2015; Accepted: 27-Apr-2015

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/jssc.201500217.

This article is protected by copyright. All rights reserved.

## Abstract

Valuable quantitative information could be obtained from strongly overlapped chromatographic profiles of two enantiomers by using proper chemometric methods. Complete separation profiles where the peaks are fully resolved are difficult to achieve in chiral separation methods, and this becomes a particularly severe problem in case that the analyst need to measure the chiral purity, i.e., when one of the enantiomers is present in the sample in very low concentrations. In this report, we explore the scope of a multivariate chemometric technique based on unfolded partial least-squares regression, as a mathematical tool to solve this quite frequent difficulty. This technique was applied to obtain quantitative results from partially overlapped chromatographic profiles of *R*- and *S*-ketoprofen, with different values of enantioresolution factors (from 0.81 down to less than 0.2 resolution units), and also at several different *S*:*R* enantiomeric ratios. Enantiomeric purity below 1% was determined with excellent precision even from almost completely overlapped signals. All these assays were tested on the most demanding condition, i.e., when the minor peak elutes immediately after the main peak. The results were validated using univariate calibration of completely resolved profiles and the method applied to the determination of enantiomeric purity of commercial pharmaceuticals.

## 1 Introduction

Chirality is now a major issue in the design and development of new drugs and agrochemicals, supported by the understanding of the role of molecular recognition in many relevant biological events [1]. Also, a sustained increase in the development of chiral drugs in pharmaceutical industries has begun since three decades ago, the main motivation is that when a patent on racemic drug expires, pharmaceutical companies have the opportunity to extend its rights through the development of the chiral switch to the enantiomer with the desired activity [2]. At the same time, the new policy statements issued by the regulatory agencies promoted that new chiral drugs have to be produced in enantiomerically enriched pure forms [3] and the trends in the agrochemical industries are somewhat similar [4]. All these policies caused a significant change in the way the production plants are managed [5].

Kinetic, pharmacological and toxicological properties of the individual enantiomers need to be fully assessed, regardless of the decision whether the chiral new product will be marketed in an enantiomerically pure form or as racemic mixture. From the analytical control point of view, this implies that a growing number of methods need to be developed aimed at the determination of the enantiomeric purity of optically active samples. In our opinion, methods based on liquid, gas or supercritical chromatography resorting to chiral columns are specially sensitive, simple and robust to be considered as gold standards.

The development of a chiral chromatographic method often requires expensive and time-consuming testing of different columns and chromatographic conditions. Furthermore, a typical target value for enantiomeric purity is 99.5% [6] and the enantiomer of interest (minor peak) has to be usually quantitatively analyzed at levels down to 1% of the main enantiomer. This experimental situation requires injecting relatively large amounts of sample to increase the sensitivity for the minor peak and, as a consequence, the major component

might become broader and tailed as a result of mass overloading. This is especially a problem when the minority peak elutes after the big one, the little peak remains under the tail unless higher enantioresolution rates can be achieved. In RP-HPLC, baseline enantioresolutions are not easily obtained and, very often, it is necessary to decrease elution strength to obtain enough enantioresolution. There are also many examples where total separation is impossible under RPLC conditions and the use of normal-phase HPLC is needed [7,8]. Either, very long analysis times in RPLC or the use of normal phase are expensive approaches. Moreover, pretreatment of biological samples usually requires specific polar solvents incompatible with normal phase [8]. Usually, the complexity of an enantioseparation problem can be simplified by the use of circular dichroism detectors (CD) [9–11].

Chemometric methods would be an interesting choice to determine enantiomeric ratios. Busch et al. reported the multivariate analysis of UV-spectral data in the presence of  $\beta$ -cyclodextrin for the determination of enantiomers of amino acids [12]. Later the authors extended the technique to model UV and fluorescence signals of a variety of chiral molecules, including pharmaceuticals, using several cyclodextrin derivatives as chiral auxiliaries [13–15], the authors demonstrated that although the spectral differences were small, the signals could be correlated with enantiomeric composition using partial least-squares (PLS) regression. More recently, the authors coupled IR spectroscopy with chemometrics for determining the enantiomeric excess of the pharmaceuticals norephedrine and phenylalanine [16]. Similarly, Zhou et al. proposed a successful chemometric analysis of attenuated total reflectance IR spectra of host-guest complexes of an M3 antagonist drug [17]. More recently, PLS and also multivariate curve resolution (MCR) were combined with traditional IR spectra *without* any chiral auxiliary, since it was observed that the identical spectrum due to the R and the S-enantiomers changed their IR pattern for the racemic

Accepted Article

mixtures, the authors modeled the regression signals of ketoprofen and also of mandelic acid solutions [18]. Valderrama and Poppi [19] developed a method based on parallel factor analysis (PARAFAC) [20] to model fluorescence excitation/emission matrix data of ibuprofen in the presence of  $\beta$ -cyclodextrin and 1-butanol. Very recently, the use of chemometrics for improving the methods of enantioseparation (experimental design) has been reviewed, explaining some specific retention mechanisms and chiral recognition (quantitative structure enantioselective retention relationships) and indicating chiral purity (enantiomeric excess) from different spectroscopic signals [21]. The determination of enantiomeric fraction of not completely resolved UV-signals of ibuprofen solutions using direct chiral LC combined with chemometric data analysis, has been recently proposed in our laboratory [22]. In that study, *R*-ibuprofen was accurately quantified even at a concentration level of 0.5 mg L<sup>-1</sup> in the presence of 99.9% of the major enantiomer, by using an unfolded partial least-squares (U-PLS) algorithm [23] applied to elution profiles at multiple wavelength detection and even though the enantioresolution factor was as low as 0.87. This was possible due to the high sensitivity of methods based on second-order data for analyte quantitation.

In the present study we were not particularly interested in the chemometric resolution of the chromatographic profiles of *R*- and *S*-ketoprofen. The aim was to focus on the scope of the U-PLS chemometric algorithm to provide precise quantitative results after subsequently decreasing enantioresolution factors well below the unit and also for samples of different enantiomeric ratios. As a proof of concept, a very simple enantiomeric pair was selected with the objective of exploring the limits in the overlapping of the profiles beyond which the quality of the results would be seriously affected.

Ketoprofen, 2-(3-benzolphenyl)propionic acid, is a largely consumed non-steroidal anti-inflammatory drug with analgesic and antipyretic effects. As with other profens, the *S*-(+) enantiomer has the desired pharmacological activity, and the *R* enantiomer is an impurity.

Complete chromatographic enantioresolution of ketoprofen is feasible under specific conditions [24–27], but our goal was to face a study taken ketoprofen as a model solute to demonstrate the applicability of chemometric algorithms to quantitatively resolve very difficult enantioseparations. These algorithms would make possible: i) faster chromatographic runs and, therefore, shorter retention times; ii) switching methods from in normal phase to RPLC, and consequently, replacement to more environmentally acceptable solvents; iii) quantitative analysis when total enantioresolution is not possible.

## 2 Materials and methods

### 2.1 Reagents and solutions

Racemic ketoprofen and (*S*)-(+)-ketoprofen were obtained from Sigma–Aldrich (St Louis, MO), HPLC-grade methanol was purchased from J. T. Baker (Edo. México, Mexico) and water was obtained from a Milli-Q System (Simplicity, Millipore, MA). Mobile phases consisted in mixtures of methanol/buffer 0.1% triethylammonium acetate (TEAA) pH=4.0 (measured in pure water). TEAA Buffer was generated by mixing 300  $\mu$ L of triethylamine (TEA, Anedra, Buenos Aires) and 700  $\mu$ L of glacial acetic acid (Merck, Darmstadt, Germany) in 1 L of Milli-Q water. Commercial pharmaceutical *S*-ketoprofen Enantyum® was purchased from Laboratorios Menarini (Badalona, Spain).

A set of 18 calibration solutions were prepared by mixing stock solutions of racemic ketoprofen and *S*-(+)-ketoprofen to obtain three duplicated levels of each one of the three *S*:*R* compositions (99.5:0.5; 99:1 and 95:5). Solution concentrations are reported in Table 1.

Thirty test solutions, with random concentrations (ten for each concentration rate), were prepared from stock solutions with the three composition rates (Table 2).

Samples of *S*-ketoprofen pharmaceutical tablets were treated as follows. Ten pills were crushed and dissolved in methanol and centrifuged at 4000 rpm for 20 min to yield two stock solutions of about  $1.6 \text{ mg mL}^{-1}$  (considering the nominal content of active declared by the manufacturer). The standard addition method was applied to determine the accurate *S*-ketoprofen concentration. Adequate dilutions were prepared from the latter two stock solutions in 50:50 (% v/v) methanol/buffer TEAA 0.1% pH: 4 and, then filtered through a  $0.45 \mu\text{m}$  nylon membrane before injection.

## 2.2 Instrumentation

All chromatographic studies were performed on an HP 1100 liquid chromatograph (Agilent Technologies, Palo Alto, CA) equipped with vacuum degasser, binary pump, autosampler, thermostatted column device, and photodiode array detector (DAD). The chiral column was a Nucleodex- $\beta$ -PM ( $200 \times 4.0 \text{ mm}$ , particles of  $5 \mu\text{m}$ ) from Macherey-Nagel (Düren, Germany). Mobile phase compositions are given in Table 3. The flow rate was set to  $0.6 \text{ mL min}^{-1}$  and temperature to  $25^\circ\text{C}$ . The injection volume was  $10 \mu\text{L}$ . The output signals from the DAD detector were acquired between 190 and 290 nm every 1 nm with a frequency of 2.5 Hz. Chromatograms obtained were exported in matrix form, from 190 to 290 nm every 1 nm and within a window time adapted according to the elution time of both peaks in each mobile phase condition. Details are also given in Table 3. However, for chemometric analysis, the high-absorbance region below 215 nm was avoided.

### 2.3 Validation of the model

There is not an official method to determination of enantiomeric purity of ketoprofen, only some guidance for impurity determinations are given by The International Conference on Harmonisation (ICH) [<http://www.ich.org/cache/compo/276-254-1.html>] [28]. Thus, to validate the results obtained from multivariate methods, classical univariate calibration was applied under conditions of enantioresolution of 1.0, using a more polar phase (see Table 3) and data were recorded at 260 nm. Figures of merit from univariate calibration were then compared with those from U-PLS chemometric method. All details about the theory of U-PLS methods are provided in Supporting Information.

### 2.4 Software

Chromatographic alignment, vectorization of data matrices and application of the U-PLS algorithm were made using MATLAB R2010a (The MathWorks, Natick, MA, USA). In the latter case, the MVC2 graphical interface toolbox was employed, which is available at [www.iquir-conicet.gov.ar/descargas/mvc2.rar](http://www.iquir-conicet.gov.ar/descargas/mvc2.rar).



### 3 Result and Discussion

#### 3.1 Chromatographic considerations

Four methods, named M1 to M4, consisting in the increase of solvent elution strength were used. Enantioresolution factors estimated with racemic mixtures were of ca. 1.0 with mobile phase 50:50 MeOH/buffer (M1). To worsen the enantioseparations, the methanol fraction was raised to 55 (M2), 60 (M3) and 65 (M4) (% v/v). Figure 1 shows the obtained chromatograms under each of these conditions. Under the less hydrophobic mobile phases the (apparent) enantioresolution factors, measured from the peak top, decreased to 0.81, 0.55 and 0.17, respectively. These resolutions corresponded to the racemic ketoprofen mixture, where both peaks have the same areas. However, for enantiomeric fraction determinations, real conditions consist in quantifying very low levels of the *R* enantiomer, therefore solutions containing *S*:*R* ratios of 95:5, 99:1 and 99.5:0.5 were prepared. Figure 2 compares the peak profiles corresponding to a 95:5 enantiomeric ratio and collected in the four eluent conditions (the best resolution one compared to the three poorly resolved ones). Clearly, a single peak without any shoulder is observed in the chromatograms corresponding to the two stronger solvent mixtures. These partially separated profiles cannot be individually analyzed with a mass detector, since enantiomers have identical ionization patterns.

Deconvolution methods are not applicable to these severely overlapped profiles with identical spectra [29], which show, in addition, lack of reproducibility in the elution profiles from sample to sample, both in peak position and shape. These challenges make the PARAFAC algorithm, for example, inapplicable to the present problem. In the case of MCR, this algorithm is able to cope with changes in elution profiles for different experimental runs; however, the fact that the spectra for both sample components are identical precludes its application [30]. The best alternative to analyze the present system is U-PLS after a simple chromatographic alignment procedure applied to peak positions, with the latent variables employed by this algorithm modeling the changes in elution profile shapes.

### 3.2 Time alignment of elution profiles

The chromatographic profiles must be registered in the same time position in the data matrix so the algorithms can recognize the signals correctly, but small changes in HPLC elution profiles are unavoidable, probably due to slight fluctuations in mobile phase composition, column equilibration time or column effective temperature, even when the column has been thermostatted (Figure 3A). For this reason, it is necessary to align the chromatographic profiles in the elution time direction. In principle, this could be done using alignment algorithms seeking to maximize the correlation between a chromatogram and a reference one [31] However, this is not feasible in cases where strong profile overlapping produces a single peak or a peak with a shoulder. Thus, in the present case the recorded data matrices of each resolution condition were aligned by selecting one of them as a reference, and digitally moving all the others in the time direction until the maximum peak for the major component was aligned with the one in the reference matrix (see Figure 3B).

### 3.3 U-PLS analysis

In our previous work [22], the latent structured method unfolded partial least-squares was successfully used, due to its intrinsic flexibility towards profile shape changes, being modeled during the calibration phase. As explained in the theoretical section, a PLS model is built using unfolded calibration data and analyte concentrations, providing regression coefficients for prediction in new samples [23]. Figures of merit, such as sensitivity, uncertainty in predicted concentration and LOD, in the latter case incorporating the latest IUPAC recommendations [32–35], can be readily estimated using known expressions.

Leave-one-out cross validation was employed to estimate the optimal number of factors for U-PLS modeling. In each of the four studied resolution cases, the time region was selected as the one leading to the minimum predicted error sum of squares (PRESS) value for cross-validation, using a moving-window strategy applied to the elution profiles at the single wavelength of 215 nm, selecting the optimum number of factors which was then applied to the determination of both enantiomers in the test solutions. A relatively large number of factors are initially selected to calculate the PRESS value for each factor, and the optimal number of factors are obtained by observing the changes of the PRESS. The finally selected values of  $A$  are quoted in Table 4 at each resolution value.

The accuracy and robustness of the calibration were demonstrated by predicting a set of independent validation samples described in Table 2. The root mean square error of prediction (RMSEP) is a measure of the variability of the difference between predicted and nominal values for a set of validation samples. It provides an idea of prediction uncertainty and bias. A plot of RMSEP as a function of resolution factor is shown in Figure 4. This plot includes the RMSEP calculated with chromatograms corresponding to M1. There is no significant difference between predictions corresponding to baseline resolution with that corresponding to  $R_s=0.81$ , so we can reduce the analysis time without loss in the quantification information quality. At the worst resolution conditions, prediction errors (REP%) grow up to 4 and 9% for the  $R$  and the  $S$  peaks, respectively, which are still reasonably precise quantitative results, considering the extremely large degree of chromatographic overlap (see Figures 1 and 2). Larger errors for the  $S$  enantiomer as compared to the minor peak in any condition are attributed to the overloading mass for the more concentrated solutions. Peak profiles developed a slight fronting for the most concentrated solutions. The use of those data during calibration undoubtedly impaired the precision. Relative error predictions within the range of 3 to more than 20% were reported

previously, values that were highly dependent on the host molecule used for their UV-absorption spectra measurements [14]. These differences were attributed to the stability of the diastereomeric complexes that can be formed and its effect on the UV spectra. Predictions based PLS models constructed from IR spectra showed somehow lower standard error of prediction (1.3%), but the validation samples were prepared with enantiomeric excess below 98.2% [17]. Similarly, relatively low predictive errors (lower than 4%) were found for ibuprofen fluorescence excitation/emission measurements in combination with PARAFAC, but once again, the mole fraction of enantiomers range from 50 up to 80% [19].

Table 4 shows the results of sensitivity, and limits of detection and quantification for univariate analysis and U-PLS calculations. The statistical prediction results (RMSEP, REP%,  $R^2$ ) shown in Table 4 for M1 and M2 are similar and statistically comparable, and are seemingly better than those corresponding to M3 and M4. The latter values are relevant, since they allow us to assess the detection capability of the developed method concerning the minor component in the studied mixtures. According to Table 4, at the elution conditions corresponding to M2, the LOD was as low as  $0.1 \text{ mg L}^{-1}$ . Since the mean calibration concentration of the major enantiomer is ca.  $600 \text{ mg L}^{-1}$ , the later LOD value corresponds to a ratio 6000:1, i.e., a composition rate 99.98:0.02.

With regard to the numbers of optimum latent variables, they are in general smaller in the case of the major *S* enantiomer, due to the fact that the signal from the minor *R* enantiomer is considerably small. The latter is greatly influenced by the signal from the major *S* enantiomer and also by any residual variation of the chromatographic bands from sample and to sample.

Table 5 shows a step-by-step protocol to obtain quantitative data.

### 3.4 Validation

The obtained results have been validated by:

- i.* comparison of the figures of merit obtained by univariate calibration of samples run under chromatographic condition M1 with those get from U-PLS and,
- ii.* determination of *S*-ketoprofen amount in pharmaceuticals.

The results obtained from U-PLS and from univariate calibration are compared in Table 4. Amazingly, the differences between all the figures, even those obtained under the worse enantioresolution conditions are of significantly superior quality when multivariate data were used. This is a clear indication of the potential of chemometric methods to acquire analytical information even though these profiles were severely overlapped.

Table 6 gathers the results of the *S*-ketoprofen predicted in commercial tablets, using the three most potentially useful U-PLS calibrations. Samples were analyzed by HPLC under complete enantioseparation and by applying the U-PLS algorithm after partial separation. Standard addition calibration results for the *R* enantiomer indicated its absence in these real samples (below the LOD). The relative errors between nominal and found concentrations are shown in the Table 4. The statistical analysis of the recoveries, made through paired *t*-statistics, indicates that predictions of *S*-ketoprofen in these formulations were in excellent agreement with the nominal concentrations for M1, quite good for M2, and even reasonable for M3.

## 4 Conclusions

High-quality quantitative information in chiral analysis could be achieved even when complete separation is not possible and other analytical techniques are useless. The implementation of chemometric methods as a second data mode to deal with unresolved enantiomeric profiles allows decrease chromatographic costs by replacing normal phase LC by usually cheaper RP methods or, either, by using stronger solvent conditions to reduce analysis time. For instance, there are not significant differences in precision calculated from the chromatograms partially overlapped obtained with the two weaker mobile phases, so we can shorten the time of analysis to quantify the minor peak and, depending on the requirements, it is possible to quantify enantiomeric composition from strongly overlapped profiles. In this work, enantiomeric purity well below 1% could be determined with excellent precision even from highly overlapped signals and on the worst condition, i.e., a minor peak eluting after the main enantiomer.

## Acknowledgments

The authors kindly acknowledge Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT), Universidad Nacional de La Plata (UNLP) and Universidad Nacional de Rosario (UNR) for financial support.

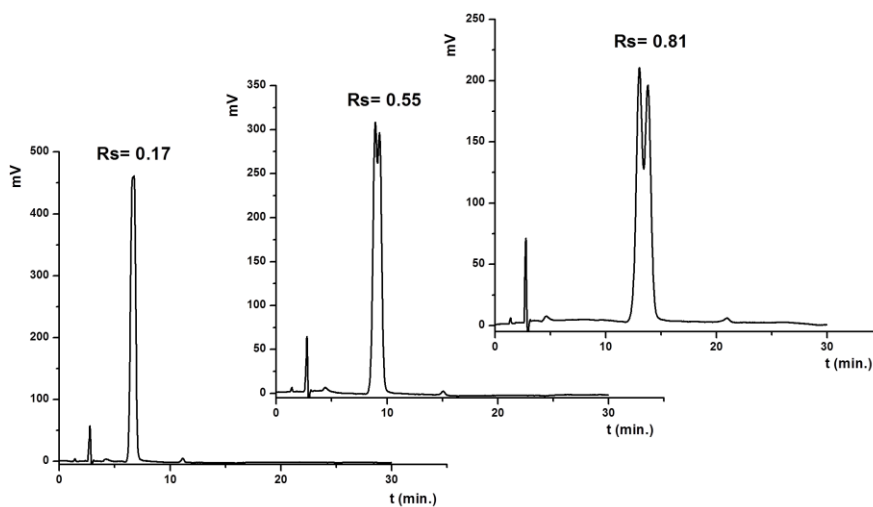
## References

- [1] Berthod, A. (Ed.) *Chiral Recognition in Separation Methods. Mechanisms and Applications*, Springer, Heidelberg 2010.
- [2] Agrat, I., Caner, H., Caldwell, J., *Nature Reviews* 2002, 1, 753–768.
- [3] Zeid, R. L. *Regulatory and Development Considerations of Chiral Compounds, in: Chiral Separation Methods for Pharmaceutical and Biotechnological Products* (S. Ahuja, Ed.), Wiley, Hoboken 2011.
- [4] Jin, L., Gao, W., Li, L., Ye, J., Lin, C., Liu, W., Enantioseparation and Enantioselective Analysis of Chiral Herbicides, In *Herbicides, Theory and Applications* (S. Soloneski, M.L. Larramendy, Eds.), , InTech open science, Rijeka 2010, pp. 281–308
- [5] Stinson, S. C., *Chem. Eng. News* 2000, 23, 55–78.
- [6] Carey, J. S., Laffan, D., Thomson, C., Williams, M. T., *Org. Biomol. Chem.* 2006, 4, 2337–2347.
- [7] Al-Othman, Z. A., Al-Warthan, A., Ali, I., *J. Sep. Sci.*, 2014, 37, 1033–1057.
- [8] Tachibana, K., Ohnishi, A., *J. Chromatogr. A*, 2001, 906, 127–154.
- [9] Eto, S. Yamaguchi, M. Bounoshita, M., Mizuloshi, T., Miyano, H., *J. Chromatogr. B*, 2011, 879, 3317–3325.
- [10] Kiesswetter, R., Brandl, F., Kastner-Pustet, N., Mannschreck, A., *Chirality*, 2003, 15, 40–49.
- [11] Purdie, N., Brittain, H.G. (Eds.), *Analytical Applications of Circular Dichroism*, Elsevier, Amsterdam 1994, pp. 279–292.
- [12] Busch, K.W., Swamidoss, I. M., Fakayode, S. O., Busch, M. A., *J. Am. Chem. Soc.*, 2003, 125, 1690–1691.
- [13] Busch, K.W., Busch, M. A. (Eds.) *Chiral Analysis*, Elsevier, Amsterdam 2006.
- [14] Fakayode, S. O., Swamidoss, I. M., Busch, M. A., Busch, K. W., *Talanta*, 2005, 65, 838–845.
- [15] Fakayode, S. O., Williams, A. A., Busch, M. A., Busch, K. W., *J. Fluoresc.*, 2006, 16, 659–670.
- [16] Ingle, J. R., Busch, K. W., Busch, M. A., *Talanta*, 2008, 75, 572–584.
- [17] Zhou, L., Zhihao, L., Welch, C. J., Ge, Z., Ellison, D., *Chirality*, 2006, 18, 306–313.
- [18] Marini, F., Bucci, R., Ginevro, I., Magri, A. L., *Chem. Intell. Lab. Syst.*, 2009, 97, 52–63.
- [19] Valderrama, P., Poppi, R. J., *Chem. Intell. Lab. Syst.*, 2011, 106, 160–165.

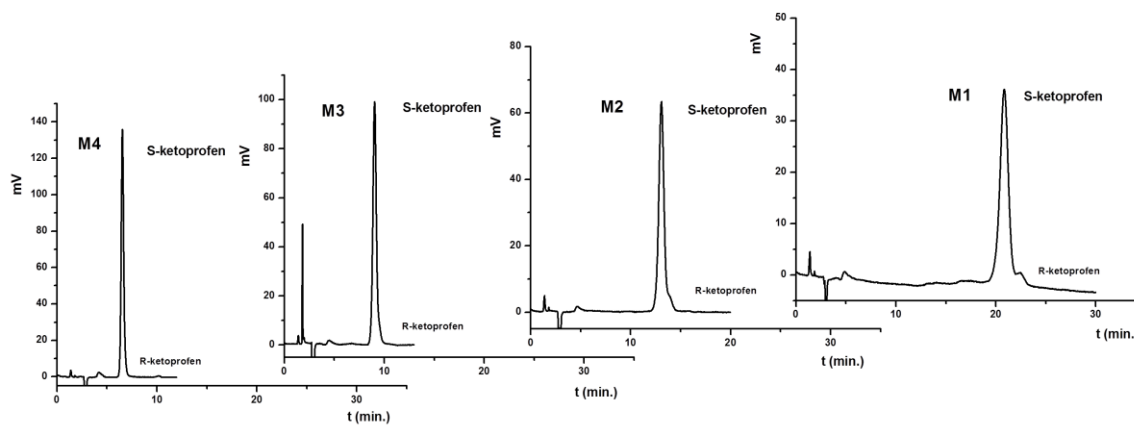
- [20] Bro, R., *Chem. Intell. Lab. Syst.*, 1997, 38, 149–171.
- [21] Chmielewska, A., Baczek, T., *J. AOAC Int.*, 2012, 95, 624–635.
- [22] Grisales, J. O., Arancibia, J. A., Castells, C. B., Olivieri, A. C., *J. Chromatogr. B*, 2012, 910, 78–83.
- [23] Wold, S., Esbensen, K., Geladi, P., *Chem. Intell. Lab. Syst.*, 1987, 2, 37–52.
- [24] Ameyibor, E., Stewart, J. T., *J. Pharm. Biomed. Anal.*, 1998, 17, 83–88.
- [25] Guo, Z., Wang, H., Zhang, Y., *J. Pharm. Biomed. Anal.*, 2006, 41, 310–314.
- [26] Hatami, M., Farhadi, K., *Can. J. Chem.*, 2013, 91, 1252–1257.
- [27] Mullangi, R., Yao, M., Srinivas, N. R., *Biomed. Chromatogr.*, 2003, 17, 423–434.
- [28] Görög, S., TrAC, *Trends Anal. Chem.*, 2006, 25, 755–757.
- [29] Holík, M., Mannschreck, A., *Chem. Intell. Lab. Syst.*, 2004, 72, 153–160.
- [30] Arancibia, J. A., Damiani, P. C., Escandar, G. M., Ibañez, G. A., Olivieri, A. C., *J. Chromatogr. B*, 2012, 910, 22–30.
- [31] Listgarten, J., Emili, A., *Mol. Cell. Proteomics*, 2005, 4, 419–434.
- [32] Haaland, D. M., Thomas, E. V., *Anal. Chem.*, 1988, 60, 1193–1202.
- [33] Olivieri, A. C., Faber, N. M., Ferré, J., Boqué, R., Kalivas, J. H., Mark, H., *Pure & Appl. Chem.*, 2006, 78, 633–661.
- [34] Olivieri, A.C., Faber, N.M., (2009) Comprehensive Chemometrics, In *Chemical and Biochemical Data Analysis* (Walczak, B., Tauler, R., Brown, S., Eds.), Elsevier, Oxford 2009, pp 91–120.
- [35] Allegrini, F. A., Olivieri, A. C., *Anal. Chem.*, 2014, 86, 7858–7866.



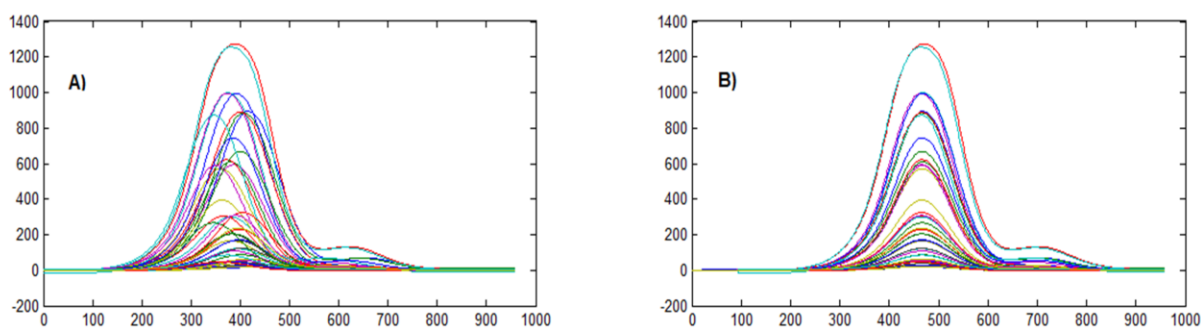
**Figure 1.** Superimposed chromatograms of racemic ketoprofen under three mobile phase conditions. Column: Nucleodex  $\beta$ -PM, temperature: 25°C. Mobile phase: methanol/buffer TEAA 0.1% pH: 4. Compositions: 55:45, 60:40 and 65:35. Flow rate: 0.6 mL min<sup>-1</sup>. Detection: 260 nm.



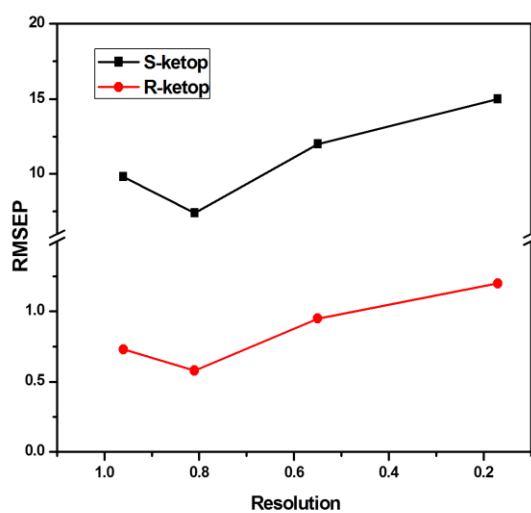
**Figure 2.** Superimposed chromatograms of partially overlapped profiles of *S*:*R* 95:5 ketoprofen. Chromatographic conditions as in Figure 1. Eluent compositions: see Table 3.



**Figure 3.** (A) Raw chromatograms at 220 nm in a selected time range for resolution M1. (B) The same chromatograms after alignment.



Accepted Article

**Figure 4.** RMSEP against resolution factor between *S*- and *R*-ketoprofen.

Accepted Article

**Table 1.** Calibration solutions sample set used and their concentrations.

Calibration simple set solutions	Composition (S:R)	Concentration of S (mg L <sup>-1</sup> )	Concentration of R (mg L <sup>-1</sup> )
1	99.5:0.5	9.75	0.05
2	99.5:0.5	9.75	0.05
3	99.5:0.5	243.87	1.32
4	99.5:0.5	243.87	1.32
5	99.5:0.5	487.75	2.65
6	99.5:0.5	487.75	2.65
7	99:1	20.60	0.21
8	99:1	20.60	0.21
9	99:1	514.95	5.35
10	99:1	514.95	5.35
11	99:1	1029.90	10.70
12	99:1	1029.90	10.70
13	95:5	32.80	1.84
14	95:5	32.80	1.84
15	95:5	820.00	46.00
16	95:5	820.00	46.00
17	95:5	1640.00	92.00
18	95:5	1640.00	92.00

**Table 2.** Test solutions set Samples used to test the calibrated model and their concentrations.

<b>Test set solutions</b>	<b>Composition (S:R)</b>	<b>Concentration S (mg L<sup>-1</sup>)</b>	<b>Concentration R (mg L<sup>-1</sup>)</b>
1	99.5:0.5	19.51	0.11
2	99.5:0.5	48.77	0.26
3	99.5:0.5	73.16	0.40
4	99.5:0.5	87.79	0.48
5	99.5:0.5	97.55	0.53
6	99.5:0.5	97.55	0.53
7	99.5:0.5	146.32	0.79
8	99.5:0.5	146.32	0.79
9	99.5:0.5	146.32	0.79
10	99.5:0.5	195.10	1.06
11	99:1	41.19	0.43
12	99:1	41.19	0.43
13	99:1	41.20	0.43
14	99:1	51.49	0.53
15	99:1	51.49	0.53
16	99:1	51.49	0.53
17	99:1	205.98	2.14
18	99:1	617.94	6.42
19	99:1	823.92	8.56
20	99:1	1029.90	10.70
21	95:5	65.60	3.68
22	95:5	164.00	9.20
23	95:5	164.00	9.20
24	95:5	196.80	11.04
25	95:5	262.40	14.72
26	95:5	262.40	14.72
27	95:5	328.00	18.40
28	95:5	492.00	27.60
29	95:5	656.00	36.80
30	95:5	820.00	46.00

**Table 3.** Characteristics of each chromatographic conditions used.

<b>Method</b>	<b>t<sub>R1</sub> (min.)</b>	<b>Mobile phase (MeOH:buffer TEAA) (%<sub>v/v</sub>)</b>	<b>Resolution factor</b>	<b>Time range (min.)</b>	<b>Original matrix size</b>	<b>Optimized time range (min.)</b>
MRs1	22	50:50	0.96	0-30	1794×102	18-27
MRs2	14	55:45	0.81	0-30	1794×102	11-18
MRs3	10	60:40	0.55	0-17	1043×102	7-15
MRs4	7	65:35	0.17	0-15	894×102	5.5-12

**Table 4.** Analytical figures of merit.

Parameter <sup>a)</sup>	S-ketoprofen					R-ketoprofen				
Condition	MRs 1 <sup>b)</sup>	MRs 1 <sup>c)</sup>	MRs 2	MRs 3	MRs 4	MRs 1 <sup>b)</sup>	MRs 1 <sup>c)</sup>	MRs 2	MRs 3	MRs 4
Number of latent variables ( <i>A</i> )	-	4	2	2	2	-	7	4	4	9
RMSEP	28	9.8	7.4	12	15	4.5	0.73	0.58	0.95	1.2
REP%	-	1.8	1.4	2.2	3.8	-	4.1	4.2	8.9	9.1
<i>R</i> <sup>2</sup>	0.997	0.996	0.998	0.995	0.996	0.982	0.997	0.997	0.988	0.982
Sensitivity (mAU L mg <sup>-1</sup> )	0.63	24	48	36	56	0.81	64	63	90	90
Analytical sensitivity (L mg <sup>-1</sup> )	0.037					0.24				
LOD (mg L <sup>-1</sup> )	74	0.3	1.0	2.4	3.3	11	0.4	0.1	1.5	2.0
LOQ (mg L <sup>-1</sup> )	220	0.9	3.0	7.3	10	33	1.2	0.3	4.5	6.1

a) RMSEP = root mean squared error of prediction, REP% = relative error of prediction, *R*<sup>2</sup> = squared correlation coefficient, analytical sensitivity = ratio of sensitivity to instrumental noise, LOD = limit of detection (in the U-PLS calibration, the maximum value of the LOD range), LOQ = limit of quantitation; LOQ= 10 × SD(0).

b) Univariate calibration.

c) U-PLS calibration.



**Table 5.** Step-by-step protocol to obtain quantitative information

1. Prepare the calibration solutions and the test solutions
2. Select chromatographic conditions, including wavelength range, number of wavelengths and data-acquisition rate
3. Run these calibration samples, and collect the exportable matrix data (txt files)
4. Align the chromatographic profiles at a given maximum time using MATLAB routine
5. Convert each calibration data matrix into a vector and build the PLS model using MATLAB and the MVC2 graphical interface
6. Estimate the number of latent variables  $A$  by using the leave-one-out method
7. Obtain the loadings and weights, as well as the regression coefficients for predictions
8. Run the test solutions, collecting the matrix data as txt files, align the test solution profiles
9. Estimate of analyte concentrations from the vectorized data for the test samples
10. Estimate the analytical figures of merit

**Table 6.** Prediction results for the real sample using U-PLS calibration at the three best resolutions and its statistical analysis. ( $C_{nom}$  = nominal concentration,  $C_{found}$  = found concentration and  $E_r$  = relative error).

$C_{nom}$ (mg L <sup>-1</sup> )	$C_{found}$ MRs1 (mg L <sup>-1</sup> )	$C_{found}$ MRs 2 (mg L <sup>-1</sup> )	$C_{found}$ MRs 3 (mg L <sup>-1</sup> )	$E_r$ MRs1	$E_r$ MRs2	$E_r$ MRs3
326.6	347.3	331.4	348.8	20.7	4.8	22.2
328.2	337.3	328.4	350.3	9.1	0.2	22.1
653.2	660.7	646.0	695.1	7.5	-7.2	41.9
656.4	643.6	642.5	701.8	-12.8	-13.9	45.4
979.8	992.8	964.1	1066.6	13.0	-15.7	86.8
984.6	1006.3	969.7	1063.6	21.7	-14.9	79.0
Mean difference				9.9	7.8	49.6
Standard deviation of differences				12.5	8.6	27.7
$t_{exp}$				1.9	2.2	4.4
$t_{crit}$				2.58	2.58	2.58