

## Simultaneous Determination and Classification of Riboflavin, Thiamine, Nicotinamide and Pyridoxine in Pharmaceutical Formulations, by UV-Visible Spectrophotometry and Multivariate Analysis

Pedro L. López-de-Alba,<sup>\*,a</sup> Leticia López-Martínez,<sup>a</sup> Victor Cerdá<sup>b</sup> and Judith Amador-Hernández<sup>c</sup>

<sup>a</sup>Instituto de Investigaciones Científicas, Universidad de Guanajuato, 36000 Guanajuato Gto., México

<sup>b</sup>Departamento de Química, Universidad de las Islas Baleares, 07071 Palma de Mallorca, España

<sup>c</sup>Instituto de Ecología, Universidad del Mar, 70902 Puerto Ángel Oax., México

*Soft Independent Modeling of Class Analogy (SIMCA) e Regressão por Mínimos Quadrados Parciais foram usados nesse trabalho para a identificação e quantificação de tiamina, riboflavina, nicotinamida e piridoxina, por espectrofotometria UV-Vis, sem realizar os procedimentos analíticos de separação e pré-concentração. Para a quantificação, os intervalos de trabalho estabelecidos foram 1-14 mg L<sup>-1</sup> para riboflavina, 2-26 mg L<sup>-1</sup> para tiamina, 2-30 mg L<sup>-1</sup> para nicotinamida, e 2-22 mg L<sup>-1</sup> para piridoxina. Os resultados de recuperação obtidos foram superiores a 95% em todos os casos, para a análise de amostras sintéticas e comerciais. Na busca de cada vitamina alvo, foi construído um modelo de classificação com duas categorias: (i) com a vitamina de interesse e (ii) sem a vitamina de interesse. A capacidade discriminatória de cada modelo de classificação foi avaliada para aprendizagem, para um conjunto de amostras teste e para amostras comerciais, com resultados satisfatórios, com exceção da riboflavina. Assim, um método simples e confiável é proposto para a estimativa simultânea desses compostos.*

Soft Independent Modeling of Class Analogy and Partial Least Squares Regression were used in this work for the identification and quantification of thiamine, riboflavin, nicotinamide and pyridoxine by UV-Vis spectrophotometry, without separation or preconcentration steps in the analytical procedure. For quantitative purposes, the working range established was 1-14 mg L<sup>-1</sup> for riboflavin, 2-26 mg L<sup>-1</sup> for thiamine, 2-30 mg L<sup>-1</sup> for nicotinamide, and 2-22 mg L<sup>-1</sup> for pyridoxine. Recovery results higher than 95% were obtained in all cases during the analysis of synthetic and commercial samples. In the screening of each target vitamin, a classification model was built with two classes: (i) with the vitamin of interest, and (ii) without it. The discriminate capability of each classification model was evaluated for learning, independent testing and commercial samples, resulting in satisfactory findings with exception of riboflavin. Thus, a simple and reliable method is proposed for the simultaneous estimation of these compounds.

**Keywords:** vitamins, thiamine, riboflavine, pyridoxine, nicotinamide, SIMCA, PLS

### Introduction

Vitamins are essential compounds in living systems which differ in their chemical structure and physiological action. Analytical methods have been developed for vitamin identification and/or quantification, using a wide variety of strategies. Thus, food and pharmaceutical industries have taken advantage of these reliable methods and used them for the estimation of vitamins from simple to complex matrices.

Liquid chromatography has been widely employed for vitamins determination,<sup>1-3</sup> in spite of the fact that the sample preparation usually requires laborious and time-consuming steps.<sup>4</sup> Other chromatographic techniques such as micellar electrokinetic capillary chromatography have been reported with satisfactory results.<sup>5,6</sup> Some analytical methods based on UV-Vis spectrophotometry and spectrofluorimetry have also been optimized for the simultaneous quantification of vitamins. However, these methods usually require some chemical reaction or separation steps,<sup>7-9</sup> where sample manipulation can

\* e-mail: judith@angel.umar.mx, amadorjudith@yahoo.com

increase the risk of human error in the results and the cost of the analysis.

Nowadays, chemometrics is considered a suitable alternative to analytical separation procedures. Selectivity limitations presented by optical detection techniques can be overcome with the application of some mathematical algorithms during the interpretation of instrumental data.<sup>10,11</sup> Therefore, different research groups have proposed the use of chemometric strategies for the simultaneous determination of vitamins. As an example, there are several reports about the quantification of vitamins in pharmaceutical formulations by derivative spectrophotometry.<sup>12-14</sup> Multiple linear regression and artificial neural networks were also applied to resolve overlapping spectrophotometric signals of multivitamin samples.<sup>15,16</sup>

Partial Least Squares (PLS) is a multivariate calibration technique used in the resolution of multicomponent systems.<sup>17-19</sup> Berzas *et al.*<sup>20</sup> reported the fluorimetric determination of pyridoxal, pyridoxamine and pyridoxic acid at low concentration levels, using non-linear variable angle synchronous spectra and PLS. Collado *et al.*<sup>21</sup> applied PLS to quantify nicotinamide and inosine in ophthalmic solutions by UV-Vis spectrophotometry with good results. The application of the PLS algorithm was an excellent tool to eliminate the spectral interferences in the quantification of the analytes of interest, which was satisfactorily demonstrated by Ghasemi and Vosough<sup>22</sup> and Aberásturi *et al.*<sup>23</sup> in the resolution of four-component mixtures of vitamins.

Soft Independent Modeling of Class Analogy (SIMCA) is a supervised pattern recognition technique used for the reliable classification of unknown samples which is based on Principal Component models for the training of set categories.<sup>24</sup> The flow injection screening of PAHs in water by laser induced spectrofluorimetry,<sup>25</sup> the identification of pharmaceutical excipients by NIR reflectance spectroscopy and the confirmation of the authenticity of Galician wines from the Ribeira Sacra area are some successful applications of SIMCA for sample classification.<sup>26,27</sup>

Excellent monographs have been published, which discuss the principles and algorithms of PLS and SIMCA, such as interesting applications.<sup>28-30</sup> In the present study, SIMCA and PLS have been used to identify and quantify B-group vitamins (thiamine, riboflavine, nicotinamide and pyridoxine) without separation or preconcentration steps. In mixtures with one to four components, spectral interferences among the target substances were avoided by using chemometric tools. Finally, the predictive ability of classification and calibration models was evaluated in commercial formulations and showed satisfactory results.

## Experimental

### Apparatus

A Spectronic 3000 Diode Array Milton Roy spectrophotometer with a resolution of 0.35 nm was used, coupled to a 486 PC. A User Data version 2.01 Milton Roy Inst. Software was employed for spectral data acquisition, storage, and manipulation. Data treatment was carried out using a Pentium IV PC equipped with the GRAMS /386 tm software package, version 3.01A (Galactic, USA) and the Pirouette software package, V. 3.1 (Infometrix, USA).

### Reagents

All chemicals were of analytical reagent grade. Thiamine hydrochloride (THIA), riboflavine (RIB), nicotinamide (NIC), and pyridoxine (PYR) were obtained from Aldrich. Water purified with a Milli-Q system was used throughout.

Standard solutions of THIA (1000 mg L<sup>-1</sup>), NIC (1000 mg L<sup>-1</sup>), PYR (1000 mg L<sup>-1</sup>), and RIB (50 mg L<sup>-1</sup>) were prepared by dissolving the appropriate amounts of each analytical reagent in pure water. The solutions were stored and protected from light at 4 °C. Working standard solutions were prepared daily by appropriate dilution. A buffer solution of monochloroacetic acid/potassium hydroxide (pH 2.2; 0.1 mol L<sup>-1</sup>) was also prepared, adjusting the pH with hydrochloric acid.

For the analysis of commercial samples, BEPLEX 50 from Laboratorios Carnot de Productos Científicos S.A. (Product A), TIAMINAL TRIVALENTE from Laboratorios Silanes S.A. (Product B) and ANEREX from Aplicaciones Farmacéuticas S.A. (Product C) were used.

### Procedure

The working solutions were prepared by adding adequate volumes of the stock vitamin solutions and 5 mL of buffer solution in 25 mL volumetric flasks and filled up with pure water. The linear ranges of work were of 2-26 mg L<sup>-1</sup> for THIA, 1-14 mg L<sup>-1</sup> for RIB, 2-30 mg L<sup>-1</sup> for NIC, and 2-22 mg L<sup>-1</sup> for PYR. The samples were stable for at least 90 min protected from light, according to a stability study in the proposed experimental conditions. The absorption spectra were recorded from 200 to 400 nm against a blank, with a resolution of 0.35 nm, and smoothed using the Savitzky and Golay procedure (25 experimental points, moving average function).

For the application of SIMCA and PLS, a training set of 44 samples was used (see Table 1). A test set of 16

samples was also used (Table 2), to evaluate the qualitative and quantitative capabilities of the proposed models.

The analyses of pharmaceutical preparations were carried out by triplicate, weighing homogeneous portions of the multivitamin samples in powder (about 1 g) and dissolving them with pure water in volumetric flasks. Then each solution was maintained in an ultrasonic bath for a period of 15 min. In order to remove suspended particles, a fraction of the solution was centrifuged at 3500 rpm for 15 min. Finally, aliquots of the supernatant were analyzed by using the procedure described above.

## Results and Discussion

Absorption spectra of thiamine, riboflavine, nicotinamide and pyridoxine are shown in Figure 1. There is a substantial overlapping on the spectra from 200 to 400 nm. Therefore, the simultaneous spectrophotometric identification or quantification of the vitamins of interest requires the application of a separation step before their detection, or the use of a chemometric algorithm for the resolution of multicomponent mixtures.

In order to develop a new spectrophotometric method for the determination of B-group vitamins by SIMCA and PLS, the influence of pH in absorption spectra was studied. It is well known that THIA and RIB are not stable in alkaline conditions. Therefore, solutions with these vitamins in KCl 0.1 mol L<sup>-1</sup> were prepared for each one, with pH values between 1.0 and 8.0 adjusted with HCl or NaOH.

As an example, Figure 2 shows the influence of pH in PYR spectra. Four absorption bands were clearly observed; the maxima were located at 220, 254, 291 and 324 nm. The bands between 200-210 were not considered, owing to major interferences from electrolytes could be expected. With exception of the band centered at 291 nm (third, from left to right), the rest of absorption bands showed a hyperchromic effect while the pH decreased. Contrarily, the band located between the isosbestic points at 267 and 305 nm (with a maximum in 291 nm), showed a hyperchromic effect with the increase of the pH.

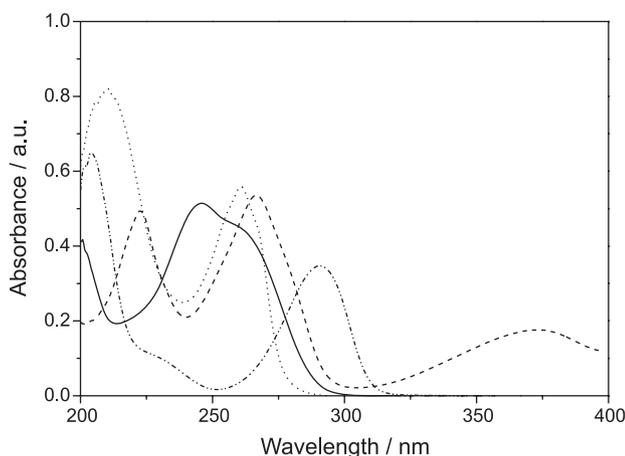
It was desirable to obtain mayor differences between the spectral shapes of the analytes to resolve the multicomponent system by multivariate strategies. Taking into account this premise and the fact that THIA and RIB were not stable at basic conditions, a pH 2.2 was selected as optimum. At this pH value, a major differentiation between the absorption bands of the four compounds were observed, since pyridoxine shows only one absorption band (maximum at 291 nm) instead of three (maxima at 220, 254, and 324). A buffer solution of monochloroacetic acid/potassium hydroxide 0.1 mol L<sup>-1</sup> provided an adequate buffering capacity.

**Table 1.** Composition of the training set of samples used in the multivitamin analysis (concentrations in mg L<sup>-1</sup>)

Sample	THIA	RIB	NIC	PYR
1	0	1	0	11
2	2	7	0	0
3	0	0	30	2
4	13	0	2	0
5	0	1	30	6.5
6	2	4	0	11
7	7.5	7	2	0
8	13	0	16	2
9	2	4	2	0
10	13	0	30	6.5
11	13	0	30	11
12	7.5	0	2	2
13	0	4	30	11
14	2	1	0	6.5
15	13	7	16	0
16	0	7	30	11
17	0	1	16	2
18	7.5	7	0	11
19	13	7	0	11
20	13	7	30	0
21	7.5	7	16	2
22	2	7	2	2
23	7.5	7	16	6.5
24	13	1	16	6.5
25	7.5	4	2	11
26	13	1	2	2
27	13	4	16	6.5
28	2	4	30	6.5
29	2	1	30	2
30	7.5	4	30	6.5
31	2	1	2	11
32	7.5	4	16	11
33	1	0	0	0
34	7	0	0	0
35	13	0	0	0
36	0	1	0	0
37	0	4	0	0
38	0	7	0	0
39	0	0	2	0
40	0	0	16	0
41	0	0	30	0
42	0	0	0	2
43	0	0	0	6.5
44	0	0	0	11

### Application of PLS to the multicomponent system

Due to the complexity of the system, a large number of training samples were necessary. Firstly, typical one-compound calibration experiments were carried out to establish the concentration ranges for the determination. Linearity was observed between 2-26 mg L<sup>-1</sup> for THIA, 1-14 for RIB, 2-30 mg L<sup>-1</sup> for NIC and 2-22 mg L<sup>-1</sup> for PYR. However, the concentration of THIA, RIB and PYR used in multicomponent samples were below or in the middle of their linear calibration curves; this latter was to avoid an excessive absorbance of the mixtures. In contrast, a range from 2 to 30 mg L<sup>-1</sup> of NIC was



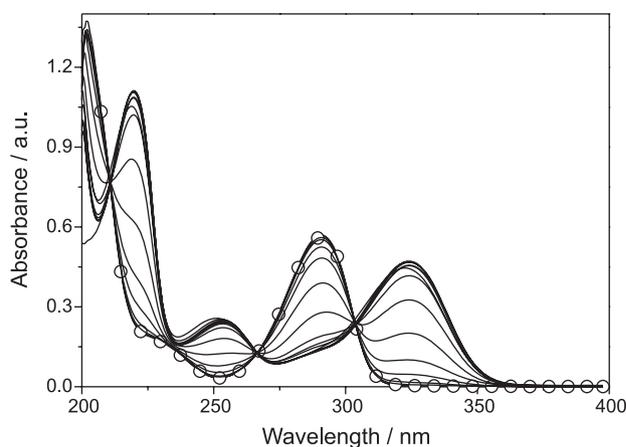
**Figure 1.** Absorption spectra of THIA (13 mg L<sup>-1</sup>, solid line), RIB (7 mg L<sup>-1</sup>, dashed line), NIC (15 mg L<sup>-1</sup>, dotted line) and PYR (11 mg L<sup>-1</sup>, dashed dotted line), under the proposed experimental conditions (pH 2.2).

**Table 2.** Composition of test samples used for the multivitamin analysis (concentrations in mg L<sup>-1</sup>)

Sample	THIA	RIB	NIC	PYR
1	0	5	0	0
2	3	0	0	0
3	0	0	25	0
4	0	0	0	4
5	0	0	8	7
6	9	3	0	0
7	0	6	0	5
8	6	0	25	0
9	4	0	21	3
10	0	2	11	8
11	5	6	0	10
12	10	5	19	0
13	9	2	6	6
14	8	3	11	9
15	4	5	24	7
16	12	6	18	3

considered in calibration samples, since this vitamin is commonly present at higher concentrations than the others in pharmaceutical formulations.

The composition of the training set of 44 samples is shown in Table 1, which corresponds to some practical guidelines given by several reports.<sup>17-19,28</sup> According to the experience of the authors, the incorporation of approximately ten samples *per* component leads to a satisfactory prediction capability of the calibration model in a complex system. Also, it is necessary to include samples with different ratios between components, to incorporate as much variability as possible to the system. Finally, the incorporation of samples with only one component at different levels of concentrations increases the efficiency of the model.<sup>17</sup> For this multivitamin system, samples of one to four components are included in different ratios,



**Figure 2.** Influence of pH in the absorption spectra of PYR (12 mg L<sup>-1</sup>) in KCl 0.1 mol L<sup>-1</sup>, from 1.0 to 8.0 pH units. The spectrum marked with open circles indicates the pH condition selected in this work.

due to the fact that the composition of commercial samples shows a wide variety. Spectra did not exceed absorbance values higher than 1.2 in all cases. The external validation of calibrations models was carried out using the set of 16 test samples described in Table 2.

For the resolution of four-vitamin mixtures, Aberásturi *et al.*<sup>23</sup> used an experimental design of two levels plus one central point ( $4^2 + 1$ ) for calibration matrix, while Ghasemi and Vosough<sup>22</sup> proposed a calibration set of solutions in which absorbances did not exceed a value of 1.0. In the first work, there were not considered samples without at least one of the components of interest, while in the work of Ghasemi and Vosough were considered. Satisfactory prediction capabilities of PLS models were observed in both cases, although the sets of validation (synthetic mixtures) included samples with all components. However, pharmaceutical formulations did not always contain the four vitamins; therefore, in the present work were considered samples with two to four-components for calibration and validation.

Some exploratory analyses based on PCA were carried out to evaluate the quality of the spectral information provided. The original spectral range from 200 to 400 nm was reduced to the region of 215 to 310 nm, to avoid noise and irrelevant information in the numerical analysis. Mean centered data was used as independent variables. The Mahalanobis distance and sample residual tools were applied in the evaluation of outliers; no outliers were identified.

Under consideration of the training set of samples and the newly selected spectral region, the leave-one-out cross validation was developed for each one of the target compounds. The PRESS (Prediction Error Sum of Squares) vs. number of factors (or principal components) plot was constructed for each vitamin. After that, the criteria of the first local minimum value of PRESS<sup>18</sup> and

an F-statistic comparison of PRESS<sup>28</sup> (probability of 0.75) were used for the selection of the number of factors required for the construction of calibration models. The PRESS was estimated according to:

$$\text{PRESS} = \sum_{i=1}^I (c_i - \hat{c}_i)^2 \quad (1)$$

where  $c_i$  is the real concentration (or amount) of the analyte in sample  $i$ , and  $\hat{c}_i$  corresponds to the estimated concentration (or amount) by using the proposed calibration model,  $I$  is the overall number of samples. Other tools offered by GRAMS Software (e.g., PLS loading factors) were also considered, to avoid sub-or overfitting of models due to the inclusion of non-representative factors in their construction.

The internal cross validation of the proposed models was performed while taking into account several statistical parameters.<sup>28</sup> The  $R^2$  (square of correlation coefficient), which indicates the quality of data adjustment between real and estimated concentrations, was calculated according to:

$$R^2 = \frac{\sum_{i=1}^I (\bar{c} - c_i)^2}{\sum_{i=1}^I (\bar{c} - \hat{c}_i)^2} \quad (2)$$

where  $\bar{c}$  is the average concentration of the overall samples. SEC(P), the standard error of calibration or prediction, whatever the case, was evaluated by:

$$\text{SEC(P)} = \sqrt{\frac{\sum_{i=1}^I (c_i - \hat{c}_i)^2}{I-1}} \quad (3)$$

Similarly, the root mean square difference (RMSD), an average error index in the analysis, was calculated according to:

$$\text{RMSD} = \sqrt{\frac{\sum_{i=1}^I (c_i - \hat{c}_i)^2}{I}} \quad (4)$$

The REP (%), that is the relative error of prediction, which is considered an error average percentage in the set of samples, was established by:

$$\text{REP}(\%) = \frac{100}{\bar{c}} \sqrt{\frac{\sum_{i=1}^I (c_i - \hat{c}_i)^2}{I}} \quad (5)$$

A summary of these results is shown in Table 3, where  $r^2 > 0.999$  were obtained in all cases. Comparison of estimated statistical parameters after considering both criteria for the selection of factors (first local minimum and F-test), showed that the use of more than four factors increased the risk of overfitting, since improvement of the results was not representative; additionally, the introduction of less factors in multivariate models reduced their prediction capabilities. According to Table 3, satisfactory results were obtained using four factors in each case; error average percentages smaller than 3% were found for all cases.

The correlation plot revealed that the spectral region selected was the appropriate for the quantification of vitamins in mixtures (values higher than an absolute 0.5). In loadings plot, data features were in accordance with spectral shapes of individual components, for the factors selected as optima.

The prediction capabilities of the optimum PLS models were also tested by means of the set of samples described in Table 2. Recovery results expressed as percentages, as well as REP (%) and SEP (standard error of prediction) were estimated as part of the external validation, whose results are included in Table 4. Since the standard error rates of calibration and prediction were similar, no sub-or overfitting problems were observed.

Later, the analyses of pharmaceutical formulations were carried out. The proposed method was applied for the determination of THIA, RIB, NIC and PYR in three commercial samples (Products A, B and C). The results are shown in Table 5. Data show that the estimated concentration of vitamins agreed satisfactorily with product labels. Other substances which were not considered in the calibration model, e.g., cyanocobalamine, ascorbic acid, excipient, etc., did not interfere.

**Table 3.** Statistical parameters estimated during the optimization of PLS calibration models (internal validation), with the application of two criteria for the selection of the optimum number of factors

Compound	Criterion of first local minimum PRESS				F-statistical comparison of PRESS			
	Number of factors	RMSD	SEC	REP (%)	Number of factors	RMSD	SEC	REP (%)
THIA	4	0.18	0.19	0.9	9	0.12	0.12	2.4
RIB	4	0.05	0.05	1.8	5	0.04	0.04	0.1
NIC	4	0.32	0.33	2.9	6	0.14	0.14	1.3
PYR	4	0.07	0.07	1.5	7	0.04	0.04	1.0

**Table 4.** Statistical parameters estimated during the external validation of the proposed PLS calibration models

Compound	Number of factors	Recovery <sup>a</sup>	SEP	REP (%)
THIA	4	100 ± 2	0.11	2.4
RIB	4	100 ± 2	0.06	2.1
NIC	4	99 ± 1	0.23	2.1
PYR	4	99 ± 1	0.07	1.8

<sup>a</sup>Mean recovery percentage ± standard deviation.

### Classification of vitamins by SIMCA

Nowadays, the importance of screening methods is increasing. In many cases, the pre-classification of samples before the use of chromatographic techniques substantially reduces the time and cost of the overall analysis.

In the case of this study, the pattern recognition strategy known as SIMCA was applied to identify the presence or absence of THIA, NIC, PYR, or RIB in commercial samples, when qualitative and not quantitative estimation was of interest. To start with, the training set used for PLS models was also considered for classification by SIMCA. For each vitamin, all the learning samples (Table 1) were divided into two classes: (i) with and (ii) without the vitamin of interest. One classification model was made for each compound. For example, in the classification model of PYR, samples 1, 3, 5-6, 8, 10-14, 16-19, 21-32, 42-44, were considered class I (with PYR), while samples 2, 4, 7, 9, 15, 20, 33-41, were assigned to class II (without PYR). In contrast, for the classification model of THIA, class I included the samples 2, 4, 6-12, 14-15, 18-35, while class II considered samples 1, 3, 5, 13, 16, 17, 36-44.

Also, absorption spectra from 200 to 400 nm were assigned as independent variables. Preliminary results demonstrated a poor discriminant capability of models regarding the nature of samples. As a consequence, several strategies were applied.

The learning set was complemented with 12 samples (three samples for each one of the single analytes). For feature selection, modeling and discrimination power were calculated; as a result, variables in the range from 215 to 310 nm were also found convenient to improve class

identification. The relevance of a pretreatment strategy on a variable-basis was evaluated in the cases of the study, using original and mean centered data. Outlier diagnostics was based on Mahalanobis distance of sample residuals.

Later, the classification rules for THIA, RIB, NIC and PYR were estimated by SIMCA. In each case, the optimum number of factors for classes (i) with and (ii) without the target vitamin was determined by considering the variance related to each factor. Also, the interclass residuals and interclass distances were estimated for the selection of the proper factors.<sup>24</sup>

For example, an interclass residual for class *i*, can be defined as:

$$s_{12} = \sqrt{\frac{1}{(m-k_2)n_1} \sum_i^{n_1} e_i e_i^T} \quad (6)$$

where *m* is the number of original variables; *k*<sub>2</sub> is the number of factors in the class *ii* model; *n*<sub>1</sub> is the number of class *i* samples; *e*<sub>*i*</sub> is the row vector of residuals for class *i* samples, *i.e.* the difference between original data and its *k* factor estimate.

The interclass distance is defined as:

$$D_{12} = \sqrt{\frac{s_{12}^2 + s_{21}^2}{s_{11}^2 + s_{22}^2}} - 1 = D_{21} \quad (7)$$

where *s*<sub>11</sub> denotes the residuals of fitting class *i* samples to the class *i* model; *s*<sub>22</sub> corresponds to the residuals of fitting class *ii* samples to the class *ii* model; *s*<sub>12</sub> and *s*<sub>21</sub> represent the residuals of fitting class *i* samples to the class *ii* model and *vice-versa*. Some of the results obtained during internal validation of classification models are summarized in Table 6. It shows that more samples of both classes (*i* and *ii*) were correctly classified starting from original data than with pretreatment steps, allowing for the identification of fewer false positives and negatives. Cumulative variance was in all cases > 99%.

External validation of the classification rules estimated with SIMCA was carried out with the indepen-

**Table 5.** Results of the analyses of pharmaceutical formulations by UV/Vis spectrophotometry and PLS (content in mg)

Compound	Content of Product A		Content of Product B		Content of Product C	
	Expected	Founded <sup>a</sup>	Expected	Founded <sup>a</sup>	Expected	Founded <sup>a</sup>
THIA	50	52.0 ± 0.2	100	112.4 ± 0.8	24.15	23.6 ± 0.2
RIB	10	11.5 ± 0.1	-	-	-	-
NIC	50	56.1 ± 0.3	-	-	-	-
PYR	5	5.0 ± 0.1	50	44.3 ± 0.4	16.06	18.2 ± 0.2

<sup>a</sup>Mean predicted value ± standard deviation.

**Table 6.** Results of the internal validation of classification models obtained by SIMCA

Compound	Factors selected		Samples correctly classified (%) <sup>a</sup>	
	OD <sup>b</sup>	MCD <sup>c</sup>	OD	MCD
THIA	3; 3	4; 4	87; 93	77; 81
RIB	3; 4	4; 4	94; 100	78; 85
NIC	3; 4	4; 4	94; 92	72; 88
PYR	4; 4	4; 4	97; 96	68; 85

<sup>a</sup>Values separated by semi-colon correspond to classes *i* and *ii*, respectively; <sup>b</sup>OD: original data; <sup>c</sup>MCD: mean centered data.

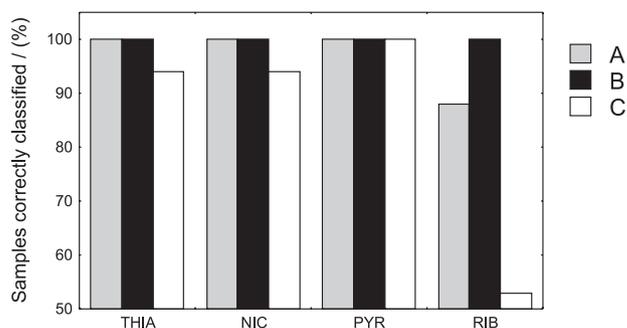
dent test samples described in Table 2. Interclass distances and the percentage of correctly classified samples in categories *i* and *ii* are shown in Table 7. It demonstrates that better results were obtained using data without pretreatment in all cases, which confirm the tendency observed during internal validation. Although mean centering is recommended for most spectral data, preprocessing increases the influence of outliers. In these studies no outliers were identified, but the large variability in the data probably caused the phenomena that were observed. Therefore, models without pretreatment strategies were proposed for the identification of the analytes in further studies.

Finally, classification rules established by SIMCA were applied for the qualitative determination of B-group vitamins in the pharmaceutical products A, B and C. Additionally, some of these samples were spiked with the vitamins of interest in different ratios, mainly in those cases which original composition not include some of the analytes. Thus, a set of 20 samples for each commercial product (spiked or not) were evaluated. The percentage of correctly classified samples in categories *i* and *ii* (which denotes the presence or absence of each of the target compounds) are represented in Figure 3. As can be observed, satisfactory results were obtained in almost all cases. However, the identification of riboflavin in product C gave rise to false positives (54% of samples belonging to group *i*). The presence of additional components in the sample matrix could produce interferences that could not

**Table 7.** Parameters estimated during the external validation of classification models derived from SIMCA

Compound	Interclass distance		Samples correctly classified (%) <sup>a</sup>	
	OD <sup>b</sup>	MCD <sup>c</sup>	OD	MCD
THIA	1.9	0.5	80; 100	90; 33
RIB	2.1	0.9	90; 67	70; 33
NIC	2.56	0.2	100; 83	70; 67
PYR	5.5	0.4	100; 100	60; 33

<sup>a</sup>Values separated by semi-colon correspond to classes *i* and *ii*, respectively; <sup>b</sup>OD: original data; <sup>c</sup>MCD: mean centered data.

**Figure 3.** Total percentage of samples belonging to the pharmaceutical formulations A, B and C, correctly classified in categories *i* and *ii*.

be excluded by the chemometric algorithm. For the rest of the products and analytes, no false positives or negatives were obtained. The remaining samples were simply did not fit in any category.

## Conclusions

A simple and reliable spectrophotometric method is proposed in this work for the analysis of multi-component samples of vitamins. The construction of calibration models with four factors for THIA, RIB, NIC and PYR did not lead to sub- or overfitting problems; statistical parameter values obtained during their internal and external validation were satisfactory in all cases. As an example, binary and quaternary mixtures of the vitamins of interest (in the presence of additional substances) were accurately resolved. The results obtained from the analyses of commercial samples confirm the predictive ability of the method to eliminate spectral interferences without use of separation steps during the analytical procedure, and the flexibility of the method to analyze samples with a variable number of vitamins as part of their composition (from single to four-component mixtures).

On the other hand, the qualitative determination of THIA, NIC, PYR and RIB was satisfactorily carried out by SIMCA. Three and four factors were necessary for the construction of classification models. Large interclass distances obtained during internal and external validation showed a good separation between classes. With exception of RIB, all vitamins were properly identified in commercial samples and no false positives or negatives were observed. Again, chemometric techniques such as SIMCA and PLS proved to be powerful tools to reduce sample pretreatment steps which are traditionally required for the proper determination of multiple components in real samples.

## Acknowledgments

The authors gratefully acknowledge financial support from CONACyT (México), projects 21118-E and J37143-E.

## References

1. Gámiz-Gracia, L.; Velasco-Arjona, A.; Luque de Castro, M.D.; *Analyst* **1999**, *124*, 801.
2. Dinc, E.; Kokdil, G.; Onur, F.; *J. Pharm. Biomed. Anal.* **2000**, *22*, 915.
3. Vinas, P.; Campillo, N.; Garcia, I.L.; Cordoba, M.H.; *Food Chem.* **1992**, *45*, 349.
4. Analytical Methods Committee; *Analyst* **2000**, *125*, 353.
5. Blanco, D.; Laviana, L.; Gutiérrez, D.; *Anal. Chim. Acta* **1999**, *396*, 55.
6. Hu, Q.; Zhou, T.; Zhang, L.; Li, H.; Fang, Y.; *Anal. Chim. Acta* **2001**, *437*, 123.
7. Monferrer-Pons, L.; Alvarez-Rodríguez, L.; Esteve-Romero, J.; *Anal. Lett.* **1999**, *32*, 51.
8. Zhihong, L.; Wang, Q.; Mao, L.; Cai, R.; *Anal. Chim. Acta* **2000**, *413*, 167.
9. Rocha, F.R.P.; Filho, O.F.; Reis, B.F.; *Talanta* **2002**, *59*, 191.
10. López-de-Alba, P.L.; López-Martínez, L.; Michelini-Rodríguez, L.I.; Wróbel, K.; Wróbel, K.; Amador-Hernández, J.; *Analyst* **1997**, *122*, 1575.
11. Amador-Hernández, J.; Fernández-Romero, J.M.; Luque de Castro, M.D.; *Anal. Chim. Acta* **2001**, *435*, 227.
12. Consiglieri, V.O.; Vals, N.R.M.; Magalhaes, J.F.; *Anal. Lett.* **2001**, *34*, 1875.
13. Dinc, E.; Kokdil, G.; Onur, F.; *J. Pharm. Biomed. Anal.* **2000**, *22*, 915.
14. Elwalily, A.F.; Elanwar, F.; Zamel, S.; *Anal. Chim. Acta* **1991**, *248*, 583.
15. Blanco, M.; Coello, J.; Iturriaga, H.; MasPOCH, S.; Gomezcotin, T.; Alaouiismaili, S.; Rovira, E.; *Fresenius J. Anal. Chem.* **1995**, *351*, 315.
16. Yin, C.S.; Shen, Y.; Liu, S.S.; Li, Z.L.; Pan, Z.X.; *Chem. J. Chin. Univ. Chin.* **2000**, *21*, 49.
17. López-de-Alba, P.L.; Wróbel-Kaczmarczyk, K.; Wróbel, K.; López-Martínez, L.; Amador-Hernández, J.; *Anal. Chim. Acta* **1996**, *330*, 19.
18. Amador-Hernández, J.; Cladera, A.; Estela, J.M.; López-de-Alba, P.L.; Cerdá, V.; *Analyst* **1998**, *123*, 2235.
19. Amador-Hernández, J.; García-Ayuso, L.E.; Fernández-Romero, J.M.; Luque de Castro, M.D.; *J. Anal. At. Spectrom.* **2000**, *15*, 587.
20. Berzas, J.J.; Murillo, J.A.; Amador-Hernández, J.; Gómez, M.A.; *Analyst* **1998**, *123*, 483.
21. Collado, M.S.; Mantovani, V.E.; Goicoechea, H.C.; Olivieri, A.C.; *Anal. Lett.* **2001**, *34*, 363.
22. Ghasemi, J.; Vosough, M.; *Spectrosc. Lett.* **2002**, *35*, 153.
23. Aberásturi, F.J.; Jiménez, A.I.; Arias, J.J.; Jiménez, F.; *Anal. Lett.* **2002**, *35*, 1677.
24. Infometrix Inc.; *Pirouette, Multivariate Data Analysis for Windows 95/98 and NT, User's Manual*, V. 3.1; Woodinville WA, 2002.
25. Amador-Hernández, J.; Fernández-Romero, J.M.; Luque de Castro, M.D.; *Anal. Chim. Acta* **2001**, *448*, 61.
26. Krämer, K.; Ebel, S.; *Anal. Chim. Acta* **2000**, *420*, 155.
27. Reboló, S.; Peña, R.M.; Latorre, M.J.; García, S.; Botana, A.M.; Herrero, C.; *Anal. Chim. Acta* **2000**, *417*, 211.
28. Martens, H.; Naes, T.; *Multivariate Calibration*, Wiley: New York, 1989.
29. Massart, D.L.; Vandeginste, B.G.M.; Buydens, L.M.C.; De Jong, S.; Lewi, P.J.; Smeyers-Verbeke, J.; *Handbook of Chemometrics and Qualimetrics: Parts A and B*, Elsevier: Amsterdam, 1997.
30. Sharaf, M.A.; Illman, D. L.; Kowalski, B.R.; *Chemometrics*, Wiley: New York, 1986.

Received: April 26, 2005

Published on the web: May 26, 2006