JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Rapid Determination of Crocetin Esters and Picrocrocin from Saffron Spice (*Crocus sativus* L.) Using UV–Visible Spectrophotometry for Quality Control

ANA M. SÁNCHEZ,[†] MANUEL CARMONA,[†] AMAYA ZALACAIN,[†] JOSÉ M. CAROT,[‡] JOSÉ M. JABALOYES,[‡] AND GONZALO L. ALONSO*,[†]

Cátedra de Química Agrícola, ETSI Agrónomos, Universidad Castilla-La Mancha, 02071 Albacete, Spain, and Departamento de Estadística e Investigación Operativa Aplicadas y Calidad, ETSI Industriales, Universidad Politécnica de Valencia, 46022 Valencia, Spain

The aim of this work was the development of multivariate models able to determine the content of the main crocetin esters and picrocrocin from spectrophotometric data that could be used for routine quality control of saffron. These compounds were determined with HPLC in Spanish saffron, and their absorbance spectra from 190 to 700 nm were simultaneously monitored. Partial least-squares regression (PLSR) models have been obtained and applied to the determination of individual crocetin esters, to the sum of crocetin esters, and to picrocrocin. A modification of the Kennard−Stone algorithm was used to divide the pool of samples into calibration and validation subsets. The best predictions were obtained with the sum of crocetin esters model, followed by the model for *cis*-crocetin (β-D-glucosyl)-(β-D-gentiobiosyl) ester, *trans*-crocetin di-(β-D-gentiobiosyl) ester, and *trans*-crocetin (β-D-glucosyl)-(β-D-gentiobiosyl) ester, whereas the worst predictions were found with the picrocrocin and *trans*-crocetin (β-D-gentiobiosyl) ester models. These models may enhance quality control in saffron enterprises.

KEYWORDS: Saffron; partial least-squares regression (PLSR); crocetin esters; picrocrocin; UV–vis spectrophotometry

INTRODUCTION

Due to consumer demands and strict regulatory requirements for food quality and safety, quality control is of utmost importance in the entire food sector. Because spice commerce in general, and saffron (the dried stigmas of *Crocus sativus* L.) commerce in particular, are no exceptions to these controls, the development of simple and quick methods of quality control and the application of chemometric tools to treat spectroscopic information is essential for medium- and small-size companies that make up this sector. Moreover, their potential use in a large number of samples without the necessity of additional instruments is of special interest.

Current international research on saffron quality is basically centered on the characterization of its attributes and their subsequent deterioration (*1–4*), the detection and quantification of food and nonfood colorings (*5*), and the revision of the standards used to certify saffron in the international saffron trade (*6*). In recent years, there has also been a growing concern for guaranteeing and defending the quality of saffron historically produced in certain regions (*7–9*). Saffron quality in the food industry is and has been mainly determined by specifications recommended by the ISO 3632, the latest revision of which has given rise to the Technical Specification ISO 3632/TS (*6*). This classifies saffron into three categories with regard to a large number of physical and chemical parameters that define saffron quality: microscopic characteristics, presence of flower waste, moisture and volatile matter content, ash content, $E^{1\%}_{1\,cm}$ 257 nm, $E^{1\%}_{1\,cm}$ 330 nm, $E^{1\%}_{1\,cm}$ 440 nm (coloring strength), etc. These last three categories are historically related to the content of picrocrocin, safranal, and crocins, respectively, the compounds associated with saffron organoleptic characteristics. Picrocrocin [4-(β-D-glucopyranosyl)-2,6,6-trimethyl-1-cyclohexene-1-carboxaldehyde] is considered to be a contributor to saffron's bitter taste, although other compounds such as kaempferols and picrocrocin-related ones with this organoleptic property have been characterized in saffron spice (*10, 11*). In the saffron volatile fraction, safranal (2,6,6-trimethyl-1,3-cyclohexadiene-1-carboxaldehyde) is the major compound, whereas the crocins make up a group of water-soluble carotenoids identified as glycosyl esters of crocetin (8,8′-diapo-Ψ,Ψ′-carotenedioic acid) with glucose, gentiobiose, neapolitanose, or triglucose sugar

* Corresponding author (telephone +34 967 59 93 10; fax +34 967 59 92 38; e-mail Gonzalo.Alonso@uclm.es).
[†] Universidad de Castilla-La Mancha.
[‡] Universidad Politécnica de Valencia.

10.1021/jf703725e CCC: $40.75
Published on Web 04/12/2008

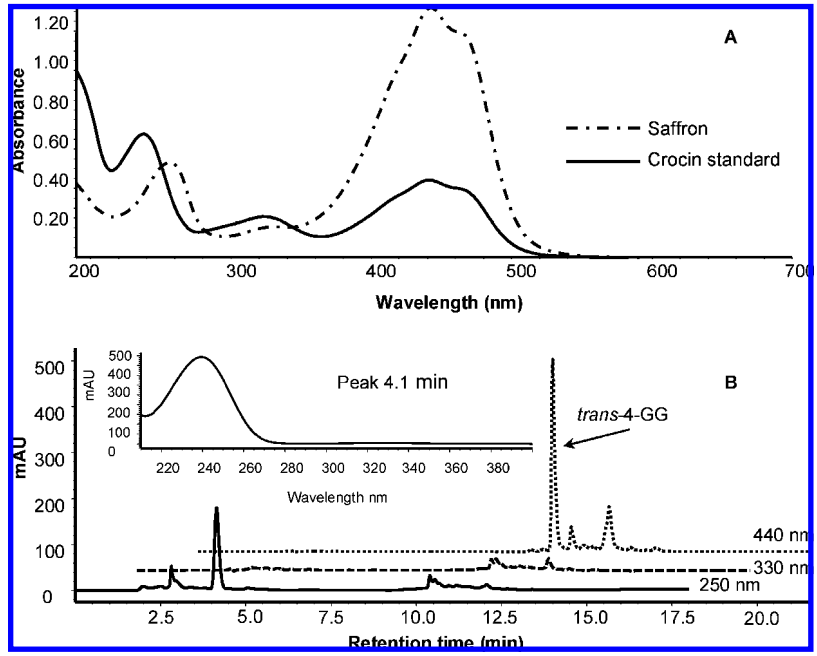

Figure 1. (**A**) UV–vis (— · —) spectrum of saffron spice from the DO "Azafrán de La Mancha", 50 mg/L in water compared to the corresponding one (—) of the commercial crocin standard, 50 mg/L in water (**A**). (**B**) Chromatograms of the commercial crocin standard and absorption spectrum (inset) of peak 4.1 min.

moieties (*1, 12*). The determination of these main characteristics of saffron by UV–vis spectrophotometry, according to ISO standards, leads to the existence of a spectrophotometer in almost all saffron companies. However, in the case of crocetin esters, this procedure by itself does not allow us to distinguish their detailed composition, that is, each *trans*-crocetin ester and each *cis*-crocetin ester. Up to now, other techniques such as thin layer chromatography (TLC) or high-performance liquid chromatography (HPLC) have been used to study these carotenoids as well as picrocrocin, with this last technique being considered to be the most effective one (*1, 2, 13, 14*). However, all of these methods are rather time-consuming and require equipment that is seldom found in small- and medium-size companies that process and package saffron spice. These methods are therefore limited to analytical or research laboratories and can hardly be used in a company to monitor raw materials, processes, or final products.

Quality differences found in saffron are due mainly to its drying process, although edaphoclimatic crop conditions, harvesting, stigma separation, handling, storing, and packaging also have an influence (*15–17*). Apart from the loss of humidity necessary to preserve the spice, drying brings about the physical, biochemical, and chemical changes needed to achieve the desired attributes and can therefore be considered the principal step in its manufacturing process. In this sense, Carmona et al. (*16, 18*) claim that temperature and rate during the drying process are factors that determine saffron's final characteristics. Pardo et al. (*19*) have studied the influence of the dehydration process on the sensory properties of saffron as well.

Despite their low selectivity, spectrophotometric measures have already been developed in other spices such as paprika and its oleoresins in order to establish simple quality criteria for the control of both raw material and finished product (*20*). In addition, multivariate chemometric methods currently play a very important role in the multicomponent analysis of mixtures with UV–vis spectrophotometry under computer-controlled instrumentation. Among the various chemometric approaches applied to multicomponent analysis, principal component regression (PCR) and partial least-squares regression (PLSR) have been successfully adopted in many fields of study, for example, quantitative assays of pharmaceutical formulations (*21*), simultaneous determination of dyes in mixtures (*22*), enology (*23*), and the olive oil (*24*) and alcohol industries (*25*). Numerous studies have also dealt with the possible presence of outliers, their treatment and alternatives for checking the validity of the calibration over time, or other perturbations (*25, 26*). Nevertheless, the studies devoted to the application of multivariate calibration to saffron spice are rather limited (*27, 28*).

The aim of this study was the development of multivariate models able to determine the content of the main crocetin esters and picrocrocin from spectrophotometric data that could then be used in saffron companies in routine quality control. A new method proposed by Galvao et al. (*29*), which is a modification of the Kennard−Stone algorithm (*30*), was used to divide the pool of samples into calibration and validation subsets.

MATERIALS AND METHODS

Samples, Chemicals, and Reagents. A total of $N = 61$ samples of Spanish saffron in filaments ($N_1 = 55$ samples) and powder ($N_2 = 6$ samples) were used in this study. Forty-five samples of the total belonged to the Protected Designation of Origin "Azafrán de La Mancha" and were obtained by means of its Regulatory Council. The remaining samples were directly obtained from the producers with the subsequent guarantee of their origin. The crocin [crocetin di-(β-D-gentiobiosyl) ester] was purchased from Fluka (Buchs, Switzerland). HPLC-grade acetonitrile and cyclohexane were from Merck (Steinhem, Germany). Ultrahigh-purity water was produced using a Millipore Milli-Q System (Bedford, MA); PTFE filters (11 mm, 0.45 μm) were also purchased from Millipore, and C_{18} packing material (125×10^{-8} cm pore size, 55–105 μm particle size) was from Waters (Milford, MA).

Experimental Measures. *Moisture and Volatile Matter Content.* These were determined by successive weighing of 1 g of powdered sample introduced into an oven set at 103 ± 2 °C for 16 h. They were calculated with the following ratio:

$$100 \times \frac{\text{initial mass} - \text{constant mass}}{\text{initial mass}} \quad (1)$$

UV–Vis Determinations. A Perkin-Elmer Lambda 25 (Norwalk, CT) spectrophotometer, accompanied by UV Winlab 2.85.04 software (Perkin-Elmer), was used for spectra recording and treatment at the following conditions: start wavelength, 700 nm; end wavelength, 190 nm; data interval, 1 nm; scan speed, 480 nm/min; 1.0 cm pathway quartz cells from Hellma (Jena, Germany). $E_{1\,\mathrm{cm}}^{1\%}$ 440 nm, $E_{1\,\mathrm{cm}}^{1\%}$ 330 nm, and $E_{1\,\mathrm{cm}}^{1\%}$ 257 nm were calculated according to ISO 3632/TS (*6*).

Identification and Analysis of Crocetin Esters and Picrocrocin by LC-DAD-MS. The identification of the main saffron components was carried out according to the method given in ref *12*, and the same procedure was followed for their analysis, the only difference being the sample injection volume (10 μL). The samples were injected into an Agilent 1100 HPLC chromatograph (Palo Alto, CA) equipped with a 150 mm × 4.6 mm i.d., 5 μm Phenomenex (Le Pecq Cedex, France) Luna C_{18} column thermostaed at 30 °C. The solvents were water (acidified with 0.25% formic acid for identification) (A) and acetonitrile (B), and the gradient system was the following: 20% B, 0–5 min; 20–80% B, 5–15 min; and 80% B, 15–20 min. The flow rate was 0.8 mL/min.

Quantification of Crocetin Esters and Picrocrocin. The only crocetin ester from saffron available on the market is the crocetin di-(β-D-gentiobiosyl) ester, although it is not suitable for quantitative use in HPLC and UV–vis spectrophotometry because it contains considerable amounts of impurities or byproduct, as could be seen by its low $E_{1\,\mathrm{cm}}^{1\%}$ 440 nm, 79, versus 257 for saffron, and its chromatogram (**Figure 1**). Due to this lack of pure standards of each crocetin ester, quantification was based on the equation

$$\%\ \text{of ester } i \text{ on dry basis} = \frac{\mathrm{Mw}_i(E_{1\mathrm{cm}}^{1\%}\ 440\ \mathrm{nm})A_i}{10\epsilon_{t,c}} \quad (2)$$

where Mw_i stands for the molecular weight of the crocetin ester i, $E_{1\,\mathrm{cm}}^{1\%}$ 440 nm is the coloring strength, A_i is the percentage peak area of the crocetin ester i at 440 nm, and $\epsilon_{t,c}$ is the molecular coefficient absorbance value [89000 for *trans*-crocetin esters and 63350 for *cis*-crocetin esters (*31*)]. Data reported represent the average of three sample replicates.

Picrocrocin was isolated by column chromatography on C_{18} packing material. Removal of nonpolar compounds was achieved with 30 mL of cyclohexane added to 5 g of powdered saffron for 24 h at room temperature in the dark with sporadic agitation. The organic solvent was discarded, and the residue was dried under vacuum before the addition of 60 mL of water bubbled with nitrogen. The resulting solution was magnetically stirred for 1 h at room temperature in the dark. Then, the extract was centrifuged at 4000 rpm for 10 min, and the supernatant was collected and transferred to a plastic LC column (8 cm high × 2.7 cm i.d.) filled with the C_{18} packing material. Picrocrocin was eluted with 90 mL of 10% acetonitrile/water (v/v) after elution of flavonoids with 20 mL of 2% acetonitrile/water (v/v) and, finally, the solvent was eliminated to dryness by evaporation under vacuum. A calibration curve as a function of its peak area was constructed. The chromatographic purity of the picrocrocin obtained was calculated as the percentage of the total peak area at 250 nm.

Multivariate Calibration. The data were analyzed by means of multivariate techniques, applying the Unscrambler software, version 9.2 (CAMO Process AS, Oslo, Norway). One PLSR model for each saffron compound as well as for the sum of crocetin esters was built; the outliers were identified and eliminated for each model. The data matrix X was formed by the UV–vis spectra of the saffron aqueous extract, and the vector Y contained the reference values for the corresponding crocetin glycoside or picrocrocin content as the dependent variable. The variables were centered. A new method proposed by Galvao et al. (*29*), which is a modification of the Kennard–Stone algorithm (*30*), was used to divide the pool of samples into calibration and validation subsets for multivariate modeling. The Kennard–Stone algorithm follows a stepwise procedure in which new selections are taken in regions of the space far from the samples already selected. For this purpose, the algorithm employs the Euclidean distances $d_x(p,q)$ between the x-vectors of each pair (p,q) of samples calculated as

$$d_x(p,q) = \sqrt{\sum_{j=1}^{J} [x_p(j) - x_q(j)]^2};\ p,q \in [1,N] \quad (3)$$

where $x_p(j)$ and $x_q(j)$ are instrumental responses at the jth wavelength for samples p and q, respectively. J denotes the number of wavelengths in the spectra,and N is the number of samples. The modification proposed by Galvao et al. for sample set partitioning based on joint $x-y$ distances, SPXY (*29*), consists of augmenting the distance defined with a distance in the dependent variables, $d_y(p,q)$. This distance can be calculated for each pair of samples p and q as follows:

$$d_y(p,q) = \sqrt{\sum_{k=1}^{K} [y_p(k) - y_q(k)]^2};\ p,q \in [1,N] \quad (4)$$

In this paper, a normalized $x-y$ distance was calculated as

$$d_{xy}(p,q) = \frac{d_x(p,q)}{\max_{p,q} d_x(p,q)} + \frac{d_y(p,q)}{\max_{p,q} d_y(p,q)};\ p,q \in [1,N] \quad (5)$$

where $d_x(p,q)$ and $d_y(p,q)$ are divided by their maximum values in the data set in order to assign equal importance to the distribution of the samples in the x and y spaces. A stepwise selection procedure similar to the Kennard−Stone algorithm could then be applied with $d_{xy}(p,q)$ instead of $d_x(p,q)$ alone. A Matlab code (version 7.0) was used for implementation of this method. The division of the 61 samples into calibration, validation, and prediction sets was carried out in the following manner. Initially, 11 prediction samples were extracted from the full set in a random manner to simulate the analysis of a batch of real unknown samples. The remaining 50 samples were divided into calibration and validation sets of 40 and 10 samples, respectively, by using the SPXY algorithm. To improve the statistical significance of the comparison, the extraction of the samples was repeated five times.

The root-mean-square error (RMSE) was calculated with the expression

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{r} (\hat{y}_i - y_i)^2}{r}} \quad (6)$$

where $\hat{y}_i$ is the predicted concentration value of the ith calibration sample, y_i is its real concentration, and r is the number of samples in each case. RMSE was calculated in calibration (RMSEC), validation (RMSEV), and prediction (RMSEP). The number of principal components (PCs) in the PLSR models was determined by testing on the validation set. The minimum RMSEV determined the number of PCs. Three forms to identify outliers were used (*32*): data with extreme leverage, unmodeled residuals in spectral data, and unmodeled residuals in the dependent variable.

Finally, to increase the model's precision and to study the values of the regression coefficients, a global model for each variable was constructed with the SPXY method, 48 calibration samples, and 13 validation ones.

RESULTS AND DISCUSSION

Experimental Measures. The majority of the samples, 55 of the total, fulfilled the ISO specifications for category I regarding moisture and volatile matter content, as well as the main characteristics using UV–vis spectrophotometry; only 4 samples belonged to category II and 2 to category III (**Table 1**). Such a distribution was representative of the saffron available in the Spanish market in a specific harvest, when saffron attains the highest category in most cases. **Table 1** also shows the composition in crocetin esters and picrocrocin expressed as a percentage of saffron on dry mass. The clear fragmentation patterns, retention times, UV–vis spectra,

Table 1. Classification of the Saffron Samples into ISO Categories and Saffron Sample Composition in Crocetin Esters and Picrocrocin Expressed as Grams of Compound per 100 g of Saffron on a Dry Basis

sample	ISO/TS 3632 category	*trans*-4-GG	*trans*-3-Gg	*trans*-2-G	*cis*-4-GG	*cis*-3-Gg	sum of crocetin esters	picrocrocin
M1	I	17.26	8.00	0.28	1.78	0.50	27.82	24.96
M2	I	17.41	7.35	1.04	0.72	0.31	26.83	26.57
M3	I	12.38	5.43	0.36	2.47	0.92	21.57	15.79
M4	I	12.03	5.60	0.57	1.53	0.64	20.38	12.66
M5	I	13.71	6.30	0.61	1.71	0.71	23.04	15.66
M6	I	13.85	5.70	0.56	1.37	0.62	22.10	14.81
M7	I	14.14	5.58	0.44	1.71	0.80	22.67	14.29
M10	III	8.37	3.14	0.23	1.06	0.35	13.15	6.91
M11	II	10.62	4.71	0.46	1.31	0.52	17.62	10.18
M12	III	7.45	2.49	0.19	1.03	0.41	11.57	5.08
M13	II	9.69	4.33	0.42	1.09	0.48	16.01	9.16
M17	I	15.55	7.37	0.58	0.87	0.00	24.37	20.97
M18	I	17.69	7.08	0.32	0.44	0.00	25.52	20.66
M19	I	15.56	6.92	0.35	0.42	0.31	23.56	16.08
M20	I	15.93	7.59	0.42	0.00	0.00	23.94	19.96
M21	I	16.34	7.86	0.64	0.00	0.00	24.84	19.44
M22	I	13.60	6.28	0.39	0.00	0.00	20.27	16.35
M23	II	11.83	4.98	0.29	0.25	0.00	17.36	14.96
M24	I	14.36	7.98	0.68	1.56	0.67	25.24	12.77
M25	I	14.63	7.31	0.58	1.56	0.72	24.80	11.98
M26	I	14.09	7.61	0.69	1.31	0.70	24.40	10.74
M27	I	13.09	7.30	0.68	1.20	0.67	22.94	11.31
M28	I	14.09	7.04	0.60	1.27	0.57	23.57	12.30
M29	I	14.53	7.50	0.64	1.80	0.86	25.32	12.42
M30	I	13.39	6.98	0.61	1.76	0.90	23.63	11.27
M31	I	14.36	7.43	0.65	1.39	0.66	24.49	12.21
M32	I	12.73	6.80	0.65	1.24	0.61	22.03	9.91
M33	I	13.48	6.52	0.50	1.51	0.71	22.73	12.28
M34	I	15.75	7.87	0.72	0.97	0.48	25.79	14.05
M35	I	13.75	6.89	0.61	0.90	0.43	22.60	12.44
M36	I	14.95	8.09	0.68	1.56	0.77	26.05	14.10
M37	I	13.42	7.21	0.59	1.34	0.65	23.21	12.55
M38	I	14.97	6.63	0.39	1.47	0.53	23.99	21.82
M39	I	15.53	7.34	0.35	1.86	0.70	25.77	22.28
M40	I	15.12	7.76	0.56	1.22	0.30	24.97	16.55
M41	I	12.97	5.62	0.19	3.15	1.14	23.07	9.84
M42	I	15.35	8.53	0.66	0.42	0.22	25.18	24.22
M43	I	14.20	6.37	0.25	0.92	0.44	22.18	16.05
M44	I	16.54	7.80	0.50	1.73	0.63	27.20	26.00
M45	I	14.83	7.69	0.42	2.01	0.78	25.73	16.51
M46	I	15.24	9.59	0.80	0.94	0.60	27.16	21.48
M47	I	17.38	9.61	0.74	0.57	0.29	28.59	17.25
M48	I	13.02	6.56	0.27	3.13	1.18	24.17	23.50
M49	I	14.06	9.35	0.67	1.00	0.60	25.68	25.76
M50	I	14.00	6.47	0.31	3.26	1.13	25.16	11.02
M51	I	14.40	7.40	0.29	0.85	0.60	23.52	10.55
M52	I	14.95	8.98	0.63	1.23	0.62	26.41	10.61
M53	I	15.27	7.35	0.37	2.04	0.76	25.79	12.52
M54	I	14.03	7.63	0.45	1.73	0.69	24.52	10.44
M55	I	13.71	8.17	0.48	1.54	0.82	24.71	15.70
M56	I	14.88	6.94	0.28	1.74	0.74	24.59	13.82
M57	I	15.26	8.00	0.48	1.32	0.48	25.53	20.36
M58	I	17.63	8.35	0.73	0.00	0.00	26.72	15.26
M59	I	19.03	7.93	0.43	0.00	0.00	27.39	17.95
M60	I	14.00	4.47	0.11	2.94	0.91	22.42	14.92
M61	I	17.81	7.66	0.38	0.47	0.29	26.61	13.64
M62	I	18.75	7.87	0.24	0.21	0.33	27.41	17.06
M63	I	15.85	7.23	0.31	0.54	0.38	24.30	13.14
M64	II	11.98	4.08	0.12	1.18	0.32	17.69	10.16
M65	I	16.65	7.55	0.40	0.56	0.52	25.68	20.50
M66	I	17.54	7.15	0.33	0.00	0.00	25.01	26.10

and their comparison with previous papers (*12, 28, 33*) allowed the identification of five crocetin esters and picrocrocin (**Figure 2**). To abbreviate the names of crocetin esters, the nomenclature used by Carmona et al. (*12*) has been adopted: first, the reference to the isomeric cis and trans forms has been written with a hyphen separating the total number of glucose moieties at both extremes of the base molecule. Then, the glucose moiety distribution has been indicated as (t) triglucoside, (n) neapolitanoside, (G) gentiobioside, or (g) glucoside. The name of the base structure, crocetin esters, was removed, because it is the same in all compounds.

The three major crocetin esters were, in decreasing order of mean values, *trans*-4-GG > *trans*-3-Gg > *cis*-4-GG. It is noteworthy that the first one alone was found to represent >60% of the total crocetin ester content of the aqueous saffron extract, whereas these three accounted for >95% of the total esters

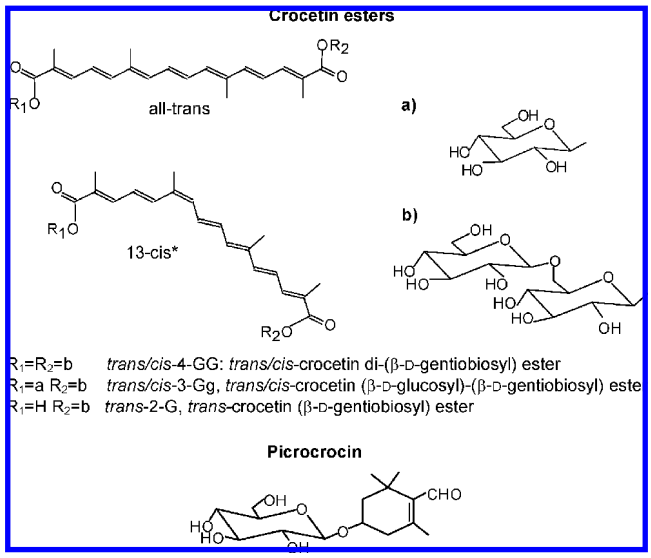

Figure 2. Structures of saffron compounds under discussion. (*) In the case of crocetin esters with cis configuration, the position of the substituents R_1 and R_2 could not be exactly determined in relation to the C_{13-14} bond.

recorded. Despite calculating $E^{1\%}_{1\,cm}$ at 330 nm for the classification of the samples into the ISO categories, safranal was not determined by HPLC due to its low water solubility.

The chromatographic purity of the picrocrocin obtained was 96%. The calibration curve of the picrocrocin concentration, *c* (mg/L), as a function of its peak area, *a*, exhibited good linear regression in the range of 2–315 mg/L: $c = 0.0290\ a + 0.5194$, *R* value = 0.999 for a total of six data points. The heterogeneity of saffron samples (of different coloring strengths and submitted to different dehydration processes), together with the lack of standards for each crocetin ester, has led to a wide range of results concerning the composition of saffron in the literature, making comparisons difficult. On comparing our composition results (**Table 1**) to those obtained by Alonso et al. (*2*) for Spanish saffron from the La Mancha region, we obtained higher contents of *trans*-4-GG, *trans*-3-Gg, and picrocrocin (approximately 3 times more) but very similar ones of *cis*-4-GG and *cis*-3-Gg. Contents of picrocrocin ranging from 0.79 to 13.9% have been previously reported (*2, 34*), even though we found samples up to 26.6% of picrocrocin.

Before multivariate calibration, linear correlations between the content expressed as percentage on a dry basis of the main *trans*- and *cis*-crocetin esters (*trans*-4-GG, *trans*-3-Gg, *trans*-2-G, *cis*-4-GG, and *cis*-3-Gg), the sum of crocetin esters, coloring strength ($E^{1\%}_{1\,cm}$ 440 nm), $E^{1\%}_{1\,cm}$ 257 nm, and $E^{1\%}_{1\,cm}$ 330 nm were studied. The highest Pearson correlation coefficients, *R*, were found between the coloring strength and the following determinations: sum of crocetin esters ($R = 0.994$), % of *trans*-3-Gg ($R = 0.929$), and $E^{1\%}_{1\,cm}$ 257 nm ($R = 0.925$).

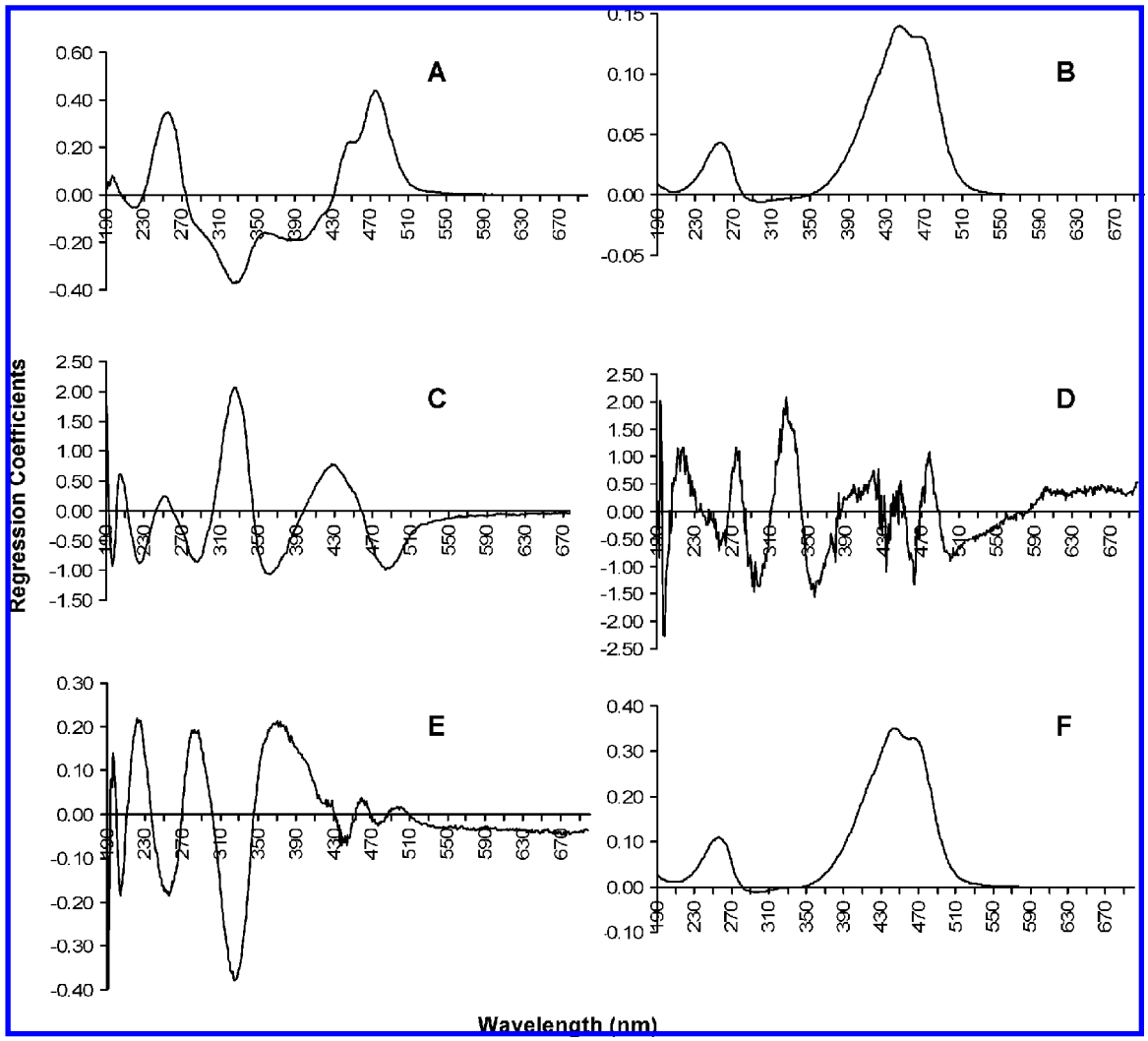

Figure 3. Regression coefficients of the PLSR models for *trans*-4-GG (**A**), *trans*-3-Gg (**B**), *cis*-4-GG (**C**), *cis*-3-Gg (**D**), *trans*-2-G (**E**), and the sum of crocetin esters (**F**).

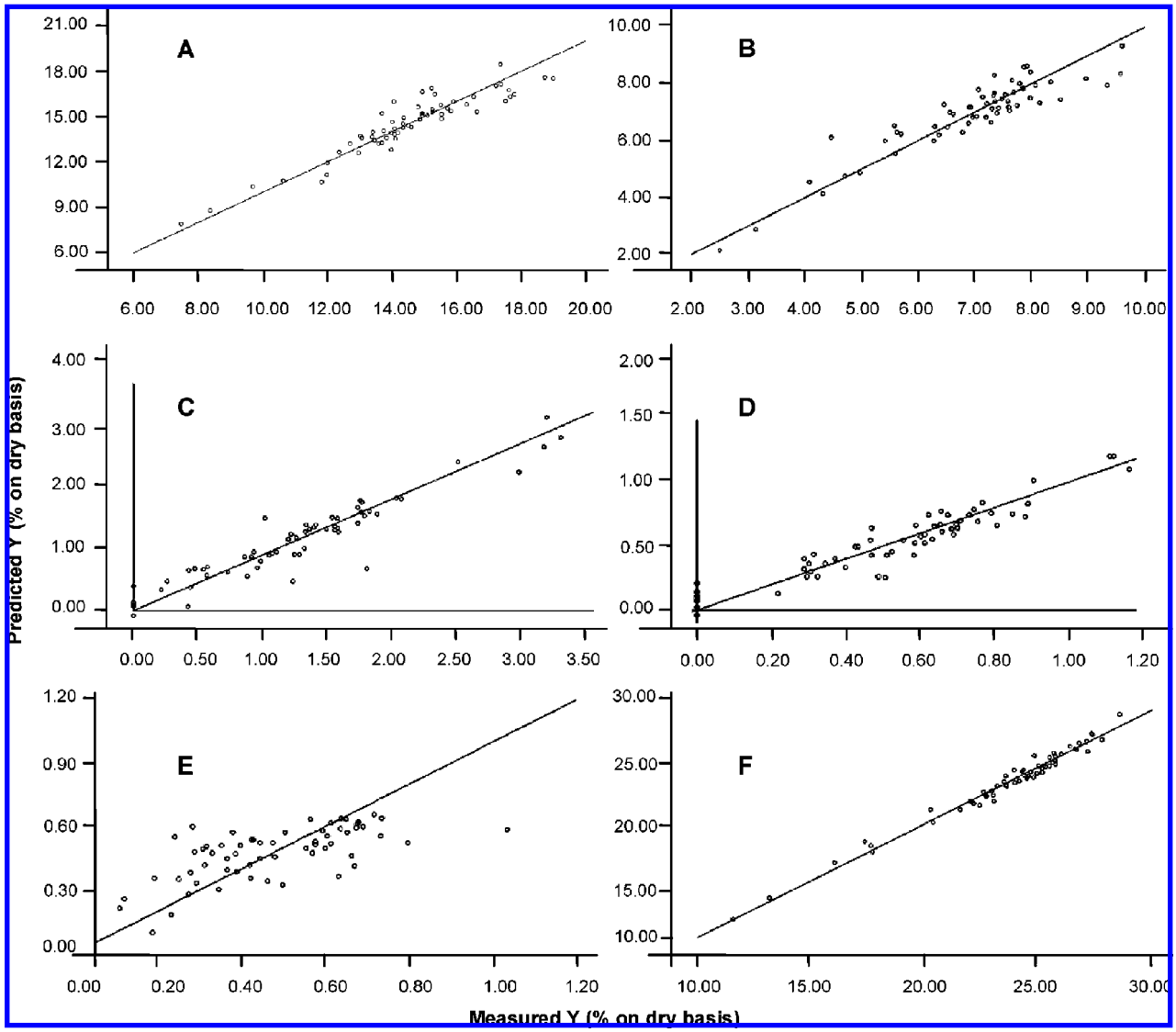

Figure 4. PLSR model predicted values of *trans*-4-GG (**A**), *trans*-3-Gg (**B**), *cis*-4-GG (**C**), *cis*-3-Gg (**D**), *trans*-2-G (**E**), and the sum of crocetin esters (**F**) versus measured ones.

Multivariate Calibration. *PLSR.* In multivariate calibration problems involving complex matrices, it can be difficult to reproduce the composition variability of real samples by means of optimized experimental design. In such cases, a representative calibration set must be extracted from a pool of real samples. Moreover, validation samples should also be selected to assess the quality of the model. Random sampling is a popular technique because of its simplicity and because a group of data randomly extracted from a larger set follows the statistical distribution of the entire set. However, random sampling does not guarantee the representativity of the set, nor does it prevent extrapolation problems. An alternative to random sampling is the Kennard−Stone algorithm (*30, 35*), which covers the multidimensional space in a uniform manner by maximizing the Euclidean distances between the instrumental response vectors of the selected samples. Therefore, this algorithm, with modifications, was used to split samples into calibration and validation sets.

First, an attempt was made to build a PLSR2, that is, a joint model for all of the compounds studied. Because the results were unsatisfactory, it was decided to develop a different PLSR1 model for each one. In the models built with 11 prediction samples and having calibration and validation data sets composed of 40 and 10 samples, respectively, the RMSEP values found, which represented the standard error of the prediction for unknown samples, were low enough for the models to be applied in practice. In all models, except for picrocrocin and *cis*-3Gg, the values of RMSEP were lower than or similar to RMSEC values. This proves the robustness and prediction capability of these models.

In the definitive PLSR model for *trans*-4-GG, the optimum model dimension determined by the minimum RMSEV was two. These two PCs accounted for 97.8% of the variability in the data set and 87.9% of the *trans*-4-GG composition variability (% explained *Y*). In **Figure 3A** the regions of the UV–vis spectrum with higher regression coefficients for this model, and therefore more informative, are shown. It was found that the wavelengths with the highest regression coefficients were at 475, 328, and 255 nm, not coinciding with the maxima of the *trans*-4-GG spectrum. It is quite interesting to point out that, although the spectrum of *trans*-4-GG does not have a maximum at 330 nm, the model gave importance to this region. Almost all of the regression coefficients except those for the region approximating 330 nm were positive. This means that the lower the absorbance is around 330 nm in the spectrum and the higher it is in the rest of the spectrum, the higher the content of *trans*-4-GG. **Figure 4A** shows the goodness of fit model for *trans*-4-GG, presented by plotting its predicted values versus its

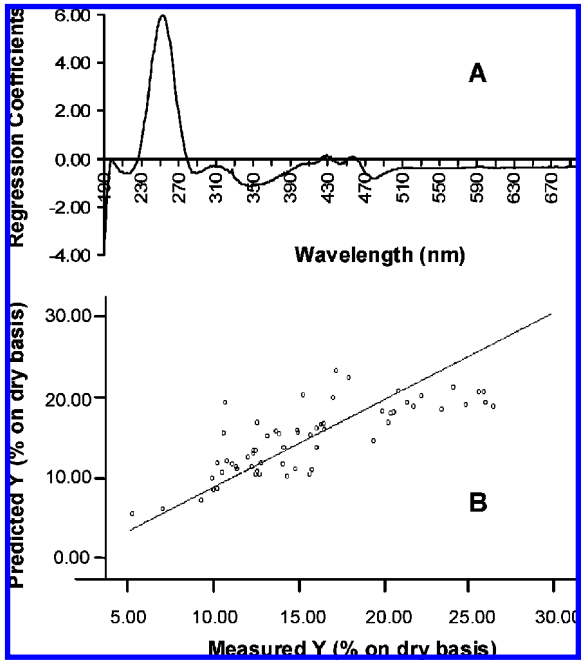

Figure 5. Regression coefficients (**A**) and predicted values of the PLSR model for picrocrocin versus measured ones (**B**).

measured ones. Good results were obtained for the RMSE, which was 0.824 in the calibration (RMSEC), and 0.470 in the validation (RMSEV).

The PLSR model for *trans*-3-Gg presented just one latent variable that accounted for 96.4% of the variability in the data set and 86.2% of the *trans*-3-Gg composition variability. In **Figure 3B** the regions of the UV–vis spectrum with higher regression coefficients are shown. The wavelengths with the highest regression coefficients were at 444, 460, and 255 nm, being very close to the maxima of the *trans*-3-Gg UV–vis spectrum. However, unlike the *trans*-4-GG model, the region around 330 nm was almost not taken into consideration. **Figure 4B** shows the predicted values for *trans*-3-Gg versus the measured ones. The RMSEC was 0.568, whereas RMSEV was 0.452.

The optimum number of PCs for the *cis*-4-GG PLSR model was five, with these accounting for 99.8% of the variability in the data set but explaining only 74.6% of the reference set. This result was understandable as the content in *cis*-4-GG was <10 times that of *trans*-4-GG and <5 times that of *trans*-3-Gg. Its higher variation coefficient could have been responsible for this behavior. In **Figure 3C**, the UV–vis regions with higher regression coefficients are shown. It was found that the wavelengths with the highest regression coefficients were at 326, 363, 485, and 197 nm. Only the maximum at 326 nm coincided with one maximum of the *cis*-4-GG UV–vis spectrum. Differing from the *trans*-4-GG and *trans*-3-Gg models, the correlation coefficients for the region around 330 nm were positive, being in accordance with the fact that *cis*-4-GG has a maximum at 327 nm. The RMSEC was 0.423, whereas the RMSEV was 0.254. **Figure 4C** shows the predicted values for *cis*-4-GG versus the measured ones.

In the PLSR model built for the content of *cis*-3-Gg, the number of PCs was nine, accounting for 99.9% of the variability in the data set, whereas the result for explained *Y* was 90.0%. **Figure 3D** shows that the highest positive regression coefficients were at 327, 193, and 478 nm, whereas the highest negative ones were at 198, 293, and 462 nm. As in the previous model explained, but different from the *trans*-4-GG and *trans*-3-Gg models, the correlation coefficients for the region around 330 nm were positive. **Figure 4D** shows the predicted values for *cis*-3-Gg versus the measured ones. The RMSEC was 0.100, whereas RMSEV was 0.083.

Despite being in the same magnitude order as *cis*-3-Gg, it was not possible to obtain such a good regression model for *trans*-2-G (**Figures 3E** and **4E**). With five PCs the model could explain 99.8% of *X* variability but only 42.1% of *Y* variability. When the model was generated from the sum of crocetin ester data, only one PC was found to explain 96.4% of the variability in the UV–vis spectrum variability and 98.0% of the crocetin ester composition variability. The wavelengths with the highest regression coefficients were exactly the same as those for *trans*-3-Gg (**Figure 3F**), being close to the maxima of the *trans*-crocetin ester spectrum. The region between 280 and 360 nm was not significant in this model. **Figure 4F** shows the predicted values for the sum of crocetin esters versus the measured ones. The RMSEC was 0.469, whereas the RMSEV was 0.339.

The optimum number of PCs in the model built for picrocrocin was four. The region of the saffron spectrum between 240 and 270 nm showed a maximum regression coefficient at 253 nm, practically coinciding with the maximum of the picrocrocin spectrum, whereas the 330–380 nm region showed the lowest weights and negative regression coefficients (**Figure 5A**). The percentage of the variability in the data set for this parameter was 99.8% but only 61.1% of the *Y* variability. **Figure 5B** shows the predicted values of the PLSR model for picrocrocin versus measured ones. Apart from the R^2 value of the *trans*-2-G, the R^2 value of picrocrocin model was the lowest among all of models herein described. The RMSEC was 3.405, whereas RMSEV was 1.892. This result gives evidence that the maximum at 257 nm of the saffron spectrum is due not only to picrocrocin but also to the possible interference of flavonoids and crocetin esters. Better results were reported for NIR by Zalacain et al. (*28*) in reference to the picrocrocin correlation model.

These models are really simple tools that, included or associated with spectrophotometer software, will afford small enterprises the opportunity to achieve quick monitoring of the quality of its products, with only appropriate and periodically checked calibration.

In conclusion, the PLS method based on spectrophotometric data has proved to be a valid tool for determining the main components of saffron spice. Seven PLS1 models have been obtained, and six of them were successfully applied to the determination of *trans*-4-Gg, *trans*-3-Gg, *cis*-4-GG, *cis*-3-Gg, sum of crocetin glycosides, and picrocrocin from saffron. The best predictions were obtained with the sum of the crocetin ester model, followed by the model corresponding to the *cis*-3-Gg, *trans*-4-GG, and *trans*-3-Gg, whereas the worst predictions were found with the picrocrocin and *trans*-2-G models. Nonetheless, with these models we managed to get better correlations with the detailed composition of saffron than by using the UV–vis parameters established by the ISO. These models may considerably enhance quality control in saffron enterprises without the liability of further investments in additional instruments, allowing a more efficient use of their spectrophotometers on a large number of samples.

ABBREVIATIONS USED

Abbreviations in nomenclature were adopted from Carmona et al. (*21*): *trans*-4-GG, *trans*-crocetin di-(β-D-gentiobiosyl) ester; *trans*-3-Gg, *trans*-crocetin (β-D-glucosyl)-(β-D-gentiobiosyl) ester; *trans*-2-G, *trans*-crocetin (β-D-gentiobiosyl) ester; *cis*-

4-GG, *cis*-crocetin di-(β-D-gentiobiosyl) ester; *cis*-3-Gg, *cis*-crocetin (β-D-glucosyl)-(β-D-gentiobiosyl) ester.

ACKNOWLEDGMENT

We thank Kathy Walsh for proof-reading the English manuscript.

Supporting Information Available: Pearson correlation coefficients (R) between the different parameters studied; RMSE results obtained for PLSR prediction models of main saffron components with 11, 40, and 10 prediction, calibration, and validation samples, respectively; and principal characteristics of PLSR prediction models of the main saffron components with 48 and 13 calibration and validation samples, respectively. This material is available free of charge via the Internet at http://pubs.acs.org.

LITERATURE CITED

(1) Tarantilis, P. A.; Tsoupras, G.; Polissiou, M. G. Determination of saffron (*Crocus sativus* L.) components in crude plant extract using high-performance liquid chromatography−UV-visible photodiode-array detection−mass spectrometry. *J. Chromatogr., A* **1995**, *699*, 107–118.

(2) Alonso, G. L.; Salinas, M. R.; Garijo, J.; Sánchez, M. A. Composition of crocins and picrocrocin from Spanish saffron (*Crocus sativus* L.). *J. Food Qual.* **2001**, *24*, 219–233.

(3) Alonso, G. L.; Salinas, M. R.; Sánchez, M. A.; Garijo, J. Safranal content in Spanish saffron. *Food Sci. Technol. Int.* **2001**, *7*, 225–229.

(4) Sánchez, A. M.; Carmona, M.; Ordoudi, S. A.; Tsimidou, M. Z.; Alonso, G. L. Kinetics of individual crocetin ester degradation in aqueous extracts of saffron (*Crocus sativus* L.) upon thermal treatment in the dark. *J. Agric. Food Chem.* **2008**, *56*, 1627–1637.

(5) Zalacain, A.; Ordoudi, S. A.; Blázquez, I.; Díaz-Plaza, E. M.; Carmona, M.; Tsimidou, M. Z.; Alonso, G. L. Screening method for the detection of artificial colours in saffron using derivative UV-vis spectrometry after precipitation of crocetin. *Food Addit. Contam.* **2005**, *22*, 607–615.

(6) ISO 3632/TS-1, 2 Saffron (*Crocus sativus* L.) Part 1: Specifications, Part 2: Test Methods; ISO: Geneva, Switzerland, 2003.

(7) DOCE (2001). European Commision 464/2001. 8 March 2001, p 29. (L 66 de 8.3.2001). http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2001:066:0029:0030:ES:PDF (accessed 12.2.2008).

(8) DOCE (1999). European Commision 378/99. 20 February 1999, p. 13. (L 46 de 20.9.1999). http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:1999:046:0013:0014:ES:PDF (accessed 12.2.2008).

(9) OJUE L 33/6, Commision Regulation (EC) 205/2005 of 4 February 2005, 5.2.2005. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2005:033:0006:0007:EN:PDF (accessed 12.2.2008).

(10) Straubinger, M.; Bau, B.; Eckestein, S.; Fink, M.; Winterhalter, P. Identification of novel glycosidic aroma precursors in saffron (*Crocus sativus* L.). *J. Agric. Food Chem.* **1998**, *46*, 3238–3243.

(11) Carmona, M.; Sánchez, A. M.; Ferreres, F.; Zalacain, A.; Tomás-Barberán, F.; Alonso, G. L. Identification of the flavonoid fraction in saffron spice by LC/DAD/MS/MS: comparative study of samples from different geographical origins. *Food Chem.* **2007**, *100*, 445–450.

(12) Carmona, M.; Zalacain, A.; Sánchez, A. M.; Novella, J. L.; Alonso, G. L. Crocetin esters, picrocrocin and its related compounds present in *Crocus sativus* stigmas and *Gardenia jasminoides* fruits. Tentative identification of seven new compounds by LC-ESI-MS. *J. Agric. Food Chem.* **2006**, *54*, 973–979.

(13) Lozano, P.; Castellar, M. R.; Simancas, M. J.; Iborra, J. L. Quantitative high-performance liquid chromatographic method to analyse commercial saffron (*Crocus sativus* L.) products. *J. Chromatogr., A* **1999**, *830*, 477–483.

(14) Sujata, V.; Ravishankar, G. A.; Venkataraman, L. V. Methods for the analysis of the saffron metabolites crocin, crocetins, picrocrocin and safranal for the determination of the quality of the spice using thin-layer chromatography, high-performance liquid chromatography and gas chromatography. *J. Chromatogr.* **1992**, *624*, 497–502.

(15) Ordoudi, S.; Tsimidou, M. Saffron quality: effect of agricultural practices, processing and storage. In *Harvest and Quality Evaluation of Food Crop*; Dris, R., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2004; pp 215–218.

(16) Carmona, M.; Zalacain, A.; Pardo, J. E.; López, E.; Alvarruiz, A.; Alonso, G. L. Influence of different drying and aging conditions on saffron constituents. *J. Agric. Food Chem.* **2005**, *53*, 3974–3979.

(17) Carmona, M.; Martínez, J.; Zalacain, A.; Rodríguez-Méndez, M. L.; de Saja, J. A.; Alonso, G. L. Analysis of saffron volatile fraction by TD-GC-MS and e-nose. *Eur. Food Res. Technol.* **2006**, *223*, 96–101.

(18) Carmona, M.; Zalacain, A.; Alonso, G. L. The dehydration process. In *The Chemical Composition of Saffron: Color, Taste and Aroma*, 1st ed.; Editorial Bomarzo: Albacete, Spain, 2006; pp 60–62.

(19) Pardo, J. E.; Zalacain, A.; Carmona, M.; López, E.; Alvaruiz, A.; Alonso, G. L. Influence of the type of dehydration process on the sensory properties of saffron spice. *Ital. J. Food Sci.* **2002**, *4*, 413–422.

(20) Hornero-Méndez, D.; Mínguez Mosquera, M. I. Rapid spectrophotometric determination of red and yellow isochromic carotenoid fractions in paprika and red pepper oleoresins. *J. Agric. Food Chem.* **2001**, *49*, 3584–3588.

(21) Ragno, G.; Ioele, G.; Risoli, A. Multivariate calibration techniques applied to the spectrophotometric analysis of one-to-four component systems. *Anal. Chim. Acta* **2004**, *512*, 173–180.

(22) Berzas, J. J.; Rodríguez Flores, J.; Villaseñor Llerena, M.; Rodríguez Fariñas, J. N. Spectrophotometric resolution of ternary mixtures of tartrazine, patent blue V and indigo carmine in commercial products. *Anal. Chim. Acta* **1999**, *391*, 353–364.

(23) Carot, J. M. Aplicación de Métodos Multivariantes al Estudio de Mezclas de Vinos Tintos Monovarietales, Thesis, Universidad Politécnica de Valencia, Valencia, Spain, 2003.

(24) Alves, M. R.; Cunha, S. C.; Amaral, J. S.; Pereira, J. A.; Oliveira, M. B. Classification of PDO olive oils on the basis of their sterol composition by multivariate analysis. *Anal. Chim. Acta* **2005**, *549*, 166–178.

(25) Valderrama, P.; Braga, J. W.; Poppi, R. J. Variable selection, outlier detection, and figures of merit estimation in a partial least-squares regression multivariate calibration model. A case study for the determination of quality parameters in the alcohol industry by near-infrared spectroscopy. *J. Agric. Food Chem.* **2007**, *55*, 8331–8338.

(26) Fernández Pierna, J. A.; Wahl, F.; Noord, O. E.; Massart, D. L. Methods for outlier detection in prediction. *Chemom. Intell. Lab. Syst.* **2002**, *63*, 27–39.

(27) Zougagh, M.; Ríos, A.; Valcárcel, M. An automated screening method for the fast, simple discrimination between natural and artificial colorants in commercial saffron products. *Anal. Chim. Acta* **2005**, *535*, 133–138.

(28) Zalacain, A.; Ordoudi, S. A.; Díaz-Plaza, E. M.; Carmona, M.; Blázquez, I.; Tsimidou, M. Z.; Alonso, G. L. Near-infrared spectroscopy in saffron quality control: determination of chemical composition and geographical origin. *J. Agric. Food Chem.* **2005**, *53*, 9337–9341.

(29) Galvao, R. K.; Araujo, M. C.; José, G. E.; Pontes, M. J.; Silva, E. C.; Saldanha, T. C. A method for calibration and validation subset partitioning. *Talanta* **2005**, *67*, 736–740.

(30) Kennard, R. W.; Stone, L. A. Computer aided design experiments. *Technometrics* **1969**, *11*, 137–148.

(31) Speranza, G.; Dada, G.; Manitto, P.; Monti, D.; Grammatica, P. 13-*Cis* crocin: a new crocinoid of saffron. *Gazz. Chim. Ital.* **1984**, *114*, 189–192.

(32) Martens, H.; Martens, M. *Multivariate Analysis of Quality. An Introduction*; Wiley: Chichester, U.K., 2001.

(33) Pfister, S.; Meyer, P.; Steck, A.; Pfander, H. Isolation and structure elucidation of carotenoid-glycosyl esters in *Gardenia* fruits (*Gardenia jasminoides* Ellis) and saffron (*Crocus sativus* L.). *J. Agric. Food Chem.* **1996**, *44*, 2612–2615.

(34) Iborra, J. L.; Castellar, M. R.; Canovas, M.; Manjón, A. TLC preparative purification of picrocrocin, HTCC and crocin from saffron. *J. Food Sci.* **1992**, *57*, 714–731.

(35) Kanduc, K. R.; Zupan, J.; Majcen, N. Separation of data on the training and test set for modelling: a case study for modelling of five colour properties of a white pigment. *Chemom. Intell. Lab. Syst.* **2003**, *65*, 221–229.

Received for review December 20, 2007. Revised manuscript received February 14, 2008. Accepted February 15, 2008. We thank the Consejería de Educación y Ciencia of the Junta de Comunidades de Castilla-La Mancha and the European Social Fund for funding this work through Grant Exp. 04/131 and the Ministerio de Educación y Ciencia and FEDER (CE) for Project AGL2007-64092/ALI.

JF703725E