JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY



# Nondestructive Application of Laser-Induced Fluorescence Spectroscopy for Quantitative Analyses of Phenolic Compounds in Strawberry Fruits (*Fragaria x ananassa*)

J. S. WULF,*,† S. RÜHMANN,‡ I. REGO,‡ I. PUHL,‡ D. TREUTTER,‡ AND M. ZUDE†

Leibniz Institute for Agricultural Engineering Potsdam-Bornim, Max-Eyth-Allee 100, 14469 Potsdam, Germany, and Technical University Munich, Alte Akademie 16, 85350 Freising, Germany

Laser-induced fluorescence spectroscopy (LIFS) was nondestructively applied on strawberries (EX = 337 nm, EM = 400–820 nm) to test the feasibility of quantitatively determining native phenolic compounds in strawberries. Eighteen phenolic compounds were identified in fruit skin by UV and MS spectroscopy and quantitatively determined by use of rp-HPLC for separation and diode-array or chemical reaction detection. Partial least-squares calibration models were built for single phenolic compounds by means of nondestructively recorded fluorescence spectra in the blue-green wavelength range using different data preprocessing methods. The direct orthogonal signal correction resulted in $r^2 = 0.99$ and rmsep < 8% for *p*-coumaroyl-glucose, and $r^2 = 0.99$ and rmsep < 24% for cinnamoyl-glucose. In comparison, the correction of the fluorescence spectral data with simultaneously recorded reflectance spectra did not further improve the calibration models. Results show the potential of LIFS for a rapid and nondestructive assessment of contents of *p*-coumaroyl-glucose and cinnamoyl-glucose in strawberry fruits.

**KEYWORDS: blue-green fluorescence; direct orthogonal signal correction; LIFS; partial least-squares regression; reflectance; strawberry**

## INTRODUCTION

Fresh fruits and vegetables deliver water, carbohydrates, proteins, nucleic acids, and lipids for the human nutrition. They possess compounds built up in the secondary metabolism such as phenolic compounds that have been further investigated in recent years in different clinical studies for their constitutional effects in the human diet (*1*). However, the health-promoting effects of single phenolic compounds in the human diet are a controversial topic of discussion (*2, 3*). Therefore, it is not the total phenolic compound content that determines the nutritional value of horticultural products but rather the composition and contents of single native phenolic compounds (*4*).

During the last three decades, spectral optical methods in optical geometries of transmittance, remittance, and diffuse reflectance have been successfully applied in nondestructive fruit and vegetable analyses addressing the pigment absorption (*5–7*), whereas the potential of fluorescence spectroscopy has not been investigated systematically. Research in fluorescence spectral analysis in agricultural and horticultural sciences has been mainly focused on changes in fruit chlorophyll fluorescence kinetic detectable in the red wavelength range (*8*). Additionally, spectroscopic methods and imaging techniques were tested to analyze the fluorescing molecules in plants and horticultural products (*9*). Furthermore, fluorimetric systems were introduced to detect changes in the blue-green fluorescence (*10, 11*). Few working groups have been dealing with quantitative fluorescence analyses of compounds in processed fruits although promising results were found in other food stuffs (*12*).

The detection of fruit compounds by fluorescence spectroscopy is more selective than the photometric determination of absorbance and reflectance attributes, because only a part of the fruit molecules possess the ability to re-emit the absorbed energy as fluorescence radiation after some pico- or nanoseconds (average 10 ns) (*13*) in which the molecule remains in the excited state. Moreover, small amounts of fluorescing compounds can also be detected fluorimetrically because of the higher sensitivity compared to the absorption and reflectance spectroscopy. However, the fluorescence spectrum recorded from complex fruit tissue represents the sum of all single fluorescence spectra of every fluorophor in the probe. As a result, the molecule fluorescence signal in vivo is influenced by both environmental and inter- and intramolecular factors, which also change during development and ripening of fruits. Reabsorption and masking effects of autofluorescent and nonfluorescent molecules as well as scattering effects can interfere with the fluorescence emission of native fluorophors. The use of statistical methods instead of generic models for evaluating horticultural products in a qualitative and quantitative manner therefore

* Corresponding author. E-mail: jwulf@atb-potsdam.de.
† Leibniz Institute for Agricultural Engineering Potsdam-Bornim.
‡ Technical University Munich.

10.1021/jf072495i CCC: $40.75 
Published on Web 04/17/2008

became more important (*14*) for eliminating these influences on the spectral intensities which are not correlated to the fluorescence signal of a specific fruit compound under question. Using multivariate processing methods, particularly partial least-squares (PLS) regression, the spectral data can be evaluated and information on a specific fruit compound can be extracted (*14, 15*). However, dealing with raw data often leads to unsatisfactory calibration models. The use of different data preprocessing methods to reduce the noise, off-sets, and scattering influences in the spectral data before calculating the calibration model or to eliminate the information in the spectral data that is not referring to the content of a specific compound might be valuable for building robust calibrations on the fluorescence spectra as well for dealing with the reabsorption effects of the fruit matrix.

Alternatively, an often applied method in fluorescence spectroscopy to correct the fluorescence spectra is the use of the information contained in simultaneously recorded diffuse reflectance spectra of the fruit tissue. This approach has been successfully undertaken in applications of Gitelson et al. (*16*) and Löhmannsröben and Schober (*17*). For this purpose, (i) the entire reflectance spectrum can be subtracted wavelength by wavelength from the fluorescence spectrum, and (ii) a single reflectance wavelength can be used at which preferably marginal fluorescence appears for the correction of the entire fluorescence spectrum.

In the present work, nondestructive fluorescence measurements using laser-induced fluorescence spectroscopy (LIFS) and chromatographical fruit analyses were applied on strawberries to evaluate the feasibility for the quantification of single phenolic fruit compounds in comparison to quantitative HPLC-DAD and HPLC-CRD analysis. An approach was carried out to correct the apparent fluorescence signal by means of reflectance readings and statistical preprocessing methods.

## MATERIALS AND METHODS

**Plant Material.** Strawberries (*Fragaria x ananassa*) cv. 'Elsanta' cultivated in hydroponic production system were harvested in November 2005 in a commercial greenhouse in the region of Brandenburg, Germany. At each harvest date, three different maturity stages were collected and visually classified as follows: In ripening stage 1, the unripe fruits were characterized by a greenly white color; ripening stage 2 contained ripe berries with the skin appearance turning from white to red; and ripening stage 3 was composed of slightly overripe strawberries of overall red color. Parallel readings of nondestructive measurements and chromatographic analyses were accomplished at the fruit equator. Therefore the fruits were marked on the sepals and the measuring point for spectroscopic readings as well as sampling for chemical analyses was chosen on the fruit equator under the marked sepal. The calibration models were calculated using the spectral data and the concentrations of phenolic compounds estimated by rp-HPLC-analysis ($n = 32$) on the first picking date. An independent test set ($n = 35$) of fruits harvested 1 week later in the same green house was used for the model validation.

In a pre-experiment, chemical analysis has been carried out on a batch of strawberry fruits ($n = 89$) for the identification of phenolic compounds in the fruit tissue.

**Extraction and rp-HPLC Analysis of Phenolic Compounds.** Strawberry fruits were frozen in liquid nitrogen and lyophilized (BETA 1–8 LMC-1, Christ Gefriertrocknungsanlagen, Germany). From every single lyophilized fruit, the skin was scraped off with a razor blade as a layer of about 0.5 mm around the measuring point. The sample preparation corresponds to the depth of light penetration of around 0.5 mm. The sample was ground in a mortar and 2–6 mg per sample were used for extraction. For extraction, per 2 mg of powder, 500 $\mu$L of methanol containing 0.5 mg mL$^{-1}$ flavone (Carl Roth GmbH, Germany) as an internal standard and 1.8 mL of methanol were added in a test tube. Extraction was carried out for 30 min in a cooled water bath using ultrasonication. After centrifugation. the supernatant was evaporated; the residue was redissolved in 100 $\mu$L of methanol and injected for HPLC-analysis.

The HPLC analyses were performed on a Gilson-Abimed Systems (model 231, Gynkotek HPLC, Germany) using two pumps (model 422, Kontron Instruments, Germany) and a diode-array-detector (Bio Tek Kontron 540, Kontron Instruments, Germany). For the postcolumn derivatization a HPLC pump (model 300 C, Gynkotek HPLC, Germany) was used with a VIS-detector (640 nm, Kontron Detektor 432, Kontron Instruments, Germany). The phenolic compounds were separated on a Nucleosil 120–3 C18 column (250 × 4 mm, 3 $\mu$m particle diameter, Machery-Nagel HPLC, Germany) following a stepwise gradient using mixtures of solvent A (formic acid 5% in water) and solvent B (methanol) from 95:5, v/v to 10:90, v/v with a flow rate of 0.5 mL min$^{-1}$ (*26*). The gradient profile (%B in A) used was 0–5 min, isocratic, 5% B; 5–15 min, 5–10% B; 10–15 min, isocratic, 10% B; 15–35 min, 10–15% B; 35–55 min, isocratic, 15% B; 55–70 min, 15–20% B; 70–80 min, isocratic, 20% B; 80–95 min, 20–25% B; 95–125 min, 25–30% B; 125–145 min, 30–40% B; 145–160 min, 40–50% B; 160–175 min, 50–90% B; 175–195 min, isocratic, 90% B.

Identification of compounds was achieved by their chromatographic behavior, by their UV absorbance spectra and by mass spectroscopy in comparison with authentic standards and with data from literature (**Table 1**). The procyanidins B1, B3 and p-coumaroyl-glucose were previously identified (*18, 20*). Cinnamoyl-glucose was kindly provided by W. Schwab (Technical University Munich). Commercially available standards used were caffeic acid, *p*-coumaric acid, cinnamic acid, pelargonidin, cyanidin, isoquercitrin, kaempferol, catechin, and gallic acid, which were obtained from (Roth-Karlsruhe, Germany). Epicatechin-3-gallate was purchased from (Extrasynthèse-Lyon, France).

Hydroxycinnamic acids, dihydroxychalcones, and flavonols were detected at 280 nm and anthocyanins at 540 nm, whereas flavanols were estimated by chemical reaction detection (CRD) at 640 nm after postcolumn derivatization with 4-dimethylaminocinnamic aldehyde (*18, 19*). For quantification, we estimated procyanidins B1 and B3 as procyanidin B2, caffeoyl-glucoside as caffeic acid, *p*-coumaroyl-glucose and kaempferol-coumaroyl-glucoside as *p*-coumaric acid, cinnamoyl-glucose and cinnamic acid derivate as cinnamic acid, pelargonidin-glucoside and pelargonidin-malonyl-glucoside as pelargonidin, cyanidin-glucoside-succinate as cyanidin, quercetin-glucoside as isoquercitrin, kaempferol-glucuronide and kaempferol-acetyl-glucoside as kaempferol, and galloyl-ester as gallic acid.

For statistical analysis, the data were first tested for normal distribution using the Kolmogorov–Smirnov test, which was confirmed with $P < 0.01$. The data were further treated with a one-way ANOVA and the TUKEY test using Minitab 14 statistical software.

**Mass Spectroscopy.** For MS/MS analyses of single compounds, the crude extract was first fractionated by several analytical runs as described above. The fractions containing only a few phenolic compounds were further separated on Phenomenex Synergy (Fusion-RP 18, 80 A, 50 × 2 mm, 4 $\mu$m particle diameter) and a stepwise gradient using mixtures of solvent A (formic acid, 0.1% in water) and solvent B (formic acid, 0.1% in MeOH) from 94:6, v/v to 0:100, v/v with a flow rate of 0.2 mL min$^{-1}$. The gradient profile (%B in A) used was: 0–0.2 min, isocratic, 6% B; 0.2–14.2 min, 6–16.5% B; 14.2–16.2 min, 16.5–17% B; 16.2–18.2 min 17–17.5% B; 18.2–20.2 min, isocratic, 17.5% B; 20.2–22.2 min, 17.5–18.5% B; 22.2–24.2 min, isocratic, 18.5% B; 24.2–27.2 min, 18.5–20% B; 27.2–46.2 min, 20–100% B. The survey of molar mass and fragment occurred on a triple-quadrupol mass spectrometer (API 3000, Perkin-Elmer Sciex, USA) with turbo stray as ion source, polarity negative, scan-type Q1 MS in molar mass range *m/z* 100−1000.

**Laser-Induced Fluorescence Spectroscopy.** The fluorometer (Laser Fluoroscope LF401 Lambda, I.O.M., Germany) used for the detection of fluorescence spectra of strawberries was equipped with a pulsed nitrogen laser emitting short laser pulses of 350 ps at a repetition rate of 16 Hz at 337 nm (**Figure 1**). The depth of light penetration of the excitation light in the fruit tissue is about 0.5 mm. An exciting as well as receiving fiber-optic probe allowed reading spectra directly from the fruit surface. The fiber probe had a length of 3 m with a kernel

**Table 1.** Identification of Phenolic Compounds in Strawberry Fruits using HPLC-DAD and LC-MS/MS ($m/z$ = mass to charge ratio; $[M-H]^-$ = molecule ion minus H (negative mode), $t_R$ = retention time)

| $t_R$ (min) | $\lambda_{max}$ (nm) | avg $m/z$ $[M-H]^-$ | MS/MS fragments | identification |
|---|---|---|---|---|
| | | Benzoic Acids | | |
| 8 | 287 | | | galloyl-ester |
| | | Cinnamic Acids | | |
| 19 | 286 | 341 (179 + hexose) | | caffeoyl-glucose |
| 37 | 314 | 325 | 117, 145, 89, 119 | p-coumaroyl-glucose |
| 66 | 284 | 355 [M + COOH] | 103, 59, 71, 146 | cinnamoyl-glucose |
| 70 | 282 | | | cinnamic acid derivative |
| | | Flavanols | | |
| 23 | 278 | | | procyanidin B3 |
| 24 | 277 | | | procyanidin B1 |
| 25 | 278 | | | catechin |
| 80 | 278 | | | epicatechingallate |
| | | Anthocyanins | | |
| 109 | 272, 329sh, 501 | 432 | 268, 223, 146, 120 | pelargonidin-glucoside |
| 141 | 278, 518 | | | cyanidin-glucoside-succinate |
| 48 | 271, 330sh, 502 | 519 | | pelargonidin-malonylglucoside |
| | | Flavonols | | |
| 131 | 257, 352 | 477 | 151, 301, 179, 107, 109, 121 | quercetin-glucuronide |
| 147 | 264, 346 | | | kaempferol-glucuronide |
| 162 | 264, 346 | 489 | 283, 255, 227, 285 | kaempferol-acetyl-glucoside |
| 169 | 268, 314 | 593 | 284, 285 | kaempferol-coumaroyl-glucoside |
| | | Unknown Compounds | | |
| 5 | 248 | 603 | 279 | simple phenolic |
| 17 | 275 | 399 | | cinnamic acid-derivative |

HPLC-DAD and LC-MS/MS Data of Reference Compounds

| $t_R$ (min) | $\lambda_{max}$ (nm) | avg $m/z$ $[M-H]^-$ | MS/MS fragments | ref compd |
|---|---|---|---|---|
| 27 | 278 | 577 | 425, 289, 407 | procyanidin B3 |
| 30 | 277 | 289 | 245, 109 | catechin |
| 44 | 323 | 179 | 135, 89, 107, 79, 117 | caffeic acid |
| 69 | 309 | 163 | 147, 117, 119, 91, 89 | p-coumaric acid |
| 138 | 278 | 147 | | cinnamic acid |
| 166 | 254, 371 | 301 | 151, 179, 107, 121 | quercetin |
| 172 | 262, 366 | 285 | 143, 108, 93, 117 | kaempferol |

diameter of 600 $\mu$m, providing the measuring spot, and was connected to a Y-coupler with a length of 0.50 m. To avoid measuring diffuse reflectance, the surface of the fiber-optic probe was pasted into a high-grade steel tube shaped with an angle of 8°. The acousto-optic tuneable filter enables the scanning detection in 1 nm intervals in a spectral range of 350−820 nm. The signal was detected by a photomultiplier tube and subsequently processed with a high time resolution (100 ps gate width). The low frequency of the laser pulses and an additionally integrated electronic filter allowed recordings at different incident light conditions. To avoid systematical errors due to wavelength shifts and intensity changes of the experimental setup, we carried out measurements in certain intervals on different standards such as in riboflavin solution (Carl Roth GmbH, Germany), and on 100% reflecting Spectralon (Labsphere, USA).

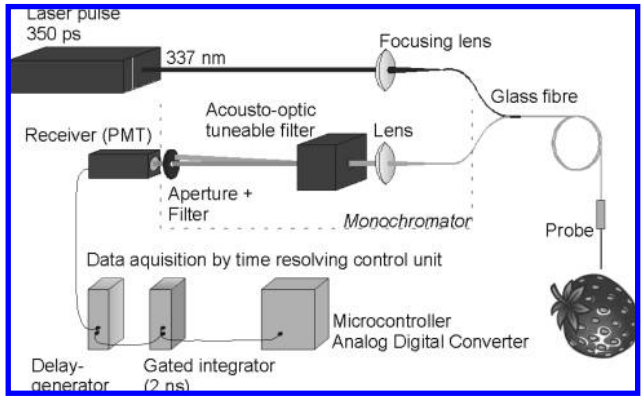


**Figure 1.** Schematic view of the experimental setup with a nitrogen laser emitting at 337 nm serving as light source and a photomultiplier (PMT) as detector.

For every strawberry, one fluorescence spectra has been recorded directly at fruit epidermis. The spectral readings were done in room temperature conditions with a measuring time of about 5 min for every single spectrum.

**Reflectance Spectroscopy.** Fruit diffuse reflectance recordings were carried out with a laboratory device as explained in Zude (*6*). The spectrophotometer (MMS1 NIR-enhanced, Zeiss-Jena, Germany) was equipped with a photodiode-array detector recording photons in the wavelength range from 300 to 1100 nm with resolution of 3.3 nm. A halogen lamp (12V/20W) was used as light source and glass fibers allowed recording the reflected light using an integrated sphere directly from the fruit surface. The system was calibrated with white Spectralon (Labsphere, USA) by measured spectrum = (raw spectrum − dark spectrum)/(white spectrum − dark spectrum).

**Data Processing.** The data sets of the strawberry fluorescence spectra were processed using different preprocessing methods (PLS-Toolbox 3.0, Eigenvector Research, USA) such as smoothing, derivation, column mean centering, and scaling. Scaling here refers to columnwise mean centering and division by the standard deviation. The dependent parameter $y$ (phenolic compound concentration) has always been scaled having previously tested different preprocessing methods. Smoothing and derivation of the fluorescence spectra have been calculated according to the Savitzky−Golay algorithm (*21*) over a window of seven values and a second-order polynome. As additional preprocessing method the direct orthogonal signal correction (DOSC) method was applied using one and two DOSC factors and a tolerance value for the orthogonality of 0.001 (*22*).

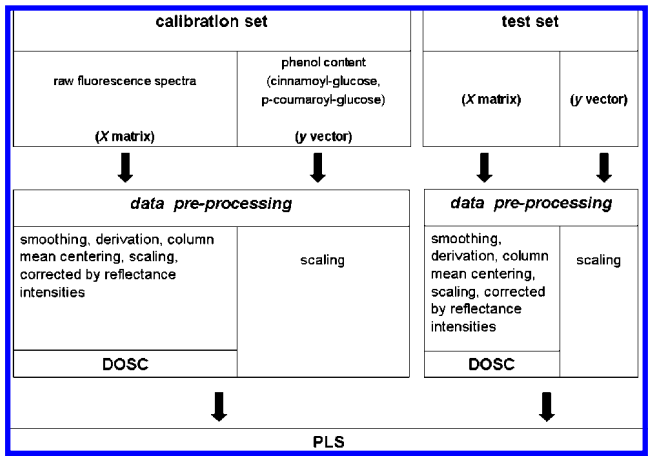


**Figure 2.** Schematic view of the different data-preprocessing steps and the direct orthogonal signal correction (DOSC) algorithm to calculate calibration models with the help of partial least-squares (PLS) regression based on the strawberry fluorescence spectral data (*X* matrix) and their phenol content (*y* vector).

The correction of fluorescence intensities using the diffuse reflectance data was carried out by (i) dividing the entire fluorescence spectra (400−820 nm) at each wavelength-dependent intensity by the corresponding reflectance intensities, or by (ii) using a single reflectance wavelength (800 nm), where low fluorescence and high scattering coefficients can be assumed. The reflectance spectra were treated with cubic spline interpolation in MATLAB 7.0 to adjust the wavelength steps of diffuse reflectance readings to the fluorescence data measured in 2 nm intervals.

The calibration models were developed using partial least-squares (PLS) regressions in MATLAB 7.0 (The Math Works, USA) applied to preprocessed data (**Figure 2**). Results of the PLS regression were determined using the SIMPLS algorithm and a leave-one-out cross validation. The percentage value of the root mean squares error of cross validation (rmsecv) has been calculated by means of dividing by the mean of *y* and multiplying by 100. The root mean squares error of prediction (rmsep) was evaluated on an independent test set. The number of latent variables was adjusted while minimizing the validation error.

## RESULTS AND DISCUSSION

**Identification and Contents of Phenolic Compounds in Strawberry Fruits.** The identification of phenolic compounds in strawberry fruits (*Fragaria x ananassa*) cv. 'Elsanta' (**Table 1**) was carried out by the UV-absorbance spectra using a diode array detector, by cochromatography with authentic samples, if available, and by mass spectrometry (LC-MS/MS). The main anthocyanins were pelargonidin-glucoside and pelargonidin-malonyl-glucoside as formerly described by Määttä-Riihinen et al. (*23*). Minor amounts of cyanidin 3-glucoside-succinate was tentatively identified according to the data described by Lunkenbein and co-workers (*24*).

In the group of benzoic acids, a galloyl-ester was measured. Based on MS/MS measurement and in comparison with data from Seeram et al. (*25*) and Zheng et al. (*26*), ellagic acid and presumably a methyl-ellagic acid-pentose were identified, which occurred in marginal concentrations in the samples. On the basis of comparison with the literature (*23, 24*) three cinnamic acids were identified: cinnamoyl-glucose, *p*-coumaroyl-glucose and caffeoyl-glucose. In the case of cinnamoyl-glucose, an additional cochromatography was undertaken with an authentic sample provided by W. Schwab (*27*). The quantitatively most important monomeric flavanol of 'Elsanta' strawberry is catechin followed by its oligomeric derivatives procyanidin B3 and procyanidin B1 identified by their reaction with 4-dimethylaminocinnamic aldehyde (*18, 19*) and by comparing their characteristics (UV spectra, retention times) with standards. The flavonols found in the fruits were quercetin-glucuronide, kaempferol-glucuronide, kaempferol-acetyl-glucoside, and kaempferol-coumaroyl-glucoside as previously found (*23, 26*).

**Table 2.** Contents of Phenolic Compounds (mg $g^{-1}$ of dry weight) in the Fruit Skin of Strawberries *Fragaria x ananassa* 'Elsanta' of Different Maturity Stages[a]

| | mean | | |
|---|---|---|---|
| | unripe $n = 29$ | ripe $n = 30$ | overripe $n = 30$ |
| **Phenolic Compounds** | | | |
| total benzoic acids | 0.090 a | 0.261 b | 0.340 c |
| total cinnamic acids | 0.397 a | 1.141 b | 1.348 b |
| total flavanols | 4.004 b | 2.425 a | 2.844 a |
| total anthocyanins | 0.111 a | 5.717 b | 10.171 c |
| total flavonols | 0.147 a | 0.522 b | 0.888 c |
| **Single Phenolic Compounds** | | | |
| galloyl-ester | 0.090 a | 0.261 b | 0.340 c |
| caffeoyl-glucoside | 0.055 a | 0.233 b | 0.243 b |
| *p*-coumaroyl-glucose | 0.038 a | 0.060 ab | 0.092 b |
| cinnamoyl-glucose | 0.019 a | 0.598 b | 0.687 b |
| cinnamic acid derivate | 0.284 a | 0.250 a | 0.326 a |
| procyanidin B3 | 1.204 a | 0.817 b | 0.942 ab |
| procyanidin B1 | 0.377 a | 0.204 a | 0.286 a |
| catechin | 2.423 a | 1.404 b | 1.616 b |
| pelargonidin-glucoside | 0.064 a | 4.942 b | 8.550 c |
| cyanidin-glucoside-succinate | 0.009 a | 0.189 ab | 0.358 b |
| pelargonidin-malonyl-glucoside | 0.038 a | 0.586 b | 1.263 c |
| quercetin-glucuronide | 0.113 a | 0.158 a | 0.504 b |
| kaempferol-glucuronide | 0.000 a | 0.244 a | 0.212 a |
| kaempferol-acetyl-glucoside | 0.019 a | 0.121 b | 0.171 c |
| kaempferol-coumaroyl-glucoside | 0.014 a | n.d. | n.d. |

[a] The data within rows were tested using Tukey-test. Different letters indicate significant differences within rows at the 5% level. n.d. = not detectable.

The concentration of cinnamic acids and flavonols are almost the same in the skin and inner fruit flesh (data not shown). The flavanols and anthocyanins were particularly located in the outer cell layers, while minor contents were found in the fruit flesh. During fruit maturation (**Table 2**) the content of flavanols (e.g., catechin) decreased, whereas the anthocyanins accumulated (e.g., pelargonidin-glucoside, pelargonidin-malonyl-glucoside).

In the group of the cinnamic acids, the cinnamoyl-glucose was increased in the ripe and overripe fruits. Furthermore, the flavonols quercetin-glucuronide, kaemperol-glucuronide, and kaempferol-acetyl-glucoside increased during ripening.

**Nondestructive Determination of Phenolic Compounds in Strawberries by Means of Fluorescence Spectroscopy.** PLS regression analysis was carried out on the fluorescence spectra of the fruits (*X* matrix) and the chemically analyzed content (*y* vector) of (i) p-coumaroyl-glucose, and (ii) cinnamoyl-glucose. The two phenolic compounds selected have shown strong fluorescence in the cochromatographic analyses. They belong to the hydroxycinnamic acids that are known to be able to fluoresce (*28, 29*). Different preprocessing methods for the spectral data were compared with respect to the regression coefficients of calibration ($r_c^2$) and the root mean squares errors of cross validation (rmsecv). For the evaluation of the calibration models the loadings were analyzed with regard to the wavelength chosen for the PLS regression.

For the calibration based on strawberry fluorescence spectra and the content of cinnamoyl-glucose, the models based on raw or smoothed spectral data resulted in rmsecv values of 58.93% and 59.23%, respectively, and coefficients of determination of $r_c^2 = 0.72$ and $r_c^2 = 0.65$ using three latent variables (LV).

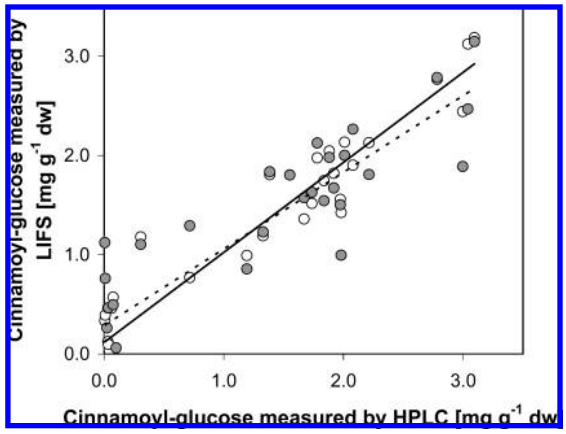


**Figure 3.** Calibration model built on the spectral data of strawberries and their content of cinnamoyl-glucose after preprocessing the fluorescence spectra by mean centering (full circle, dashed line; regression equation, $y = 0.77x + 0.28$; $r_c^2 = 0.773$) and scaling (empty circle, solid line; regression equation, $y = 0.91x + 0.12$; $r_c^2 = 0.908$).

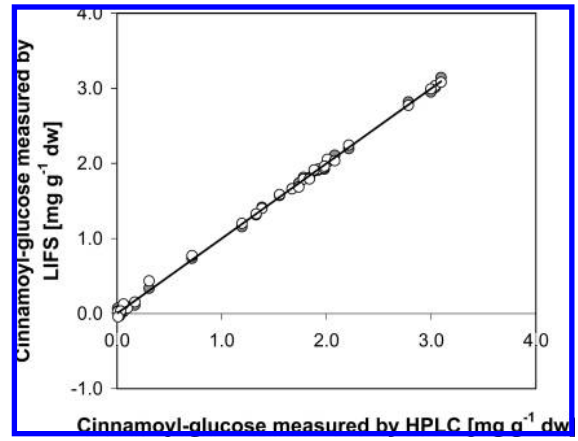


**Figure 4.** Calibration model built on the spectral data of strawberries and their content of cinnamoyl-glucose after preprocessing the fluorescence spectra by mean centering (full circle, dashed line; regression equation, $y = 0.99x + 0.00$; $r_c^2 = 0.999$) and scaling (empty circle -; regression equation: $y = 0.99x + 0.00$; $r_c^2 = 0.999$) and subtracting 1 DOSC factor.

This finding can be explained by the assumption that the noise, which was reduced by smoothing, was not the main reason for the measuring uncertainty in the apparent fluorescence spectra measured. On the other hand, derivation generally leads to increasing noise in the spectrum. Consistently a higher rmsecv value (76.07%) has been found together with an increased number of latent variables (LV = 4) for the model based on the second derivative of fluorescence spectra. However, the decision on the appropriate treatment regarding smoothing and the opportunity to apply derivative spectroscopy depends on the experimental setup used. When the calibration model using mean-centered fluorescence spectra was compared to scaled spectral data, mean centering resulted in a lower coefficient of determination ($r_c^2 = 0.77$) than scaling ($r_c^2 = 0.91$) with a slightly increased measuring uncertainty with rmsecv = 61.45% and rmsecv = 60.38%, respectively, using three latent variables (**Figure 3**). Mean centering the spectra before calibration is generally preferred to scaling, because spectral data variables (wavelengths) with high variations are assumed to be more important for the data interpretation than variables with smaller changes and should not be reduced in their valency. However, in the present study, no dominant peak occurred, but a sum signal. Concluding, the calibration models using smoothing for noise reduction, derivation to correct baseline shifts and standardization methods dealing with unequal weighting to preprocess the fluorescence spectral data did not improve or only partly improved the calibration in comparison to the calibration on raw data, but still leads to little coefficients of determination and high errors. Building calibration models on strawberry fluorescence data and their content of *p*-coumaroyl-glucose using the same different preprocessing methods confirm this conclusion.

On the basis of the assumption that the phenolic compound information contained in the fluorescence spectra may be influenced by reabsorption effects or masked by other autofluorescent compounds the direct orthogonal signal correction (DOSC) algorithm established by Westerhuis et al. (*22*) was tested to improve the calibration models. The DOSC algorithm (*7, 30*) as well as different orthogonal signal correction (*31, 32*) methods have shown promising results e.g. in near-infrared spectroscopy when applied as data preprocessing method to filter and correct the variations in the spectral data matrix that were not related to the dependent variable *y*. A following PLS regression after DOSC preprocessing led to a less complex and assumably more robust model.

The calibration models built on strawberry DOSC-corrected fluorescence spectra and the content of *p*-coumaroyl-glucose or cinnamoyl-glucose resulted in high coefficients of determination ($r_c^2 > 0.90$). When the spectra have been mean centered before correcting with the DOSC algorithm the errors of cross validation were reduced and only two latent variables were kept. This is in accordance to the preprocessing used by Westerhuis et al. (*22*) compared to scaling as preprocessing step. Thereby the rmsecv values for the models for the p-coumaroyl-glucose content showed only minor differences when removing one (rmsecv = 4.62%) or two (*rmsecv* = 4.69%) DOSC factors after mean centering from the spectral data. Similarly, marginal differences have been found when calculating models based on scaled and DOSC corrected fluorescence spectra using one (rmsecv = 17.42%) or two (rmsecv = 14.79%) DOSC factors, while keeping three LVs. The calibration models using the nondestructive measured fluorescence data and the strawberry cinnamoyl-glucose content showed similar results (**Figure 4**). When the spectral data matrix has been mean centered before the correction with DOSC and calculating the PLS regression, the models needed only two latent variables instead of three LV. When scaling was applied, the rmsecv values were higher when subtracting one (mean centering, rmsecv = 5.46%; scaling, rmsecv = 12.55%) or two DOSC factors (mean centering, rmsecv = 4.89%; scaling, rmsecv = 10.38%). To avoid overfitting of the calibration model, we did not use more than two DOSC factors (*22*).

A comparison of the original fluorescence spectra of the strawberries with those after applying the direct orthogonal signal correction algorithm without preprocessing (**Figure 5**) shows that the spectral data where one DOSC factor has been removed was now graded according to the content of cinnamoyl-glucose in the different maturity stages (**Table 3**). The corrected fluorescence spectra by subtracting one DOSC factor with respect to the content of the p-coumaroyl-glucose confirm these results. Similar results for the use of orthogonal signal correction methods compared to other preprocessing methods have been found in application on plant extracts (*33*). Considerable increasing content of anthocyanins (e.g., pelargonidin-glucoside, **Table 2**) in the higher maturity stages of the strawberries and the marginal changes in the flavonol content (e.g., quercetin-

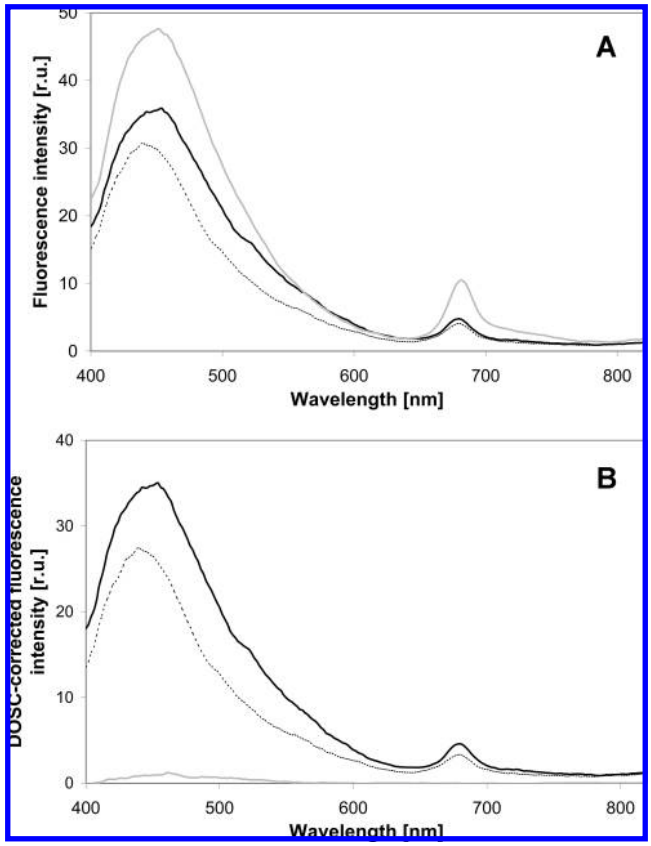


**Figure 5.** Smoothed average fluorescence spectra of (**A**) unripe (grey line; $n = 12$), ripe (black solid line; $n = 9$) and overripe (black dashed line; $n = 11$) fruits and (**B**) after removing one DOSC factor, calculated for the cinnamoyl-glucose content, from the spectral data matrix without preprocessing.

**Table 3.** Statistical Data of the Content of *p*-Coumaroyl-Glucose/Cinnamoyl-Glucose (mg $g^{-1}$ of dry weight) in Strawberries of Different Maturity Stages

| *p*-coumaroyl-glucose/ cinnamoyl-glucose | unripe | ripe | overripe |
|---|---|---|---|
| | Calibration Set ($n = 32$) | | |
| mean | 0.07/0.05 | 0.22/2.23 | 0.22/1.61 |
| standard deviation | 0.05/0.05 | 0.08/0.57 | 0.14/0.71 |
| minimum | 0.02/0.01 | 0.08/1.55 | 0.09/0.31 |
| maximum | 0.17/0.17 | 0.37/3.09 | 0.48/3.04 |
| | Test Set ($n = 35$) | | |
| mean | 0.13/0.02 | 0.20/1.61 | 0.21/1.50 |
| standard deviation | 0.09/0.03 | 0.09/0.68 | 0.11/0.90 |
| minimum | 0.03/0.00 | 0.10/0.01 | 0.08/0.01 |
| maximum | 0.29/0.11 | 0.39/2.73 | 0.41/3.03 |

glucoside, kaempferol 3-glucuronide) interfered with the variation in the fluorescence spectra with respect to the p-coumaroyl-glucose and the cinnamoyl-glucose contents. However, the increasing reabsorption effects in the more mature stages could be removed by eliminating one or two DOSC factors from the spectral data matrix. The changes of flavanol content (e.g., catechin) during maturation do not affect the fluorescence signal of the cinnamic acids (*p*-coumaroyl-glucose, cinnamoyl-glucose) because they absorb light and emit fluorescence only in the UV wavelength range (<350 nm) (*11*).

The efficiency of the different preprocessing methods was evaluated by comparing the errors of cross validation of the calibration models. However, the fruit material is a living product, which is highly dependent on endogenous and exogenous effects affecting their phenolic compounds composition as well as other factors affecting the apparent fluorescence signal. Therefore, a validation of the calibration models using an independent test set of new fruit material is essential. In the present study, strawberries from the same cultivar but different harvesting date were subjected to the validation. An influence of endogenous factors such as varying molecule allocation processes as well as exogenous factors as environmental growing conditions can be expected on the fruit material. For this purpose also the number of latent variables in the calibration models based on cross-validation results was evaluated again by comparing the rmsec and rmsecv values and the root mean squares errors of prediction (rmsep) for the first four LV. Generally, using mean centering, scaling or the DOSC algorithm as preprocessing of the fluorescence spectra of the test-set independently from the calibration data set resulted in lower validation errors.

**Table 4.** Root Mean Squares Errors of Calibration (rmsec) and Pediction (rmsep) and Coefficient of Determination of Calibration ($r_c^2$) and Validation ($r_v^2$) for Four Latent Variables (LV) of the Calibration Model Built on Fluorescence Spectra of Strawberries and Their Content *p*-Coumaroyl-Glucose and Cinnamoyl-Glucose Using Different Preprocessing Methods before Removing One DOSC Factor from the Spectral Data

| preprocessing method before DOSC | number of LV | $r_c^2$ | rmsec (%) | $r_v^2$ | rmsep (%) |
|---|---|---|---|---|---|
| | *p*-Coumaroyl-glucose | | | | |
| mean centering | 1 | 0.999 | 1.58 | 0.989 | 19.21 |
| | 2 | 0.999 | 1.04 | 0.988 | 19.19 |
| | 3 | 0.999 | 0.60 | 0.990 | 18.95 |
| | 4 | 1.000 | 0.36 | 0.990 | 19.17 |
| scaling | 1 | 0.901 | 22.57 | 0.923 | 16.69 |
| | 2 | 0.993 | 5.79 | 0.992 | 7.91 |
| | 3 | 0.998 | 3.32 | 0.981 | 10.36 |
| | 4 | 0.999 | 1.26 | 0.985 | 9.67 |
| | Cinnamoyl-glucose | | | | |
| mean centering | 1 | 0.998 | 4.08 | 0.986 | 43.02 |
| | 2 | 0.999 | 2.40 | 0.992 | 43.01 |
| | 3 | 0.999 | 0.92 | 0.995 | 43.00 |
| | 4 | 1.000 | 0.46 | 0.995 | 43.29 |
| scaling | 1 | 0.955 | 17.62 | 0.965 | 28.04 |
| | 2 | 0.996 | 5.21 | 0.984 | 24.68 |
| | 3 | 0.999 | 3.08 | 0.993 | 23.01 |
| | 4 | 0.999 | 1.39 | 0.992 | 23.07 |

Calculating the rmsep values of scaled and subsequently DOSC-corrected fluorescence spectra showed improved calibration models (rmsep < 8.00% for p-coumaroyl-glucose, rmsep < 23.10% for cinnamoyl-glucose) compared to mean centering the spectra before removing one DOSC factor (**Table 4**). The subtraction of only one DOSC factor from the preprocessed spectral data matrix using scaling showed lower validation errors than removing two DOSC factors (rmsep < 15.00% for *p*-coumaroyl-glucose, rmsep > 24.10% for cinnamoyl-glucose). The use of two latent variables gave a robust model in the case of calibrating with the p-coumaroyl-glucose content with the lowest rmsep value (7.91%) and a high coefficient of determination of validation ($r_v^2 = 0.99$). The validation of the calibration models built on the fluorescence spectra of strawberries and their content of the cinnamoyl-glucose leads to rmsep = 23.01% and $r_v^2 = 0.99$ by using three LV. In contrast, using the total phenolic compound content of both, fluorescent and nonfluorescent phenolic fruit compounds, and the scaled and DOSC corrected strawberry fluorescence spectra for the calibration models led to increased errors of validation (rmsep > 51%) compared to the models based on single fluorescent phenolic compound contents.

The coefficients of determination in the validation of the calibration models based on the fruit fluorescence spectra and their content of p-coumaroyl-glucose and cinnamoyl-glucose were always higher than 0.92. However, a rmsep value around 20% led to the assumption that even after removing non relevant information for the phenolic compound content from the spectral data with the help of the DOSC algorithm not all the variables with an interfering influence on the fluorescence signal of these two phenolic compound contents could be eliminated.

Alternatively, processing the fluorescence spectra with the help of simultaneously recorded reflectance spectra is often applied in fluorescence spectroscopy to reduce the reabsorption effects in the spectral data matrix (*16, 17*). Therefore, the fluorescence spectra were corrected by dividing the fluorescence intensity at each wavelength with the corresponding reflectance intensity in a wavelength range from 400 to 820 nm. Preprocessing the spectral data by mean centering resulted in rmsecv = 97.68% and rmsep = 95.02% and low coefficients of determinations ($r_c^2$ and $r_v^2 < 0.38$) for one LV for the content of cinnamoyl-glucose. In comparison, scaling led also to high values for the root-mean-square error of calibration (rmsecv = 92.86%) and validation (rmsep = 94.64%) and low coefficients of determination keeping two LV.

When using only a single reflectance intensity at 800 nm to correct the reabsorption effects in the fluorescence spectra for the calibration of the cinnamoyl-glucose content, the $r_v^2$ values slightly improved, but the rmsep values increased when preprocessing the spectra by mean centering (rmsep = 95.83%) using one LV and by scaling (rmsep = 123.61%) using two LV for the models. Similar results were found for the *p*-cumaroyl-glucose content. Additionally, using the DOSC algorithm on the preprocessed corrected spectra did not further improve the calibration models and the validation results.

In conclusion, using the information included in the simultaneously recorded reflectance spectra to correct reabsorption effects on the fluorescence signal leads to less robust models. It can be supposed that the use of the reflectance spectral data includes information in the spectral data matrix (*X*) that may not be relevant for the prediction of the content of *p*-coumaroyl-glucose or cinnamoyl-glucose.

To test the feasibility of using fluorescence spectroscopy to nondestructively detect the content of health-promoting compounds in fruits and vegetables, it has to be taken into account that horticultural products such as strawberries provide a complex matrix with varying optical properties leading to variation in the fluorescence signal. Particularly, reabsorption and quenching effects are influencing the fluorescence spectral data. In contrast to the general opinion that these disturbing effects are reduced when correcting the fluorescence emission spectra with the information included in the simultaneously recorded reflectance spectral data, using this often-applied approach in fluorescence spectroscopy resulted in less-robust models in the present study. In comparison by means of the direct orthogonal signal correction algorithm, two single phenolic fruit compounds (here *p*-coumaroyl-glucose and cinnamoyl-glucose) out of 18 phenolic compounds estimated by rp-HPLC in strawberry fruits were quantifiable.

The rmsep values of 8–24% were characteristic for the validation of calibration models developed for horticultural product. Therefore, the calibration models have to be improved by enlarging the calibration data set for considering the variability and the composition of fruit compound contents similar to approaches in near-infrared spectroscopy. However, the fluorescence spectroscopy seems to provide an adequate method for special purposes such as for quantification of *p*-coumaroyl-glucose and cinnamoyl-glucose in strawberry fruits compared to the time and cost-intensive, destructive, analyzing methods.

## ABBREVIATIONS USED

DOSC, direct orthogonal signal correction; rp-HPLC, reversed-phase high-performance liquid chromatography; HPLC-CRD, high-performance liquid chromatography with chemical reaction detection; HPLC-DAD, high-performance liquid chromatography with diode-array detector; LIFS, laser-induced fluorescence spectroscopy; LC-MS/MS, HPLC linked with a mass spectrometer; rmsec, root mean squares error of calibration; rmsecv, root mean squares error of cross-validation; rmsep, root mean squares error of prediction; UV, ultraviolet.

## ACKNOWLEDGMENT

The authors acknowledge Kerstin Zofka for technical assistance.

## LITERATURE CITED

(1) Hannum, S. M. Potential impact of strawberries on human health: a review of the science. *Crit. Rev. Food. Sci. Nutr.* **2004**, 1–17.
(2) Regnault Roger, C. The nutritonal incidence of flavonoids: some physiological and metabolic considerations. *Experientia* **1988**, *44*, 725–804.
(3) Santos-Buelga, C.; Scalbert, A. Proanthocyanidins and tannin-like compounds-nature, occurrence, dietary intake and effects on nutrition and health. *J. Sci. Food Agric.* **2000**, *80*, 1094–1117.
(4) Chun, O. K.; Kim, D.-O.; Smith, N.; Schroeder, D.; Han, J. T.; Lee, C. Y. Daily consumption of phenolics and total antioxidant capacity from fruit and vegetables in the American diet. *J. Sci. Food Agric.* **2005**, *85*, 1715–1724.
(5) Merzlyak, M. N.; Gitelson, A. A.; Chivkunova, O. B.; Rakitin, V. Y. Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiol. Plant.* **1999**, *106*, 135–141.
(6) Zude, M. Comparison of indices and multivariate models to non-destructively predict the fruit chlorophyll by means of visible spectrometry in apples. *Anal. Chim. Acta* **2003**, *481*, 119–126.
(7) Zude, M.; Birlouez-Aragon, I.; Paschold, P. J.; Rutledge, D. N. Non-invasive spectrophotometric sensing of carrot quality from harvest to consumption. *Postharvest Biol. Technol.* **2007**, *45*, 30–37.
(8) Krause, G. H.; Weis, E. Chlorophyll fluorescence and photosynthesis: the basics. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1991**, *42*, 313–349.
(9) Chappelle, E. W.; Wood Jr, F. M.; McMurtrey III, J. E.; Newcomb, W. W. Laser-induced fluorescence of green plants. 1: A technique for the remote detection of plant stress as species differentiation. *Appl. Opt.* **1984**, *23*, 134–138.
(10) Lichtenthaler, H. K.; Lang, M.; Sowinska, M.; Heisel, F.; Miehé, J. A. Detection of vegetation stress via a new high resolution fluorescence imaging system. *J. Plant Physiol.* **1996**, *148*, 599–612.
(11) Wulf, J. S. Polyphenol Analysis in Horticultural Products Based on Laser-Induced Fluorescence Spectroscopy. PhD Thesis, Faculty of Agriculture and Horticulture, Humboldt University of Berlin, Berlin, Germany, 2006.
(12) Cheikhousman, R.; Zude, M.; Bouveresse, D. J.; Rutledge, D. N.; Birlouez-Aragon, I. Fluorescence spectroscopy for monitoring deterioration of extra olive oil during heating. *Anal. Bioanal. Chem.* **2005**, *382*, 1438–1443.
(13) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*; 3rd ed.; Springer; New York, 2006; p 9.
(14) Bro, R.; van den Berg, F.; Thybo, A.; Andersen, C. M.; Jorgensen, B.; Andersen, H. *Trends Food Sci. Technol.* **2002**, *13*, 235–244.
(15) Geladi, P. Some recent trends in the calibration literature. *Chemom. Intell. Lab. Syst.* **2002**, *60*, 211–224.

(16) Gitelson, A. A.; Buschmann, C.; Lichtenthaler, H. K. Leaf chlorophyll fluorescence corrected for re-absorption by means of absorption and reflectance measurements. *J. Plant Physiol.* **1998**, *152*, 283–296.

(17) Löhmannsröben, H. G.; Schober, L. Combination of laser-induced fluorescence and diffuse-reflectance spectroscopy for the in situ analysis of diesel-fuel-contaminated soils. *Appl. Opt.* **1999**, *38*, 1404–1410.

(18) Treutter, D.; Santos-Buelga, C.; Gutmann, M.; Kolodziej, H. Identification of flavan-3-ols and procyanidins by high-performance liquid chromatography and chemical reaction detection. *J. Chromatogr.* **1994**, *667*, 290–297.

(19) Treutter, D. Chemical reaction detection of catechins and proanthocyanidins with 4-dimethylaminocinnamic aldehyde. *J. Chromatogr.* **1989**, *467*, 185–193.

(20) Roemmelt, S.; Peterek, S.; Treutter, D.; Andreotti, C.; Costa, G.; Halbwirth, H.; Rademacher, W.; Speakman, J. B.; Sponza, G.; Tortoreto, L.; Bazzi, C.; Zimmermann, N.; Stich, K.; Forkmann, G. Alteration of phenylpropanoid biosynthesis of fruit trees as a tool for enhancement of fire blight resistance. *Acta Hortic.* **2002**, *590*, 477–484.

(21) Savitzky, A.; Golay, M. Smoothing and differentiation of data by simplified least square procedures. *Anal. Chem.* **1964**, *36*, 1627–1639.

(22) Westerhuis, J. A.; de Jong, S.; Smilde, A. K. Direct Orthogonal Signal Correction. *Chemom. Intell. Lab. Syst.* **2001**, *56*, 13–25.

(23) Määttä-Riihinen, K. R.; Kamal-Eldin, A.; Törrönen, A. R. Identification and quantification of phenolic compounds in berries of *Fragaria* and *Rubus* species (family *Rosaceae*). *J. Agric. Food Chem.* **2004**, *52*, 6178–6187.

(24) Lunkenbein, S.; Coiner, H.; De Vos, C. H. R.; Schaart, J. G.; Boone, M. J.; Krens, F. A.; Schwab, W.; Salentijn, E. M. J. Molecular Charcterization of a Stable Antisense Chalcone Synthase Phenotype in Strawberry (*Fragaria X ananassa*). *J. Agric. Food Chem.* **2006**, *54*, 2145–2153.

(25) Seeram, N. P.; Lee, R.; Scheuller, S.; Heber, D. Identification of phenolic compounds in strawberries by liquid chromatography electrospray ionization mass spectroscopy. *Food Chem.* **2006**, *97*, 1–11.

(26) Zheng, Y.; Wang, S.; Wang, C. Y.; Zheng, W. Changes in strawberry phenolics, anthocyanins, and antioxidant capacity in response to high oxygen treatments. *Food Sci. Thechnol.* **2007**, *40*, 49–57.

(27) Schwab, J. F.; Klucas, R. V.; Grayer, R. J.; Abian, J.; Harborne, J. B. Characterization of Phenolic Glucosides from Soybean Root Nodules by Ion-Exchange High-Performance Liquid Chromatography, Ultraviolet Spectroscopy and Electrospray Mass Spectrometry. *Phytochem. Anal.* **1998**, *9*, 171–176.

(28) Buschmann, C.; Langsdorf, G.; Lichtenthaler, H. K. Imaging of the blue, green, and red fluorescence emission of plants: An overview. *Photosynthetica* **2000**, *38* (4), 483–491.

(29) Morales, F.; Cartelat, A.; Alvarez-Fernandez, A.; Moya, I.; Cerovic, Z. G. Time-resolved spectral studies of blue-green fluorescence of artichoke (Cynara cardunculus L. Var. Scolymus) leaves: Identification of chlorogenic acid as one of the major fluorophores and age-mediated changes. *J. Agric. Food Chem.* **2005**, *53* (25), 9668–9678.

(30) Luypaert, J.; Heuerding, S.; de Jong, S.; Massart, D. L. An evaluation of direct orthogonal signal correction and other preprocessing methods for the classification of clinical study lots of a dermatological cream. *J. Pharm. Biomed. Anal.* **2002**, *30*, 453–466.

(31) Sjöblom, J.; Svensson, O.; Josefson, M.; Kullberg, H.; Wold, S. An evaluation of orthogonal signal correction applied calibration transfer of near infrared spectra. *Chemom. Intell. Lab. Syst.* **1998**, *44*, 229–244.

(32) Sáiz-Abajo, M. J.; González-Sáiz, J. M.; Pizarro, C. Orthogonal signal correction applied to the classification of wine and molasses vinegar samples by near-infrared spectroscopy. Feasibility study for the detection and quantification of adulterated vinegar samples. *Anal. Bioanal. Chem.* **2005**, *382*, 412–420.

(33) Qu, H.; Ou, D.; Cheng, Y. Background correction in near-infrared spectra of plant extracts by orthogonal signal correction. *J. Zhejiang Univ. Sci.* **2005**, *6*, 838–843.

**Received for review August 20, 2007. Revised manuscript received March 7, 2008. Accepted March 13, 2008. The authors thank the Deutsche Forschungsgemeinschaft for financial support of the project HESPERIDES.**

JF072495I