JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Rapid Assessment of Quality Parameters in Processing Tomatoes Using Hand-Held and Benchtop Infrared Spectrometers and Multivariate Analysis

Elizabeth D. Wilkerson,[†] Gordon E. Anthon,[‡] Diane M. Barrett,[‡] Glynda Fe G. Sayajon,[†] Alejandra M. Santos,[†] and Luis E. Rodriguez-Saona*,[†]

[†]Department of Food Science & Technology, The Ohio State University, 110 Parker Food Science and Technology Building, 2015 Fyffe Road, Columbus, Ohio 43210, United States

[‡]Department of Food Science & Technology, University of California—Davis, 1 Shields Avenue, Davis, California 95616, United States

ABSTRACT: Two portable infrared sensors were evaluated for the rapid determination of quality parameters in processing tomatoes. A total of 370 hot-break juices were prepared from ~40 processing tomato varieties grown in 5 California counties. The levels of sugars, acids, soluble solids, titratable acidity, and pH in these juices were determined using standard reference methods. Juices were processed, filtered, and directly applied to the FT-IR crystal (15−40 μL) to obtain spectra. Partial least-squares regression (PLSR) was used to generate correlation models, both calibration and validation. The PLS validation models showed good ability (R_{val} > 0.80; <10% SEP) in estimating the sugars, acids, and especially soluble solids in tomato for both the transmission DialPath portable system and benchtop unit using triple-bounce attenuated total reflectance (ATR). The IR portable unit may provide the tomato processing industry with an efficient method for in-plant, high throughput quantification of quality parameters in tomatoes.

KEYWORDS: *infrared spectroscopy, tomato, sugar, acid, soluble solids*

INTRODUCTION

Tomato (*Lycopersicon esculentum*) is the second most important vegetable crop next to potato with production occurring in 144 countries.[1] Fresh and processed tomatoes are one of the mostly widely produced and consumed vegetables in the U.S. and generate annual an revenue of approximately $2 billion in farm cash receipts.[2]

In the U.S., the state of California currently produces about 12 million tons of processing tomatoes annually, accounting for more than 90% of the U.S. crop and about 40% of the processing tomato production worldwide.[3] Processing tomatoes are used to produce a variety of products, including whole peeled fruit and diced tomatoes, as well as various juices and purees. The largest portion of this crop is thermally processed and concentrated into tomato paste. The most important quality attributes in processing tomatoes are soluble solids, pH, titratable acidity, viscosity, and color.[4] Tomato fruit composition is approximately 93% water and 7% solids. Total solids are further classified according to their water solubility as soluble and insoluble solids.[5] Approximately half of the total solids are reducing sugars, with slightly more fructose than glucose. The remaining solids consist of organic acids (citric and malic), amino acids, proteins, lipids, minerals, pectic substances, cellulose, and hemicellulose.[6]

Soluble solids are a key parameter in tomato paste production. Tomato paste is produced and sold based on its soluble solids content; thus, soluble solids dictate the factory yield. Higher soluble solids in the incoming fruit means that fewer tons of tomatoes will be needed to produce a given amount of paste.[7] Furthermore, water removal during evaporation of juice to paste is an energy intensive process. Producing paste from fruit with high levels of soluble solids is less expensive since less water needs to be removed to obtain the desired soluble solids content.[8]

A second key parameter in paste production is pH, which plays a vital part in microbiological safety and food spoilage. Tomatoes are high-acid foods and thus require less drastic thermal treatments than foods classified as low-acid (pH > 4.6) for the destruction of spoilage microorganisms to ensure food safety.[9] Generally, the pH of tomatoes has been reported to range from 3.9 to 4.9 or in standard cultivars from 4.0 to 4.7;[10] hence, the USDA standards of identity allow organic acids to be added to lower the pH when needed during the processing of high-acid foods.[6] The pH of the tomato is determined by its organic acid content with citric acid being the most abundant.

Sugars and organic acids are responsible for the sweetness and tartness, and are major factors affecting flavor acceptability.[11−14] The development of processing tomato varieties with altered compositions requires the efficient and accurate evaluation of thousands of tomato breeding lines.[15] Traditional methods for sugar and acid determination include the use of enzymatic kits and chromatography.[16] Chromatographic methods are typically accurate and allow for the determination of multiple juice components from a single sample. Limitations include time-consuming sample preparation, use and disposal

Received: November 21, 2012
Revised: February 3, 2013
Accepted: February 3, 2013
Published: February 4, 2013

of hazardous solvents, low sample through-put, and a high skill-set for the testing personnel. Enzyme based methods are rapid, require little sample preparation, and allow for a high sample throughput.[17] However, by their nature these tests can only determine a single juice component at a time. If a complete profile of the juice is desired, multiple tests will need to be run, greatly increasing the time and expense involved with these methods. Infrared (IR) spectroscopy presents an ideal alternative for assessing processing tomatoes as it is a simple, time- and cost-efficient technique that can potentially provide information on many juice components from a single measurement. It has already shown promise in analyzing other food and agricultural products.[18] The mid-IR region (4000−400 cm^{-1}) produces absorption bands for most functional groups, which allows for a direct correlation between specific chemical parameters of interest.[19] Fourier-transform infrared (FT-IR) techniques combined with chemometrics offer tomato processors and breeders powerful tools for the rapid assessment of tomato quality attributes. Portable IR units would enable the food manufacturer to quickly assess the quality of the incoming raw material, the quality of their product, and allow for timely corrective measures during manufacture. Portable systems are simple to use and require minimal or no sample preparation, thus reducing assay time and helping to streamline the analytical procedure so that it is more applicable to higher sample throughput.

The objective of the present research was to develop a simple, quick, and reliable methodology for the determination of processing tomato quality parameters (°Brix, pH, titratable acidity, fructose and glucose, and citric and glutamic acids). Additionally, we evaluated the performance of two novel portable infrared units against benchtop IR spectrometers.

MATERIALS AND METHODS

Tomato Plant Material. A total of 40 processing tomato varieties were obtained from five counties (Fresno, Kern, Merced, San Joaquin, and Yolo) located in central California during the 2010 and 2011 growing seasons. The counties provided one to three replicates of each variety totaling 370 samples. In each California county, tomatoes were planted by cooperating commercial growers as part of a long-term program of evaluation of new tomato cultivars under the coordination of the Department of Food Science and Technology and the Cooperative Extension Program at the University of California—Davis.[20] All tomatoes were manually harvested at the commercial red-ripe stage during midseason (mid-August through early October). Tomatoes were selected from the middle of the plant, avoiding the top and bottom set of tomatoes.

Sample Preparation. All tomatoes were washed, towel dried, and sorted for defects. A "microwave break" method, developed by the Department of Food Science and Technology at University of California—Davis[21] to simulate a hot break process, was used to prepare tomato juice samples as described previously.[22] The juice samples were immediately evaluated for soluble solids, pH, and titratable acidity (% citric). The remaining portion of the blended tomato juices were stored in 15 mL Fisherbrand plastic centrifuge tubes (Waltham, MA) at −40 °F for later chemical and FTIR analysis. During all analytical stages, special care was taken to protect samples from unnecessary light and exposure to heat.

Reference Analysis: Sugars, Acids, and Soluble Solids. *Enzymatic Determination of Sugars and Acids.* Sugars and acids were quantified using an enzymatic procedure as previously described.[17] Analysis was done using enzyme reagent kits (R-Biopharm, Marshall, MI) for use in 96-well microplates with a final assay volume of 200 μL. The procedure followed kit instructions except that the volumes of water used to prepare the reagents were modified so that the final reagent concentrations were the same for the 200 μL volume used here as for the 3 mL procedure described in the kit instructions. Tomato juice samples were clarified by centrifuging for 5 min at 16,100*g*. Aliquots of the supernatants were diluted 100-fold with water for acid analysis and 1000-fold with water for sugar analysis. To perform the assay, 100 μL aliquots of the diluted supernatants were mixed with 100 μL of the modified kit reagents in the microplate wells. Absorbance at 340 nm was measured, then 4 μL of the appropriate enzyme suspension was added and the absorbance at 340 nm monitored until a stable new reading was obtained. Concentrations of sugars and acids were calculated from the absorbance differences at 340 nm and the extinction coefficient for NADPH of 6,300 $M^{-1}cm^{-1}$. Standard solutions provided with the kits were used to verify the accuracy of the method.

pH and Titratable Acidity. Pulped juice tomato samples were evaluated for titratable acidity using titration with NaOH.[23] The remaining juice was then deaerated and the temperature was adjusted to 25 °C before pH determination.

Soluble Solids (°Brix) Determination. For estimation of soluble solids content, 1.5 mL of tomato puree was centrifuged at 10,000 rpm (15 min, 25 °C), and the supernatant was filtered through Whatman nonsterile syringe filters (0.45 μm). The filtered tomato serum (40 μL) was measured using the Leica Mark II Plus Abbe Refractometer Model 10494 (Leica, Buffalo, NY).

Measurements were taken once for each sample, and 70% ethanol was used to clean in between samples. The refraction index was expressed as % soluble solids in °Brix.

Infrared Spectroscopy Analysis. Samples were divided into calibration and validation sets to compare the performance of the regression models developed by collecting mid-infrared (MIR) spectra with different systems. A total of 210 samples were analyzed from the 2010 growing season and an additional 160 samples from the 2011 growing season. Two thirds of the samples (245 tomato juices) were used for calibration models, and the remaining one-third (125 tomato juices) was used for prediction models. Samples were randomly selected from each county for inclusion in calibration and validation models using a random sampling approach (Matlab 6.1).

Aliquots (1.5 mL) from each thawed, blended tomato juice sample were centrifuged at 10,000 rpm for 15 min at 25 °C. The supernatant was filtered through Whatman nonsterile syringe filters (0.45 μm) and collected in 1.5 mL Fisherbrand centrifuge tubes. Spectral data from all tomato juice samples were collected using 3 spectrometers: a benchtop system equipped with a single-bounce and triple-bounce zinc selenide (ZnSe) attenuated total reflectance (ATR) accessories, a hand-held (FlexScan, Agilent, Santa Clara, CA) unit equipped with a single-bounce diamond ATR, and a Cary 630 portable IR spectrometer using a ZnSe dial-path transmittance accessory.

The filtered juice (15−50 μL) was applied directly at the surface of the crystals (ATR or ZnSe transmittance) for spectral acquisition in less than 2 min. Duplicate, independent measurements were taken on each sample, and background spectra were acquired every sample for the hand-held and portable units or every 5 samples for the benchtop model to account for environmental variations. In-between measurements, the crystal was carefully cleaned with 70% ethanol and dried with Kimwipe tissue (Kimberly-Clark Corp. LLC, Roswell, GA).

The benchtop model used was an Excalibur Series 3100 Fourier-Transform infrared spectrometer (Agilent Technologies Inc., Santa Clara, CA) with a potassium bromide beam splitter and a deuterated triglycine sulfate (DTGS) detector, operating at 4 cm^{-1} resolution. A horizontal ATR sampling accessory, coupled with a ZnSe crystal plate with a refractive index of 2.5, allowed triple reflection within the sample at an incidence angle of 45° for the highest infrared sample throughput (Pike Technologies, Madison, WI). An additional ZnSe crystal was obtained for single-bounce benchtop analysis. The samples were scanned at room temperature, and mid-infrared spectra were collected at wavenumbers from 4000 to 700 cm^{-1}. Subsequently, the sample spectra were corrected against the background spectrum of air. Interferograms of 32 scans were co-added followed by Beer-Norton apodization. Spectra were displayed in terms of absorbance and viewed with Win-IR Pro Software (Agilent Technologies Inc., Santa Clara,

Table 1. Reference Method Results of Quality Parameters in Tomato Samples from the 2010 and 2011 Growing Seasons

county	number of varieties		glucose (g/L)	fructose (g/L)	citric acid (g/L)	glutamic acid (g/L)	pH[a]	soluble solids (°Brix)	titratable acidity[a] (% citric)
Kern	42	range	13.2–21.0	13.8–20.6	2.20–3.86	1.41–2.57	4.31–4.76	5.1–6.7	0.171–0.302
		average ± SD	15.0 ± 2.2	15.4 ± 2.3	3.03 ± 0.40	2.19 ± 0.49	4.59 ± 0.12	5.41 ± 0.73	0.253 ± 0.034
Yolo	40	range	13.2–21.4	13.3–20.1	1.97–3.76	1.00–2.50	4.22–4.77	3.5–5.4	0.185–0.348
		average ± SD	16.3 ± 2.5	16.6 ± 2.3	2.64 ± 0.51	1.84 ± 0.45	4.51 ± 0.10	4.65 ± 0.46	0.253 ± 0.038
Merced	42	range	12.1–19.4	12.6–20.0	1.52–3.59	1.92–3.08	4.34–4.73	5.0–6.4	0.179–0.335
		average ± SD	16.0 ± 2.2	16.6 ± 2.4	2.34 ± 0.38	2.26 ± 0.35	4.55 ± 0.10	5.58 ± 0.34	0.232 ± 0.032
San Joaquin	38	range	10.0–18.5	11.0–18.1	1.29–2.53	1.54–2.61	4.39–4.87	3.2–5.4	0.152–0.283
		average ± SD	13.7 ± 1.7	14.6 ± 1.7	1.98 ± 0.30	2.19 ± 0.33	4.57 ± 0.10	4.35 ± 0.49	0.218 ± 0.031
Fresno	39	range	12.3–18.6	12.8–20.0	1.98–3.39	1.68–3.21	4.30–4.66	5.0–6.2	0.184–0.338
		average ± SD	17.3 ± 2.4	17.7 ± 2.4	2.68 ± 0.36	2.21 ± 0.43	4.57 ± 0.14	5.59 ± 0.34	0.265 ± 0.037

[a]TA and pH results were unreliable for samples from the 2011 growing season; therefore, results display 2010 samples only.

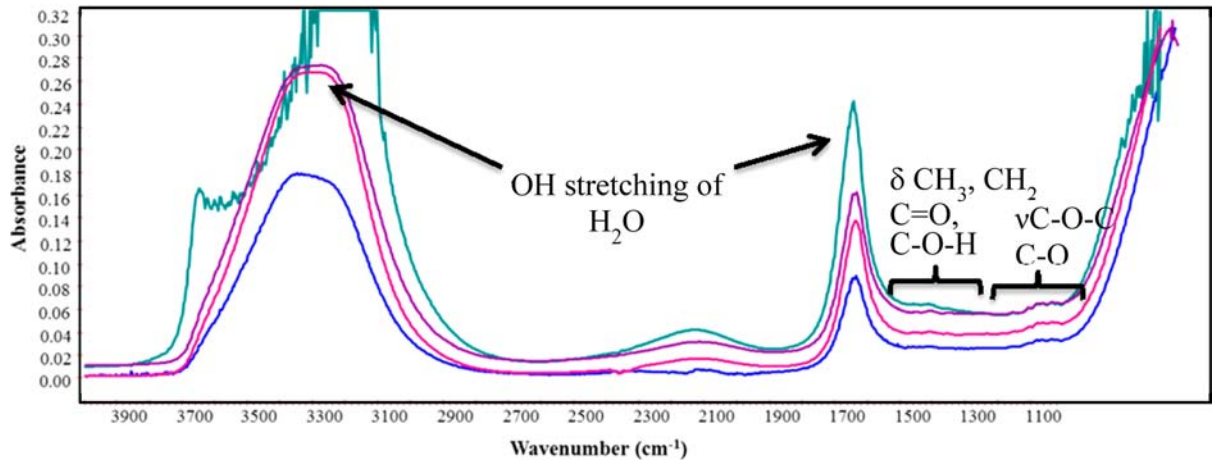

Figure 1. Infrared absorption spectrum of the tomato samples on benchtop (single-bounce ATR in pink and triple-bounce ATR in purple), FlexScan hand-held (blue), and Cary portable (green) spectrometer systems.

CA). The instrument was constantly purged with CO_2-free dry air from a CO_2RP140 dryer (Dominick Hunter, Charlotte, NC).

The portable Cary 630 FTIR unit (Agilent Technologies Inc., Santa Clara, CA) used a 30 μm transmission sampling accessory with a temperature-stabilized DTGS detector, ZnSe beam splitter, operated at a spectral resolution of 4 cm^{-1}, and the mid-infrared spectral range of 4000 − 700 cm^{-1} was collected. A total of 64 scans were co-added to improve the signal-to-noise ratio.

The hand-held FlexScan FTIR unit (Agilent Technologies Inc., Santa Clara, CA) used a single-bounce diamond ATR crystal with a ZnSe beam splitter and a temperature-stabilized DTGS detector operating at 4 cm^{-1} resolution. The samples were scanned at room temperature using the 4000–700 cm^{-1} region. Interferograms of 64 scans were co-added to improve the signal-to-noise ratio.

Multivariate Analysis: Partial Least Squares Regression (PLSR). PLSR, a pattern recognition technique, was used to analyze the complex data sets and to generate predictive models to estimate the different tomato juice quality parameters. PLSR is a bilinear regression method that compresses a large number of variables into a few much smaller number of latent factors that are linear combinations of the spectral frequencies (X) and use these factors to ascertain for the analyte's concentration (Y), explaining much of the covariance of X and Y.[24] This technique has the potential to estimate the component concentration, as well as the chemical properties of the spectra.[25] The degree of competence and versatility of this analytical approach offer a more information-rich data set of reduced dimensionality and eliminates data noise, which results in more accurate and reproducible calibration models.[26]

Spectral data were exported from the spectrometers as GRAMS.spc files and imported to the multivariate statistical program Pirouette for Windows Chemometrics Modeling Software, version 4.0 (Infometrix, Inc., Bothell, WA). These were then analyzed by PLSR that was cross-validated using a leave-one-out approach to generate calibration models and subsequently transformed using the multiplicative scatter correction (MSC) function. The calibration models correlated the spectra against the concentration of each tomato analyte (Brix, pH, titratable acids, glucose, fructose, and citric and glutamic acids). Performance of these models was evaluated in terms of outlier diagnostics, standard error of cross-validation (SECV), correlation coefficient (R^2), and number of factors.

Graphic evaluation of the calibration models was done to guarantee an arbitrary distribution of residuals. Residuals and leverage were used for the evaluation of outliers. An outlier was indicated either by a large residual or an unusual residual pattern. Conversely, leverage was used to determine its potential contribution to the estimated calibration model. Hence, any observation with atypical and large residual or leverage was reanalyzed and eliminated if it was considered a substantial outlier; thereafter, the model was recalculated. Standard error of cross-validation (SECV) is an approximation of the standard error of prediction, that is, the weight of an anticipated error when independent samples are predicted using the model.[27] External validation was performed on the prediction sample set to assess the ability of the calibration model to withstand unknown variability. The

Table 2. Performance Statistics for PLS Regression Models Generated for Quality Parameters in Processing Tomatoes from the 2010 Growing Season on FT-IR Systems[a]

analyte	technique	N	range of concn	factors	SECV[b]	R_{cv}[c]
glucose (g/L)	benchtop: 3 bounce[d]	205	10.0−21.4	4	1.13	0.76
	benchtop: 1 bounce[d]	200		4	1.16	0.77
	portable: Cary[e]	195		4	1.12	0.80
	hand-held: FlexScan[e]	180		5	1.20	0.76
fructose (g/L)	benchtop: 3 bounce[d]	205	11.0−20.6	3	1.11	0.73
	benchtop: 1 bounce[d]	190		4	1.08	0.73
	portable: Cary[e]	200		4	1.11	0.73
	hand-held: FlexScan[e]	190		5	1.17	0.73
citric acid (g/L)	benchtop: 3 bounce[d]	200	1.29−3.86	9	0.21	0.88
	benchtop: 1 bounce[d]	190		8	0.23	0.89
	portable: Cary[e]	210		7	0.25	0.88
	hand-held: FlexScan[e]	190		5	0.32	0.80
glutamic acid (g/L)	benchtop: 3 bounce[e]	205	1.00−3.21	8	0.19	0.90
	benchtop: 1 bounce[d]	190		9	0.19	0.88
	portable: Cary[e]	195		9	0.15	0.93
	hand-held: FlexScan[e]	170		5	0.21	0.84
titratable acidity (% citric)	benchtop: 3 bounce[d]	205	0.152−0.348	9	0.014	0.92
	benchtop: 1 bounce[e]	190		10	0.015	0.93
	portable: Cary[e]	205		9	0.018	0.88
	hand-held: FlexScan[e]	180		6	0.025	0.74
pH	benchtop: 3 bounce[d]	200	4.22−4.87	8	0.04	0.82
	benchtop: 1 bounce[e]	190		9	0.04	0.83
	portable: Cary[e]	195		9	0.05	0.81
	hand-held: FlexScan[e]	160		6	0.05	0.61
°Brix	benchtop: 3 bounce[e]	200	4.2−6.7	8	0.17	0.96
	benchtop: 1 bounce[d]	190		8	0.20	0.96
	portable: Cary[e]	200		7	0.19	0.96
	hand-held: FlexScan[e]	170		5	0.25	0.93

[a]Note: Models were generated using multiple specular component transformation. Smoothing was used with FlexScan and benchtop (single bounce) systems to reduce noise interferences. [b]SECV: standard error of cross-validation. [c]R_{cv}: correlation coefficient of cross-validation. [d]Spectral region of 1800−900 cm^{-1} was included. [e]Spectral region of 1500−900 cm^{-1} was included.

correlation coefficient (R^2) is a statistical measure that allows us to determine the amount of variation in the data that is adequately modeled by the calibration equation as a fraction of the total variation. An R^2 of 1.0 indicates that the calibration model represents 100% of the variance within the data and is illustrated by a regression line perfectly fitting the data.[27] The calibration model that generated an excellent combination of the minimum SECV, higher R^2, and optimized numbers of latent factors was selected the best for spectral data set on each tomato parameter.

RESULTS AND DISCUSSION

Quality Parameters of Tomato Fruit. The large number of tomato varieties (~40) grown in 5 different California counties resulted in wide ranges of compositional levels (Table 1). Overall, values ranged from 10.0 to 21.4 g/L (glucose), 11.0−20.6 g/L (fructose), 1.29−3.86 g/L (citric acid), 1.00−3.21 g/L (glutamic acid), 3.2−6.7 (°Brix), 4.22−4.87 (pH), and 0.152−0.348 (% citric) (Table 1). Similar contents were reported in other processing tomato juice studies.[3,28,22] The data also is consistent with findings from Daood[29] where 0.90−1.62 g/100g for glucose, 1.25−1.70 g/100g for fructose, and 6.04 mg/g citric acid nutrient levels in tomato fruits were reported. In general, the values found in this study are within those reported in the literature taking into account that nutrient levels may be affected by variety, maturity, temperature, and soil nutrients among others.[30]

Calibration Model Development Using Tomato Juice Samples from the 2010 Growing Season. Tomato samples from the 2010 growing season were screened on the benchtop (single-bounce and triple-bounce ATR accessories), FlexScan hand-held, and Cary portable spectrometer systems. Typical ATR and transmission spectra obtained from tomato juice supernatants (no pulp) showed strong water absorption bands (1582−1692 and 2971−3627 cm^{-1}). As seen in Figure 1, the ATR triple-bounce accessory used on the benchtop spectrometer provided increased absorbance intensity through multiple interactions between the incident IR radiation and the sample which improved the signal resolution of low concentration components.[31] Important bands in the fingerprint region (1500−900 cm^{-1}) were associated with C−O and C−C stretching modes (900−1153 cm^{-1}) and O−C−H, C−C−H, and C−O−H bending vibrational modes (1474−1199 cm^{-1}).[32,33] The Cary portable system generated higher absorption intensity compared to the single- and triple-bounce ZnSe ATR benchtop system (Figure 1). The effective path length (EPL) obtained by using the multireflection ATR accessory was ~13.08 μm,[31] a 3-fold increase from a single-bounce ATR accessory (EPL = 4.6 μm), while the DialPath transmittance accessory for the Cary 630 system provided an EPL of 30 μm for data acquisition. The saturated signals between 3800 and 2800 cm^{-1} are due to the strong infrared absorptivity of water (104.4 M^{-1} cm^{-1}) and are unfortunately void of any information.[34]

The cross-validated leave-one-out PLSR model performance statistics for the 2010 tomato samples on the benchtop,

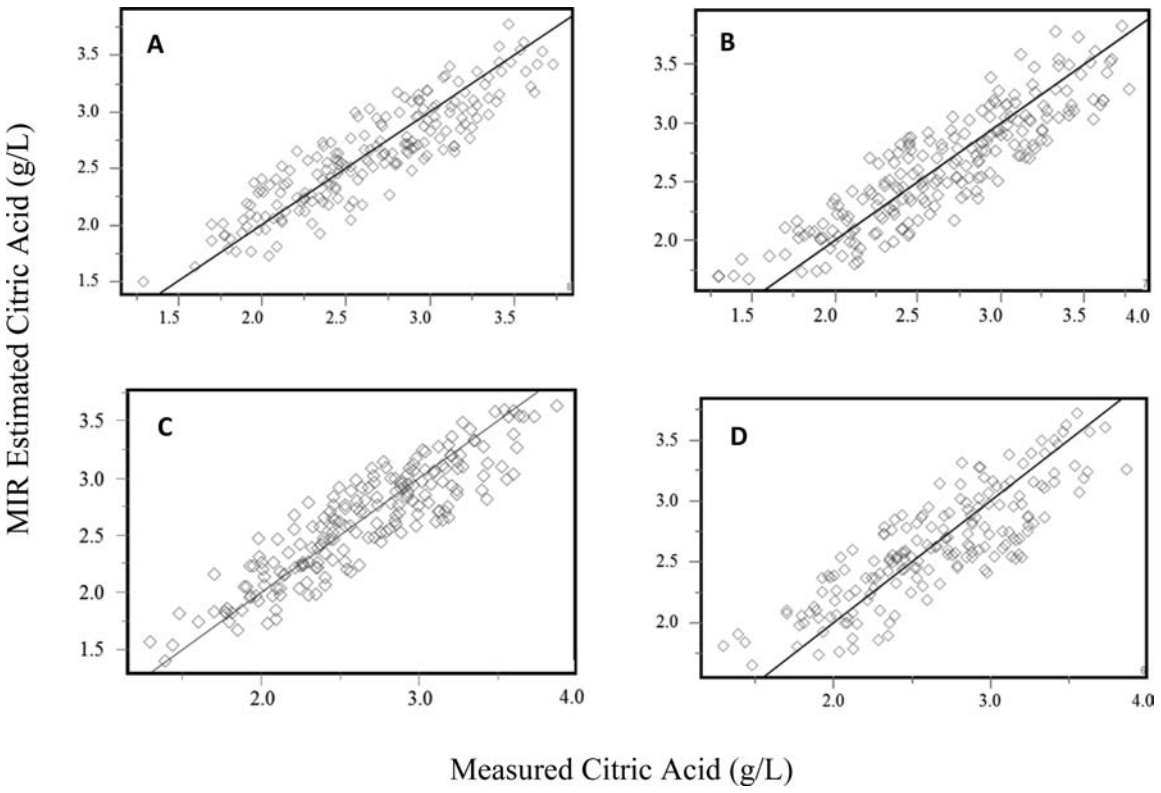

Figure 2. Generated PLSR calibration models for citric acid on the (A) single-bounce ATR benchtop, (B) triple-bounce ATR benchtop, (C) portable Cary, and (D) hand-held FlexScan.

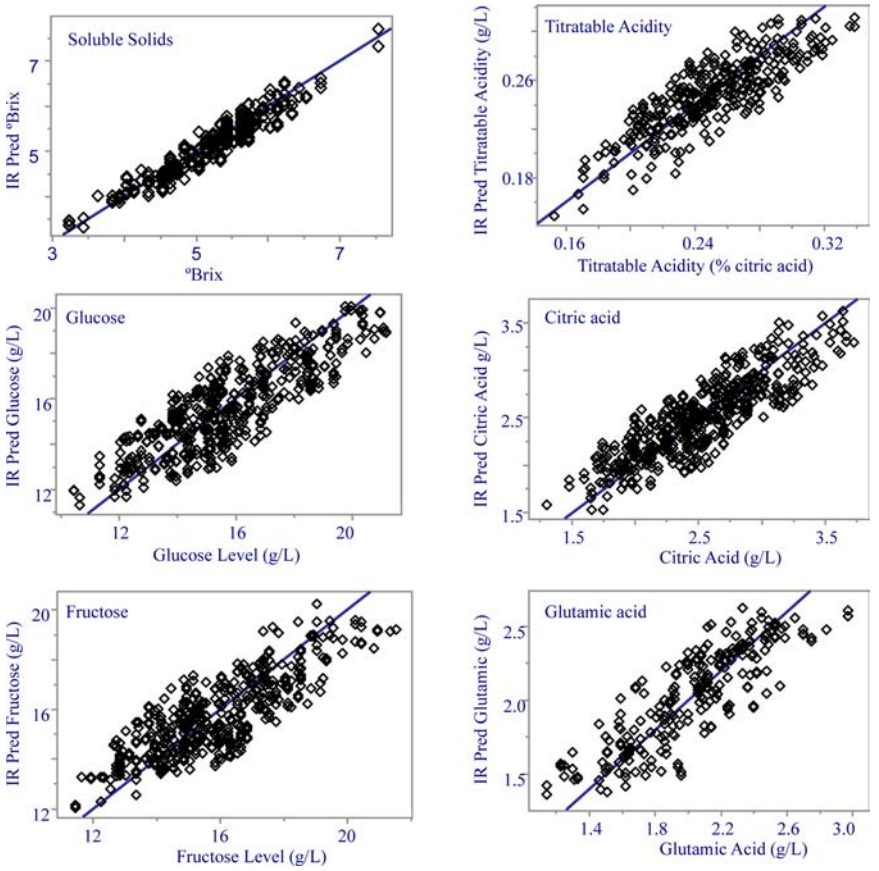

Figure 3. PLSR correlation plots for estimating quality parameters in processing tomatoes using a portable Cary 630 unit equipped with a transmittance Dialpath accessory.

dx.doi.org/10.1021/jf304968f | J. Agric. Food Chem. 2013, 61, 2088−2095

Table 3. Calibration and Validation Performance Statistics for PLS Regression Models Generated for Quality Parameters in Processing Tomatoes from the 2010−2011 Growing Season on Benchtop (Triple-Bounce ATR) and Cary Portable Systems

analyte	technique	N	range of concn	factors	SECV[a]	R_{cv}[b]	SEP[c]	R_{val}[d]
glucose (g/L)	benchtop: 3 bounce[e]	285	10.0−21.4	8	1.20	0.85	1.47	0.83
	portable: Cary[g]	345		4	1.22	0.85	1.14	0.88
fructose (g/L)	benchtop: 3 bounce[e]	285	11.0−20.6	3	1.19	0.82	1.23	0.80
	portable: Cary[g]	325		4	1.14	0.81	1.21	0.80
citric acid (g/L)	benchtop: 3 bounce[f]	270	1.29−3.86	8	0.21	0.86	0.21	0.85
	portable: Cary[g]	360		8	0.24	0.87	0.24	0.88
glutamic acid (g/L)	benchtop: 3 bounce[f]	270	1.00−3.21	7	0.22	0.84	0.25	0.79
	portable: Cary[g]	330		9	0.18	0.88	0.19	0.86
°Brix	benchtop: 3 bounce	300	4.2−6.7	8	0.20	0.94	0.23	0.92
	portable: Cary[g]	370		6	0.21	0.95	0.21	0.96

[a]SECV: standard error of cross-validation. [b]R_{cv}: correlation coefficient of cross-validation. [c]SEP: standard error of prediction. [d]R_{val}: correlation coefficient of validation. [e]Spectral region included was 1800−900 cm^{-1}. [f]Spectral regions included were 1800−1600 and 1500−900 cm^{-1}. [g]Spectral region included was 1500−900 cm^{-1}.

portable, and hand-held FTIR systems are displayed in Table 2. By selecting the optimum number of latent variables (factors) that minimized the standard error of cross-validation (SECV) of the model, ∼95−99% of the cumulative variance for the systems was explained. The number of factors (3−10) used in the models reduced spectral noise and potential overfitting, which could impair its ability to estimate composition in unknown samples.

In general, the single- and triple-bounce ATR accessories used in the benchtop and portable Cary systems performed similarly to each other in terms of the correlation coefficient of cross-validation (R_{cv}) and SECV (Table 2). The SECV for the quality parameters on the benchtop and Cary portable systems were ∼1.1 g/L (sugars), 0.23 g/L (citric acid), 0.20 g/L (glutamic acid), 0.16% citric (TA), 0.05 pH, and 0.18 °Brix (soluble solids) (Table 2). The FlexScan hand-held system showed slightly inferior performance to the other systems as seen by the lower R-value and higher SECV for nearly all quality parameters (Table 2).

Overall, a high number of outliers were identified from the FlexScan unit based on leverage and studentized residual analysis (Table 2). Interferences from the water regions' strong absorption bands (1582−1692 cm^{-1} and 2971−1395 cm^{-1}) are likely reducing the signal-to-noise ratio in the FlexScan, thus masking the signal for the bending (O−C−H, C−C−H, and C−O−H) vibrational modes in the 1500−1200 cm^{-1} region. The estimated quality parameter contents measured by the ATR-IR or transmission spectroscopy showed coefficient of determination (R_{cv}) ranging from 0.61 to 0.96 with reference values (Table 2).

Figure 2 shows the PLSR models generated for citric acid on the single and triple-bounce ATR benchtop, the portable Cary 630 system with DialPath transmission accessory, and the hand-held FlexScan diamond ATR system. Figure 3D illustrates that higher data scattering, or a higher SECV (Table 2), was observed from the FlexScan hand-held as the data points were more scattered along the regression line. The Cary 630 portable system provided much stronger correlation as seen in the tighter predictions or lower SECV (Figure 2B). These models are helpful visualizations of the relationship between measured quality parameters and predictive capability of the infrared systems.

Recalibration Model Development Using Tomato Juice Samples from the 2010 and 2011 Growing Season by Selected FT-IR Systems. On the basis of the performance of PLSR models to estimate quality parameters in tomato juices from the 2010 growing season, the triple-bounce ATR containing benchtop and portable Cary FT-IR systems were selected to develop a validated model to rapidly assess processing tomatoes. The benchtop acted as control technology, and the portable Cary 630 unit was compared to the performance of the benchtop. PLSR calibration models were built using two-thirds of the processing tomato samples from the 2010 and 2011 growing seasons. The included spectral range for the Cary portable FT-IR was 1500−900 cm^{-1} to exclude saturated water absorption signals. The 1800−1500 cm^{-1} signal was not saturated for the benchtop system; therefore, some or all parts of this region were included when developing the models (Table 3). Multiplicative scatter correction (MSC) transformation was applied to the data as a preprocessing treatment. Calibration model summary statistics for five quality parameters (glucose, fructose, citric acid, glutamic acid, and °Brix) are presented in Table 3. Reference data for titratable acidity and pH were not reliable for tomato juices from the 2011 growing season and therefore were not included in calibration model development (Table 1).

Overall, PLSR calibration models generated by tomato juices from the 2010−2011 seasons yielded higher correlation coefficients (R_{cv} = 0.81−0.95) than previously found with 2010 samples only (R_{cv} = 0.73−0.96, Table 2) and similar SECV levels. This indicated that by increasing the number and diversity of the samples, one could allow for the reduction of impact from irrelevant spectral-variations (noise) in the calibration model. This capability provided a more information-rich data set of reduced dimensionality and eliminated data noise that resulted in more accurate and reproducible calibration models.[35−37] Both instruments displayed very similar, and adequate, estimating capacity for all sugar and acid quality parameters (R_{cv} = 0.82−0.94 benchtop, 0.81−0.95 portable (Figure 3)), which indicated that the selected principal components were modeling about 81−95% of the variance within the samples.[27]

The use of loadings plots enabled the identification of portions of the original spectra that were important for discrimination. Loadings identified which areas of the IR spectra were more related to sample variation; when they were particularly large (either below or above zero), they denote areas of the original spectra that were important for discrimination. The most relevant spectral region for sugars was 1200−1000 cm^{-1}, which includes the C−O and C−C

stretching modes of carbohydrates.[38] The most intense peaks for sugars were 1062 cm^{-1} on the benchtop system and 1082 cm^{-1}, characteristic of the C−O stretch vibration. The band at 1062 cm^{-1} is in accordance with results reported by Sivakesava and Irudayaraj.[39] Prominent peaks for citric acid discrimination included the 1020−1105 cm^{-1} region for both systems and, additionally, 1716 cm^{-1} on the benchtop system, which reflected the carbonyl stretch (C═O) vibration.[32] Interestingly, relevant spectral regions for glutamic acid were in the 1020−1105 cm^{-1} region but also included a prominent band at 1405 cm^{-1} for symmetric stretching of the carboxylic group.[40] Discrimination for soluble solids utilized attributes of the spectral regions of both acids and sugars, relying on prominent peaks in the 1000−1130 cm^{-1} region and a small peak at 1406 cm^{-1}.

External Validation Using an Independent Set of Tomato Juices. External validation of the PLSR models, obtained with the calibration samples, was performed using the remaining third of tomato juice samples from 2010 to 2011 growing seasons. Although several authors have reported that SECV gives a realistic estimate of the error of prediction of samples not included in the calibration,[41−43] this step is necessary to obtain an independent measurement of the equation's accuracy, expressed as SEP, i.e., standard error of prediction.[44] Examination of the validation statistics shown in Table 3 reveal similar SECV and SEP values for MIR predictions of all quality parameters examined, indicating that the application of these MIR models of both systems gave robust predictions under practical conditions.[45] It is a normal finding that the standard errors obtained by validation were slightly higher than those of cross-validation.[27]

The PLS regression models showed good ability in estimating the sugars, acids, and especially soluble solids in tomato for both triple-bounce ATR benchtop and Cary 630 portable FT-IR systems judging by the high R_{val} and low SEP results (Table 3). Comparatively, a study completed by Schibisz and others[38] obtained similar mid-IR modeling results in quantifying sugars, citric acid, and soluble solids in tomatoes using a benchtop system, with reported R_{val} >0.92 for glucose, fructose, citric acid, and soluble solids in tomatoes and standard error of predictions of 0.87 g/L, 1.04 g/L, 0.39 g/L, and 1.84°Brix, respectively. Other works have been reported by Pedro and Ferreira[46] supporting the use of benchtop FT-IR systems for assessing tomato quality; however, no research has yet been reported on the use of portable systems. Interestingly, the Cary 630 unit equipped with a transmission DialPath accessory provided superior performance in estimating sugars, acids, and soluble acids in processing tomatoes compared to the single-bounce ZnSe ATR benchtop and FlexScan hand-held that utilized a single-bounce diamond ATR, and showed performance similar to that of the triple-bounce ZnSe ATR benchtop (Table 3).

In summary, our findings support the use of a portable FTIR with a transmission sampling accessory for rapid assessment of quality parameters in processing tomatoes. Novel portable FT-IR systems may provide the tomato processing industry with a rapid method to evaluate processing tomatoes with equivalent levels of reliability and sensitivity as benchtop systems but allow for more flexibility since the unit can be easily carried and transferred.

AUTHOR INFORMATION

Corresponding Author

*The Ohio State University, 2015 Fyffe Court, Columbus, OH 43210. Phone: 614-292-3339. E-mail: rodriguez-saona.1@osu.edu.

Funding

This work was supported by a grant from the California League of Food Processors (CLFP).

Notes

The authors declare no competing financial interest.

REFERENCES

(1) FAOSTAT Statistical Database; Food and Agriculture Organization of the United Nations. http://faostat.fao.org/faostat (accessed 2004).

(2) United States Department of Agriculture. Economic Research Service: Tomato. http://www.ers.usda.gov/topics/crops/vegetables-pulses/tomatoes.aspx#processing (accessed Aug 2012).

(3) Anthon, G. E.; LeStrange, M.; Barrett, D. M. Changes in pH, acids, sugars and other quality parameters during extended vine holding of ripe processing tomatoes. *J. Sci. Food Agric.* **2011**, *91*, 1175−1181.

(4) Salveit, M. E. Fruit Ripening and Fruit Quality. In *Tomatoes*; Heuvelink, E., Ed.; CABI Publishing: The Netherlands, 2005; pp 145−170p.

(5) Nielsen, S. S. *Food Analysis*, 2nd ed.; Aspen Publishers: Gaithersburg, MD, 1998.

(6) Barringer, S. A. Vegetables: Tomato Processing. In *Food Processing: Principles and Applications*; Smith, J. S., Hui, Y. H., Eds.; Blackwell Publishing, Ltd.: Oxford, UK, 2004.

(7) Gould, W. A. *Tomato Production, Processing and Technology*, 3rd ed.; CTI Publishers: Baltimore, MD, 1992.

(8) Nichols, M. A. Towards 10 t/ha Brix. *Acta Hort.* **2006**, No. ISHS, 724217−724223.

(9) Draft Guidance for Industry: Acidified Foods, 2010. U.S. Department of Health and Human Services; Food and Drug Administration Center for Food Safety and Applied Nutrition. http://www.fda.gov/Food/GuidanceComplianceRegulatoryInformation/GuidanceDocuments/AcidifiedandLow-AcidCannedFoods/ucm222618.htm (accessed Jan 2013).

(10) Sapers, G. M.; Douglas, F. W., Jr.; Ziolkowski, M. A.; Miller, R. L.; Hicks, K. B. Determination of ascorbic acid, dehydroascorbic acid and ascorbic acid-2-phosphate in infiltrated apple and potato tissue by high-performance liquid chromatography. *J. Chromatogr., A* **1990**, 503431−503436.

(11) Baldwin, E. A.; Goodner, K.; Plotto, A. Interaction of volatiles, sugars, and acids on perception of tomato flavor and flavor descriptors. *J. Food Sci.* **2008**, *73* (6), 294−307.

(12) De Bruyn, J. W.; Garretsen, F.; Kooistra, E. Variation in taste and chemical composition of the tomato (*Lycopersicon Esculentum* Mill.). *Eupytica* **1971**, *20*, 214−227.

(13) Stevens, M. A.; Kader, A. A.; Albright-Halton, M.; Algazi, M. Genotypic variation for flavor and composition in fresh market tomatoes. *J. Am. Soc. Hortic Sci.* **1977**, *102* (5), 680−689.

(14) Petró-Turza, M.; Teleky-Vámossy, G. Study on taste substances of tomato Part 3. Sensory evaluations. *Food/Nahrung* **1989**, *33* (5), 387−394.

(15) De Nardo, T.; Shiroma-Kian, C.; Halim, Y.; Francis, D.; Rodriguez-Saona, L. E. Rapid and simultaneous determination of lycopene and ß-carotene contents in tomato juice by infrared spectroscopy. *J Agric. Food Chem.* **2009**, *57* (4), 1105−12.

(16) Horowitz, W. *Official Methods of Analysis of AOAC International*. AOAC International: Gaithersburg, MD, 2000, Chapters 37, 44, 45.

(17) Vermeir, S.; Nicolai, B. M.; Jans, K.; Maes, G.; Lammertyn, J. High-throughput microplate enzymatic assays for fast sugar and acid

quantification in apple and tomato. *J Agric. Food Chem.* **2007**, *55*, 3240−3248.
(18) Chen, P.; Sun, Z. A review of non-destructive measures for quality evaluation and sorting of agricultural products. *J Agric. Eng. Res.* **1991**, *49*, 85−98.
(19) Ellis, D. I.; Goodacre, R. Metabolic fingerprinting in disease diagnosis: biomedical applications of infrared and Raman spectroscopy. *Analyst (Cambridge, U.K.)* **2006**, *131* (8), 875−885.
(20) Murray, M.; Cahn, M.; Caprile, J.; May, D.; Miyao, G.; Mullen, B.; Valencia, J.; Weir, B. University of California cooperative extension processing tomato cultivar evaluation program. *HortTechnology* **1999**, *9* (1), 36−39.
(21) Leonard, S.; Buhlert, J.; Marsh, G.; Wolcott, T.; Heil, J. R. *Procedures for Evaluating Utilization Potential and Quality in Processing Tomatoes and Tomato Products*; Department of Food Science and Technology: Davis, CA, 1980.
(22) Garcia, E.; Barrett, D. M. Evaluation of processing tomatoes from two consecutive growing seasons: quality attributes, peelability, and yield. *J. Food Proc. Pres.* **2006**, *30*, 20−36.
(23) AOAC International. *Official Methods of Analysis*, 17th ed.; AOAC Int.: Gaithersburg, MD, 2000.
(24) Martens, H.; Martens, M. *Multivariate Analysis of Quality: An Introduction*, 1st ed.; John Wiley and Sons: New York, NY, 2001; p 445.
(25) Haaland, D. M.; Thomas, E. V. Partial least squares methods for spectral analyses. Relation to other quantitative calibration methods and extraction of qualitative information. *Anal. Chem.* **1998**, *60*, 1193−1208.
(26) Wold, S.; Trygg, J.; Berglund, A.; Antti, H. Some recent developments in PLS modeling. *Chemom. Intell. Lab. Syst.* **2001**, *58* (2), 131−150.
(27) Mark, H.; Workman, J. *Statistics in Spectroscopy*; Academic Press, Inc.: San Diego, CA, 1991; pp 294−300.
(28) Barrett, D. M.; Weakley, C.; Diaz, J. V.; Watnik, M. Qualitative and nutritional differences in processing tomatoes grown under commercial organic and conventional production systems. *J. Food Sci.* **2007**, *72* (9), 441−451.
(29) Daood, H. G. Ion-pair chromatography and photodiode-array detection of vitamin C and organic acids. *J. Chromatogr. Sci.* **1994**, *32* (11), 481.
(30) Gould, W. A. *Tomato Production and Processing and Quality Evaluation*; The AVI Publishing Company, Inc.: Westport, CT, 1974; p 441.
(31) Pike Technologies. ATR Theory and Applications: Application note 0401. http://www.piketech.com/files/pdfs/ATRAN611.pdf (accessed Oct 10, 2012).
(32) Stuart, B. *Infrared Spectroscopy: Fundamentals and Applications*; John Wiley & Sons: Chichester, England, 2004; p 224.
(33) Irudayaraj, J.; Tewari, J. Simultaneous monitoring of organic acids and sugars in fresh and processed apple juice by fourier transform infrared-attenuated total reflection spectroscopy. *Appl. Spectrosc.* **2003**, *57* (12), 1599.
(34) Grdadolnik, J. An attenuated total reflection infrared spectroscopy of water solutions. *Int. J. Vib. Spectrosc.* **2004**, *6*, 2.
(35) Martens, H.; Naes, T. *Multivariate Calibration*; John Wiley & Sons, Ltd.: West Sussex, U.K., 1989.
(36) Warnock, D. G.; Peck, C. C. A roadmap for biomarker qualification. *Nat. Biotechnol.* **2010**, *28* (5), 444−445.
(37) De Maesschalck, R. C.; Masart, A.; Heuerding, D. L. Decision Criteria for Soft Independent Modeling of Class Analogy Applied to Near Infrared Data. *Chemom. Intell. Lab. Syst.* **1999**, *47*, 65−77.
(38) Scibisz, I.; Reich, M.; Bureau, S.; Gouble, B.; Causse, M.; Bertrand, D.; Renard, C. Mid-infrared spectroscopy as a tool for rapid determination of internal quality parameters in tomato. *Food Chem.* **2011**, *125*, 1390−1397.
(39) Sivakesava, S.; Irudayaraj, J. Determination of sugar in aqueous mixtures using mid-infrared spectroscopy. *Appl. Eng. Agric.* **2000**, *16* (5), 543−550.
(40) Barth, A. Review: the infrared absorption of amino acid side chains. *Prog. Biophys. Mol. Biol.* **2000**, *74*, 141−173.
(41) Meuret, M.; Dardenne, P.; Biston, R.; Poty, O. The use of NIR spectroscopy in predicting nutritive value of Mediterranean tree and shrub foliage. *J. Near Infrared Spectrosc.* **1993**, *11*, 45−54.
(42) Shenk, J. S. Calibration the ISI Way. In *Near-Infrared Spectroscopy: The Future Waves*; Davies, A. M. C., Williams, P. C., Eds.; NIR Publications: Chichester, England, 1996; pp 198−202.
(43) Martens, H. A.; Dardenne, P. Validation and verification of regression in small data sets. *Chemom. Intell. Lab. Syst.* **1998**, *44*, 99−121.
(44) Windham, W. R.; Mertens, D. R.; Barton, F. E., II. Protocol for NIRS Calibration: Sample Selection and Equation Development and Validation. In *Near-Infrared Spectroscopy (NIRS): Analysis of Forage Quality, Agriculture Handbook*; Martens, G. C., Shenk, J. S., Barton, F. E., II, Eds.; USDA-ARS: Washington, DC, 1989; Vol. *643*, pp 96−103.
(45) Fontaine, J.; Horr, J.; Schirmer, B. Near-infrared reflectance spectroscopy enables the fast and accurate prediction of the essential amino acid contents in soy, rapeseed meal, sunflower meal, peas, fishmeal, meat meal products, and poultry meal. *J Agric. Food Chem.* **2001**, *49*, 57−66.
(46) Pedro, A.; Ferreira, M. Simultaneously calibrating solids, sugars, and acidity of tomato products using PLS2 and NIR spectroscopy. *Anal. Chim. Acta* **2007**, *595* (1−2), 221−227.