

Rapid Prediction of Amylose Content of Polished Rice by Fourier Transform Near-Infrared Spectroscopy

Jin-Cheol Lee, Yeon-Hee Yoon, Sun-Min Kim, Byong-Sik Pyo, Fu-hung Hsieh¹, Hak-Jin Kim², and Jong-Bang Eun^{3*}

Biotechnology Industrialization Center, Dongshin University, Naju, Jeonnam 520-811, Korea

¹Biological Engineering Department, University of Missouri, Columbia, MO 65211, USA

²National Institute of Agricultural Engineering, Rural Development Administration, Suwon, Gyeonggi 441-744, Korea

³Department of Food Science and Technology and Institute of Agricultural Science and Technology, Chonnam National University, Gwangju 500-757, Korea

Abstract Fourier transform near-infrared (FT-NIR) spectroscopy and partial least squares (PLS) regression were used to predict the amylose content of polished rice. Spectral reflectance data in a wavelength range of 1,000 to 2,500 nm were obtained with a commercial spectrophotometer for 60 different varieties of Korean rice. For a comparison of this spectroscopic method to a standard chemical analysis, the amylose contents of the tested rice samples were determined by the iodine-blue colorimetric method. The highest correlation for the rice amylose ($R^2=0.94$, standard error of prediction=0.20% amylose content) was obtained when using the FT-NIR spectrum data pre-treated with normalization, the first derivative, smoothing, and scattering correction.

Keywords: Fourier transform near-infrared (FT-NIR), partial least squares (PLS), amylose content, polished rice

Introduction

Since rice, the staple food in Korea (1), imports to Korea have increased in recent years, it is imperative that Korean farmers produce good quality rice to enhance the competitiveness and meet consumer needs. However, according to World Trade Organization agreements, the opening of the Korean rice market to foreign countries is an inevitable fact. In order to survive in the global market, enhancing the quality of milled rice should occur prior to completing the opening of the rice market (2). Even though rice quality is evaluated based on a combination of subjective and objective factors, in general, rice of good quality has the following characteristics: the rice grain is fully matured; the kernel is transparent, glossy, and has a traditional Korean rice flavor and taste; the grain is free from impurities; and the grain has proper low protein and amylose contents. In Korea, the amylose content of good quality rice usually ranges from 17 to 20% (3). It is well known that the amylose content is directly related to water absorption, volume expansion, fluffiness, and separability of the cooked grains, and it is inversely related to cohesiveness, tenderness, and glossiness (4, 5). Therefore, monitoring the amylose content of rice can provide useful information for the efficient production of high quality rice. Yet, conventional methods for amylose analysis, such as the colorimetric method (6-8), are costly and time consuming.

Due to their attractive advantages such as rapid and non-destructive measurements, near-infrared (NIR) spectroscopy techniques have been attempted by many

researchers to discriminate the geographical origins of green tea, and to predict various food properties, including protein, lipid, starch, and amylose, (9-12). Czuchajowska *et al.* (13) reported that four wavelengths in the NIR region could be used to determine barley amylose and starch contents. Moon *et al.* (6) and Hwang *et al.* (7) developed prediction models to estimate rice amylose content using two (2,044 and 2,208 nm) and six wavelengths (1,772, 1,784, 1,580, 1,716, 1,728, and 1,724 nm), respectively. However, the predictions by the developed models were poor, showing low correlation coefficients and high standard errors.

The main limitations of the NIR technique are the use of wavelengths in a relatively narrow range, and the strong dependence of reflectance on the scattering properties of the sample (14). To overcome these problems, several researchers have attempted to use Fourier transform near-infrared (FT-NIR) spectroscopy because it can improve spectral reproducibility and wave number precision (15). Predicting rice amylose content from spectral data requires the application of multivariate calibration techniques. Among the various multivariate techniques, including multiple linear regression (MLR), principle component regression (PCR), and partial least squares (PLS) regression, PLS has been widely used as a regression method for determining the chemical properties of food samples using FT-NIR spectroscopy data (16) due to its superior predictive capability (17).

The objective of this study was to investigate the feasibility of using FT-NIR spectroscopy for the determination of amylose content in Korean rice. Toward that goal, a prediction model was developed using PLS regression by relating the FT-NIR spectrum data to the amylose contents, and its predictive capability was evaluated by comparing the quantity of rice amylose that was obtained with the

*Corresponding author: Tel: 82-62-530-2145; Fax: 82-62-530-2149

E-mail: jbeun@chonnam.ac.kr

Received March 19, 2007; accepted April 24, 2007

spectroscopic method and that obtained by a standard analysis via simple linear regression.

Materials and Methods

Materials Forty-seven varieties of rice (*Oryza sativa* L.: Japonica type), which had been harvested at the experimental field of Jeonnam Agricultural Research and Extension Services (JARES), Naju, Korea in 2002, were used in this study. The rice paddies were de-hulled using a rice sheller (SY88-TH; Ssangyong Machine Ind. Co., Ltd., Incheon, Korea) and then milled using a rice mill (MC90A; Toyo Co., Ltd., Wakayama, Japan) to produce polished rice, 92% in the degree of milling. After vacuum packaging (Vacuum packing machine, SW-600; Leetech System, Seoul, Korea), the samples were stored in a deep freezer at -85°C . An additional 13 varieties of rice were purchased from a local market, and stored at the same conditions until use. Before analysis, the samples were brought to ambient temperatures (about 23°C) by keeping them overnight at ambient conditions.

Amylose analysis *Sample preparation:* The rice samples for analysis were first ground using a blender (51BL31; Waring Commercial, Torrington, CT, USA) and then a sample mill (Cyclotec 1093; Foss Tecator AB, Hoeganaes, Sweden) to pass thorough a 100-mesh sieve. The rice flour was then vacuum-packaged and kept in a deep freezer at -85°C until use.

Amylose determination: The amylose contents were determined by the iodine colorimetric procedure described by Juliano (18, 19). Rice flour (100 mg) was weighed and placed in a 100 mL flask, and 1 mL of 95% ethanol and 9 mL of 1 M NaOH were added. The suspension was thoroughly mixed and gelatinized for 10 min in a boiling water bath. After cooling, the suspension was diluted to 100 mL by distilled water. Then, 1 mL of 1 M acetic acid and 2 mL of iodine reagent were added to 5 mL of the solution, which was again diluted to 100 mL with distilled water. After 20 min of standing, the absorption of this solution at 620 nm was determined by a UV-VIS spectrophotometer (UV-1650; Shimadzu, Kyoto, Japan). The amylose contents were estimated from a standard curve, constructed from solutions of rice amylose (Type III; Sigma Chemical Co., St. Louis, MO, USA) at various concentrations (20).

FT-NIR spectrum measurements All FT-NIR spectra were recorded using an NIRLab FT-NIR system (NIRLab N-200/MCS 100; Büchi, Postfach, Switzerland). Approximately 100 g of rice were placed into an NIR measurement cell with a glass petri dish base. After carefully excluding foreign substances, 10 random samples for each variety were taken, and each sample was measured three times. The raw spectra for each sample were averaged and used for the regression analysis.

Reflectance spectra (R) in the range of 1,000 to 2,500 nm were collected using a computer program (NIRLab Ware Ver. 3.0, Büchi). They were converted to absorbance for regression analysis by taking the logarithm of the reciprocal of reflectance or $\log(1/R)$.

Multivariate Analyses Calibration: Partial least squares (PLS) regression, implemented in the NIRCal Chemometric Software (Ver. 4.21, Buchi), was performed on the spectra and their corresponding chemical constituents to develop calibration models (16, 17). Two-thirds of the averaged spectra were used as the calibration dataset, which included both the highest and lowest spectrum to encompass the calibration range. The remaining one-third of the averaged spectra was used as the validation dataset to verify the validity of the calibration equation (21). In addition, all of the chemical analysis data were included in the range of the calibration dataset, and the validation data were impartially scattered on the calibration dataset acceptable above 95% in order to avoid extrapolation (22). The wavelength used for regression analysis was 1,000–2,500 nm, the same as the NIRS measuring region. The values obtained from the spectrum data were classified as primary and secondary. The number of the former was determined by the minimized X-PRESS (predicted residual error sum square of spectra X) value, and the latter was the bias and standard error of prediction (SEP), within the primary factor range. The raw spectrum data were pretreated with mathematical techniques such as derivatives and multiplicative scatter correction (MSC) to reduce parameters of the base line statistical treatment and variations, due to the physical elements of the samples and light scattering.

The optimal conditions were selected during each, or combinative, pretreatment processing (e.g., the first and second derivatives, normalization, smoothing, MSC, etc.) when error values were determined, and minimized with the selected conditions (23–25). The predictive capabilities of the developed calibration models were evaluated in terms of the standard error of calibration (SEC), SEP, coefficient of determination (R^2), and bias (25, 26). The developed models were expressed as the first function graph of calibration and validation. Finally, accuracy of the calibration equation selected as an optimum model was examined by relating the values that were determined by the calibration model, and by the chemical analysis via a simple linear regression.

Results and Discussion

Amylose contents of polished rice Table 1 shows the amylose contents of the different varieties of polished rice as obtained by the colorimetric method (18, 19). The amylose contents ranged from 10.2 to 26.1%, including high (*Goamibyeo*, 26.1%) and low (*Baekjinju*, 10.2%) amylose varieties. Excluding these two varieties, the amylose contents ranged from 16.8 to 21.3%, similar to previous reports (6–8).

Development of calibration models for amylose determination Regression analysis was conducted by correlating the raw spectra obtained from the tested rice samples (Fig. 1) within the 1,000 to 2,500 nm range, and the amylose contents determined by the colorimetric procedure (18, 19). Although the raw spectra had many bands with overtone and combination due to certain components and the wavelength region (25), the coefficient of determination (R^2) was 0.74, which was higher than

Table 1. Amylose contents of the 60 Korean polished rice varieties used in the study¹⁾

Variety	Amylose content (%)	Variety	Amylose content (%)	Variety	Amylose content (%)
Baekjinju	10.2±0.01 ¹⁾	Jongnambyeo	18.7±0.12	Sobibyeo	18.8±0.17
Chucheongbyeo	19.2±0.23	Joonambyeo	19.4±0.51	Soojinbyeo	19.7±0.57
Daepyeongbyeo	18.7±0.58	Joonghwabyeo	17.6±0.20	Soorabyeo	18.2±0.35
Damagum	19.9±0.65	Mangumbyeo	19.3±0.40	Taebongbyeo	19.1±0.38
Dongganbyeo	19.2±0.61	Manpoongbyeo	19.1±0.20	Yangjobyeo	20.9±0.06
Dongjinbyeo	19.9±0.23	Manweolbyeo	19.2±0.32	Younganbyeo	18.7±0.53
Goamibyeo	26.1±0.12	Mihyangbyeo	19.9±0.30	Younghaebyeo	18.7±0.20
Gumnambyeo	20.5±0.35	Namgangbyeo	19.0±0.32	Chucheongbyeo I*	20.9±0.60
Haepyeongbyeo	19.6±0.75	Nampyeongbyeo	19.8±0.52	Chucheongbyeo II*	21.3±0.76
Hoanbyeo	21.1±0.15	Nonghobyeo	21.2±0.59	Chucheongbyeo III*	16.8±0.40
Hojinbyeo	19.2±0.35	Odaebyeo	19.6±0.45	Ilmibyeo I*	19.8±0.46
Hwaanbyeo	19.0±0.35	Saechucheongbyeo	20.1±0.31	Ilmibyeo II*	19.2±0.36
Hwabongbyeo	19.5±0.06	Saegyechwabyeo	19.1±0.21	Ilpumbyeo*	19.1±0.55
Hwaseongbyeo	19.4±0.21	Saesangjubyeo	17.7±0.60	Kumseongbyeo*	20.1±0.42
Hwashinbyeo	19.6±0.06	Samcheonbyeo	18.4±0.25	Nampyeongbyeo*	18.3±0.61
Hwayoungbyeo	18.0±0.32	Sangmibyeo	18.1±0.25	Nampyeong+Ilmibyeo*	19.8±0.56
Hyangnambyeo	19.6±0.40	Seoganbyeo	19.6±0.21	Odaebyeo*	18.1±0.40
Ilmibyeo	20.3±0.45	Seokjeongbyeo	19.3±0.49	Sambakilpumbyeo*	18.5±0.64
Ilpumbyeo	20.3±0.45	Seolgaeng	18.4±0.26	Samcheonbyeo*	19.3±0.12
Inweolbyeo	20.0±0.38	Sindongjinbyeo	18.9±0.45	Sangjuilpumbyeo*	19.1±0.61

¹⁾Mean±SD (n=4); *Varieties of polished rice as a regional brand from different locations.

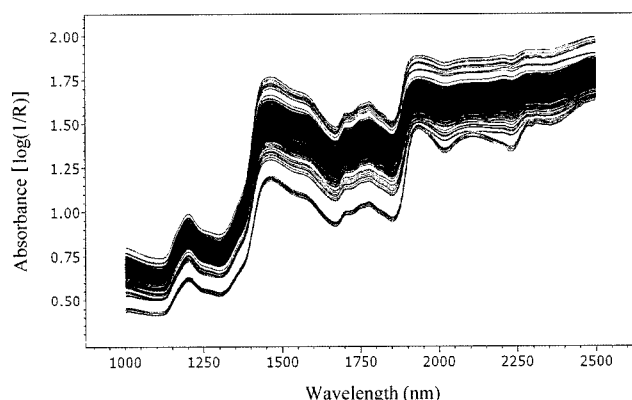


Fig. 1. Mean Fourier transform near-infrared (FT-NIR) spectra for Korean rice varieties (n=600). Measured 3 times and expressed as a mean spectrum.

expected. This is probably because FT-NIR spectroscopy usually yields a higher signal to noise ratio than does NIR spectroscopy due to its higher energy light source. Despite the relatively high R^2 , when used to predict the chemical component values, the model based on the raw spectra was not accurate enough as compared to commercial real-time analysis applications (data not shown). Therefore, it was necessary for the raw spectra to be mathematically treated, including stabilizing the baseline and removing the

Table 2. Results of the partial least squares regression analysis obtained with different mathematical pre-treatments of FT-NIR raw spectra for the prediction of amylose content of polished rice¹⁾

Pre-treatment	R^2	SEC	SEP
Raw	0.74	2.22	2.98
NOR	0.79	1.57	1.75
MSC	0.77	1.85	2.58
1st Derivative	0.81	0.77	1.08
NOR+1st derivative	0.86	0.36	0.38
1st Derivative+NOR	0.85	0.41	0.42
NOR+1st derivative+MSC	0.92	0.21	0.24
NOR+1st derivative+SMT+MSC	0.96	0.18	0.20
1st Derivative+NOR+SMT+MSC	0.88	0.25	0.29
1st Derivative+MSC+NOR+SMT	0.88	0.26	0.31

¹⁾ R^2 , coefficient of determination; SEC, standard error of calibration; SEP, standard error of prediction; raw, NIR raw spectra (not treated); NOR, normalization; MSC, multiplicative scattering correction; SMT, smoothing.

noise, to improve the accuracy.

The raw spectra were pretreated in various ways to obtain a model with the highest R^2 and the lowest SEP (Table 2). The raw spectra were first subjected to

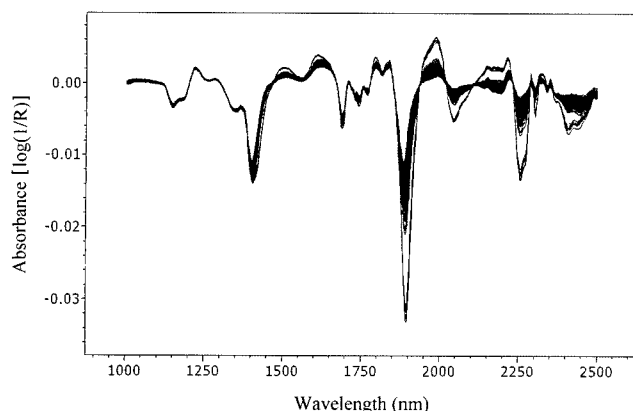


Fig. 2. Pre-treated* FT-NIR spectra for Korean rice varieties (n=460). *normalization, 1st derivative, smoothing, and multiplicative scattering correction.

normalization (NOR) treatments for baseline stability (27). After these treatments, the R^2 was slightly improved from 0.74 to 0.79, but the SEP (=1.75) was still high (Table 2).

The MSC method provided an R^2 similar to that obtained with the NOR pre-treatment, but the SEC and SEP were enhanced. The first derivative method, which can reduce baseline effects and eliminate a linear ordinate offset, increased R^2 (0.81) and decreased both SEC and SEP (0.77 and 1.08, respectively).

A combination of NOR and the first derivative method provided a lower SEP (0.38) than that obtained when the NOR and the first derivative were separately used. As shown in Table 2, it was also found that the pretreatment order affected the regression. In other treatments such as MSC and smoothing (SMT), which are usually used to reduce noise levels in spectra, the combination of all four methods yielded the best correlations between amylose and the spectrum data, showing the lowest SEP of 0.20 and the highest R^2 of 0.96.

Based on the above results, the combination of the four pretreatments (NOR, 1st derivative, SMT, and MSC) was selected as the most promising dataset for the PLS regression to predict the amylose content of the polished rice. The transformed FT-NIR spectra are shown in Fig. 2.

Table 3 shows the results of the regression analysis between the pre-treated spectra data (Fig. 2) and chemical data to predict the amylose content of the polished rice. X-PRESS and PRESS determined the optimum numbers of the primary and secondary factors, respectively. Primary was 15, and secondary was 1-15 within the primary factor numbers. In determining the factor numbers, over-fitting was prevented through the change of SEP as the factor number increased. For PLS regression, the calibration and validation of amylose were calculated using the pretreated spectra above. The model equation provided acceptable results for predicting the amylose contents of the polished rice, showing an R^2 of > 0.94, a SEC of 0.17, and an SEP of 0.20. The quality value, an index of overall quality for the predicted model, was stable (0.9676) and the bias (i.e., the difference between the measured and predicted values) was close to zero.

Table 3. Calibration and validation results of partial least squares regression analysis obtained using the pre-treated FT-NIR data for the prediction of amylose contents of polished rice¹⁾.

Parameter	Values for amylose content	
Factors	Primary	15
	Secondary	1-15
Calibration	n	308
	R^2	0.96
Validation	SEC	0.1706
	n	152
Validation	R^2	0.94
	SEP	0.1961
Validation	Bias	0.0383
	Q-value	0.9676

¹⁾n, number of points used for calibration or validation; R^2 , coefficient of determination; SEC, standard error of calibration; SEP, standard error of prediction; Bias, difference between reference and NIR values (predicted values); Q-value, quality value.

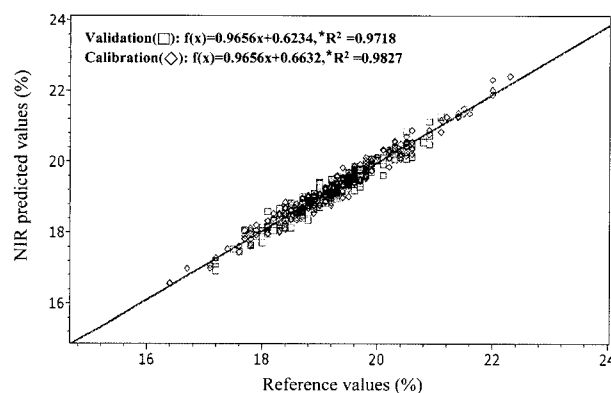


Fig. 3. Relationship between amylose contents determined by the developed model and by the standard method (n=460. calibration set=308, validation set=152, SEP=0.1961, * R^2 =coefficient of determination).

Figure 3 shows the regression relationship between the amylose contents of the polished rice that were determined by the spectroscopy technique and by the standard method for the calibration and validation datasets. The points on the calibration and validation were highly concentrated within the inner analysis range. Highly significant relationships ($R^2 > 0.94$) were found between the amylose contents obtained with the developed model and the standard instrument. Since the regression slopes were close to one and there were only small y-intercepts of 1%, it is reasonable to assume there were little differences in the amylose quantities estimated by the two methods.

References

- Chi HY, Lee CH, Kim KH, Kim SL, Chung IM. Induction of apoptotic cell death by red pericarp rice (*Jakwangchalbyeol*) extracts. Food Sci. Biotechnol. 15: 534-542 (2006)

2. Kim SS, Kim DC. Quality of rice marketed in Korea. *Food Sci. Biotechnol.* 7: 255-228 (1998)
3. Chae JC, Jung MS, Jun DK, Son YM. Relationship between yield and quality of rice varieties grown in reclaimed saline paddy field. *Korean J. Crop Sci.* 47: 259-262 (2002)
4. Son JR, Kim JH, Lee, JI, Youn YH, Kim JK, Hwang HG, Moon HP. Trend and further research of rice quality evaluation. *Korean J. Crop Sci.* 47: 33-54 (2002)
5. Chae JC. Present situation, research, and prospect of rice quality and bioactivity in Korea. *Food Sci. Ind.* 37: 47-54 (2004)
6. Moon SS, Lee KH, Cho RK. Application of near-infrared reflectance spectroscopy in quality evaluation of domestic rice. *Korean J. Food Sci. Technol.* 26: 718-725 (1994)
7. Hwang HG, Cho RK, Sohn JK, Lee SK. Rapid evaluation of chemical components of rice grain using near infrared spectroscopy. *Korean J. Crop Sci.* 39: 7-14 (1994)
8. Delwiche SR, Mckenzie KS, Webb BD. Quality characteristics in rice by near-infrared reflectance analysis of whole-grain samples. *Cereal Chem.* 73: 257-263 (1996)
9. Kim YS, Scotter C, Voyiagis M, Hall M. Potential of NIR spectroscopy of discriminating geographical origin of green tea from Korea and Japan. *Food Sci. Biotechnol.* 6: 74-78 (1997)
10. Kim SK, Chae JC, Lim MS, Ree JH. Interrelationship between amylose content and physical properties of milled rice. *Korean J. Crop Sci.* 30: 320-325 (1985)
11. Kim KH, Kim DH, Cho YH. Development of near isogenic lines and genetic analysis for amylose content of rice grain. *Korean J. Breed.* 33: 280-286 (2001)
12. Lee SH, Han O, Lee HY, Kim SS, Chung DH. Physicochemical properties of rice starch by amylose content. *Korean J. Food Sci. Technol.* 21: 766-771 (1989)
13. Czuchajowska Z, Szczodrak J, Pomeranz Y. Characterization and estimation of barley polysaccharides by near-infrared spectroscopy. I. Barleys, starches, and β -D-glucans. *Cereal Chem.* 69: 413-418 (1992)
14. Osborne BG, Fearn T. Theory of near-infrared spectrophotometry. pp. 13-35. In: *Near Infrared Spectroscopy in Food Analysis*. Osborne BG, Fearn T (eds). Longman Scientific and Technical, Essex, UK (1986)
15. McClure WF, Maeda H, Dong J, Liu Y, Ozaki Y. Two-dimensional correlation of Fourier transform near-infrared and Fourier transform Raman spectra. I: Mixtures of sugar and protein. *Appl. Spectrosc.* 50: 467-475 (1996)
16. Chung HI, Kim HJ. Near-infrared spectroscopy: principles. *Anal. Sci. Technol.* 13: 1A-14A (2000)
17. Miyamoto K. Classification of high acid fruits by PLS using the near infrared transmittance spectra of intact satsuma mandarins. *J. Near Infrared Spec.* 6: 267-271 (1998)
18. Juliano BO. Criteria and tests for rice grain qualities. *Rice Chem. Anal. Technol.* 55: 443-524 (1985)
19. Juliano BO. A simplified assay for milled-rice amylose. *Cereal Sci. Today* 16: 334-338 (1971)
20. Kim SK, Jung SO. Physicochemical properties of Japonica non-waxy and waxy rice during kernel development. *Food Sci. Biotechnol.* 15: 289-297 (2006)
21. Reeves JB. Discriminant analysis of selected food ingredients by near infrared diffuse reflectance spectroscopy. *J. Near Infrared Spec.* 5: 209-221 (1997)
22. ASTM. American Society of Testing and Materials Annual Book of Standards. Vol. 03.06, Doc. E1655-94. Standard practices for infrared, multi-variate, quantitative analysis. American Society of Testing and Materials, West Conshohochon, PA, USA (1995)
23. Cho HJ, Ha YL. Determination of honey quality by near infrared spectroscopy. *Korean J. Food Sci. Technol.* 34: 356-360 (2002)
24. Kim JM, Choi CH, Min BK, Kim JH. Development of prediction model for moisture and protein content of single kernel rice using spectroscopy. *J. Korean Soc. Agric. Machinery* 23: 49-56 (1998)
25. Lee SM, Kim S, Park JB, Hwang IK. Prediction of chemical compositions for on-line quality measurement of red pepper powder using near infrared reflectance spectroscopy (NIRS). *Food Sci. Biotechnol.* 14: 280-285 (2005)
26. Sohn MR, Cho RK. Possibility of nondestructive evaluation of pectin in apple fruit using near-infrared reflectance spectroscopy. *J. Korean Soc. Hort. Sci.* 41: 65-70 (2000)
27. Norris KH, Williams PC. Optimization of mathematical treatments of raw near-infrared signal in the measurement of protein in hard red spring wheat. I. Influence of particle size. *Cereal Chem.* 61: 158-165 (1984)