JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Analysis of Fiber Content in Flax Stems by Near-Infrared Spectroscopy

F. E. BARTON, II,*,† D. E. AKIN,† W. H. MORRISON,† A. ULRICH,‡ AND D. D. ARCHIBALD§

Richard B. Russell Agricultural Research Center, USDA-Agricultural Research Service, P.O. Box 5677, Athens, Georgia 30604-567, Biolin Research, Inc., 103−110 Research Drive, Saskatoon, Canada S7N 3R13, and Department of Crop and Soil Science, Pennsylvania State University, University Park, Pennsylvania 16802

The conventional means of measuring the fiber content of flax is time-consuming and laborious, and the results obtained vary with the analysis technique used. The plant tissues must first be "retted", a process by which the fibers are separated from the rest of the stem, either by indigenous organisms in the soil when the stems are left in the field or by water (anerobic bacteria) or enzymatic retting. The fiber content is then determined by mechanical or manual separation. In this study, fiber content of flax stems was measured rapidly and objectively by near-infrared spectroscopy (NIRS) using whole pieces of stem in a large cell, in reflectance mode. Compared to the conventional method, the standard error of performance of the NIRS method was between 0.96 and 1.45% (dry matter basis), depending on the model and data processing used. NIRS calibrations were generated by hand separation of fiber from water-retted specimens. The water retting procedure takes several days to complete and requires considerable trained labor to complete the hand separation step. The NIRS procedure was conducted on pieces of stem to simulate measurement in the field.

KEYWORDS: Flax; fiber; near-infrared; NIRS; chemometrics

INTRODUCTION

Flax (*Linum usitatissimum*) has been cultivated for thousands of years and is an important crop in many regions of the world. Flax is gaining in importance in the United States for linen and linen/cotton blend textile products. However, in the United States, the largest per capita consumer of flax fiber in the world, all textile-grade flax fiber and yarn for incorporation into textiles is imported. The fiber from flax is used in textiles, such as linen and linen cotton blends; in composites as a substitute for glass fiber; and in very fine paper for numerous applications. In general, cultivars which are grown for seed (linseed) in the United States are different from those grown in Europe to produce fiber. In Canada, straw remaining from the harvested seed flax has little value and is usually burned. Environmental concerns make it imperative to find ways to utilize the fiber from seed flax as well as the seed. The process of obtaining the fiber from the plant is part microbial/chemical and part mechanical. The microbial/chemical part is called "retting". The retting process depends on chemical breakdown of pectic materials between the fibers and around fiber bundles by the action of indigenous soil organisms during "dew retting", where the straw remains in contact with the ground for a period of some weeks. Alternatively, in water retting, anaerobic bacteria act on the same pectic material to effect separation of fiber from the nonfiber material (*1*−*3*). The retted straw is then cleaned by a mechanical process which removes the fiber from the cuticle and shive or core part of the stem. The amount and quality of fiber depends on the flax type (fiber or seed), cultivar, year, and retting conditions. With the substantial variation in fiber content among fiber and seed varieties, which is compounded by differences in climates, a reliable method of determining fiber content of flax stems in the field before processing would be of great benefit to growers and processors alike.

Near-infrared spectroscopy (NIRS) has been used to measure fiber in forages for many years and is an AOAC International Official Method for determining acid detergent fiber in grasses (*4*, *5*). The determination of degree of retting has been measured by NIRS and mid-infrared (MIR) (*6*). In this case, Kessler and Kohler used the absorbances for pectin (CH═O at 1740 cm^{-1}) and lignin (AR-ring stretch at 1515 cm^{-1}) as an indication of retting in the MIR and the ratio of the absorbance at 1000 nm to the absorbance at 1370 nm in the NIR. Another study, by Faughey and Sharma (*7*), with NIRS modeled fiber properties of retted straw. There has been no attempt to measure fiber content of intact straw. However, a model for the degree of retting, i.e., the Fried test scores (*8*, *9*) has been developed with NIRS (*10*). The composition of flax and the changes occurring

† USDA-Agricultural Research Service.
‡ Biolin Research, Inc.
§ Pennsylvania State University.

10.1021/jf020367i CCC: $22.00
Published on Web 11/06/2002

during retting have been reported (*11*). These authors utilized nuclear magnetic resonance and Raman spectroscopy, GC−mass spectrometry, and light microscopy to analyze the separate botanical parts of the flax plant and to observe what happens, chemically and structurally, when flax is enzymatically retted. Archibald et. al. (*12*) and Himmelsbach et al. (*13*, *14*) utilized MIR and Raman spectroscopy and MIR, Raman, and NMR imaging techniques, respectively, to characterize the components of flax and determine what constituents are changed during the retting process. These authors were able to identify the pectin layers and observe the removal of pectin during retting to free the fibers. All of these chemical and spectral data suggest that the fiber content should be measurable in intact flax stems.

The main objective of the present study was to determine the extent to which fiber content of flax stems, using intact stem pieces, could be measured by NIRS. In many instances, seed flax contains sufficient fiber to justify processing for textiles and high-end composites. The straw from these plants is traditionally burned in the field. There were two other objectives in the study that were to be tested when it was determined that a reasonable model for fiber in intact flax stems could be developed. The first was to determine how sophisticated an NIR instrument was needed to make the measurements. Was a full-scanning monochromator required, or could one of the less expensive models, such as an InGaAs diode array spectrometer, work? The importance of this was established by Barton et al. (*15*) when they examined the best optical geometries to measure the compositional properties of rice. This study showed that the accuracy of the reference method could greatly influence the optimal optical geometry. The second task was to test the efficacy of some of the more modern chemometric techniques, such as Marten's regression and artificial neural networks, to see if they offered improved results over the standard PLS1 algorithms. In this paper, we present a model which can be used to measure the fiber and conversely the shive content of flax straw.

MATERIALS AND METHODS

Selection of Samples. A set of samples of flax stems cut to 19.1 cm lengths was provided from 200 selections from the flax germplasm collection, Saskatoon Research Centre, Agriculture and Agri-food Canada, by Biolin Research, Inc. These 200 selections were chosen from some 1700 cultivars grown in the University of Saskatchewan flax nursery in Saskatoon, Saskatchewan, Canada, over a three-year period (1999−2001). One sample had a redundant label and was eliminated to preclude having a duplicated sample in the calibration set. A second set of 34 samples was acquired from the total 1700 sample population for validation from the 2000 collections.

Near-Infrared Spectroscopy. Visible/near-infrared spectra were obtained on a scanning monochromator (model 6500, Foss NIRSystems, Silver Spring, MD) in reflectance mode over a wavelength range of 400−2498 nm. The instrument was equipped with a sample transport device and a large sample cup (21 cm long × 5 cm wide × 4 cm diameter) used. The flax stems were arranged in the cup in several layers, filling the cup. The amount of flax stems was approximately 50 g. After the sample was scanned, the stems were rearranged and the spectra retaken. The samples were rearranged a total of five times each and the spectra averaged at the end. The instrument was operated by the software package WINISI version 2.01 (Infrasoft International, Inc., Port Matilda, PA), which includes modules for acquisition and processing of spectra.

Chemometrics. Calibrations were developed using the modified PLS1 and artificial neural network programs in the WINISI software and the PLS1 and Marten's regression programs in the Unscrambler software package, version 7.6 (CAMO, Trondheim, Norway). The modification to the PLS1 algorithm was to scale the reference method data and spectral data at each wavelength to have a standard deviation of 1.0 before each PLS1 term (*16*). This typically results in fewer PLS1 terms in the model and fewer *T* outliers. The global *H* (Mahalanobis distance) for outlier detection of spectra was set at 2.5 and *T* for reference data at 2.5 as well.

Water Retting Procedure. A weighed bundle of straw, 19.1 cm in length by 4.1 cm in diameter, was formed from each cultivar of the samples collected. Batches of up to 150 bundles were retted in a warm (50 °C) water tank for 4−6 days until the results of the Fried test (*8*, *9*) showed that the majority of the bundles were optimally retted. The Fried test involved placing 15 pieces of retting straw, each approximately 10.0 cm in length, each from a different bundle, in individual test tubes half full of boiling water. The stoppered tubes were placed in a machine designed in the Biolin Laboratory to violently shake the tubes for 15 s. After shaking, the straw in each tube was visually scored for loose fibers on an integer scale from 0 to 3, with 0 representing no loose fibers and 3 representing total loosening of all fibers on the stem. When the average score was above 2.8, there were no 0 scores, and at least 12 test tubes had a score of 3, the bundles in the batch were considered to be fully retted. The bundles were then taken from of the tank, rinsed in tap water, and set in drying racks under a fume hood for approximately 4 days to dry at room temperature.

The dried bundles were weighed, and the fiber was extracted with a reciprocating blade-type breaker/decorticator. The fiber was hand-cleaned to the point where no more shives could be seen on the fiber or no fiber could be found mixed with the shives. After cleaning, the fiber and shives were weighed and placed in separate sample bags and labeled with the accession number and a sub-catalog letter to distinguish fiber and shive subsamples. After all bundles were processed, a master spreadsheet was compiled with accession sample number, fiber weight, shive weight, dust weight, and moisture. Fiber and shive contents were determined by weight of fiber or shive extracted divided by weight of straw before retting. These data were added to the spreadsheet, imported into the NIRS spectral data file, and used to develop the calibration models. For this study, only the fiber content data were used.

Table 1. Percent Fiber Reference Data on Flax Data Set

data set	number	minimum	maximum	mean	SD
calibration	200	15.6	35.0	23.5	3.5
validation	34	17.2	34.1	22.9	3.7

RESULTS AND DISCUSSION

The flax samples in this study were selected from the increased plantings from a collection of germplasm of the Saskatoon Research Centre, Agriculture and Agri-food Canada. While most of the varieties were Canadian or European in origin, roughly 25% came from flax oilseed varieties in the United States. The varieties were grown over a three-year period, and the samples used for calibration were from years one and two, validation from year three. The samples were selected to cover a broad range of fiber content. This can be seen in the data (**Table 1**). The varieties with the lower fiber content were predominantly seed flax cultivars, while those with the high fiber content were fiber flax varieties. **Table 1** includes the data for the validation set of 34 samples. The range and standard deviation of this set are very close to those of the large set, indicating it should be a representative validation set.

Figures 1 and **2** summarize the results of this study graphically. **Figure 1** is a scatter plot of predicted vs reference for percent fiber in flax stems. The model for this plot is the Marten's regression in **Table 4**, which uses six PLS1 factors. All of the models in **Tables 2−4** use only 6−8 PLS1 factors.The American Society for Testing and Materials (ASTM) has published a Standard (E 1655-94) (*17*) covering the protocols and procedures for establishing a spectroscopic chemometric method. The standard covers all the usual linear model develop-

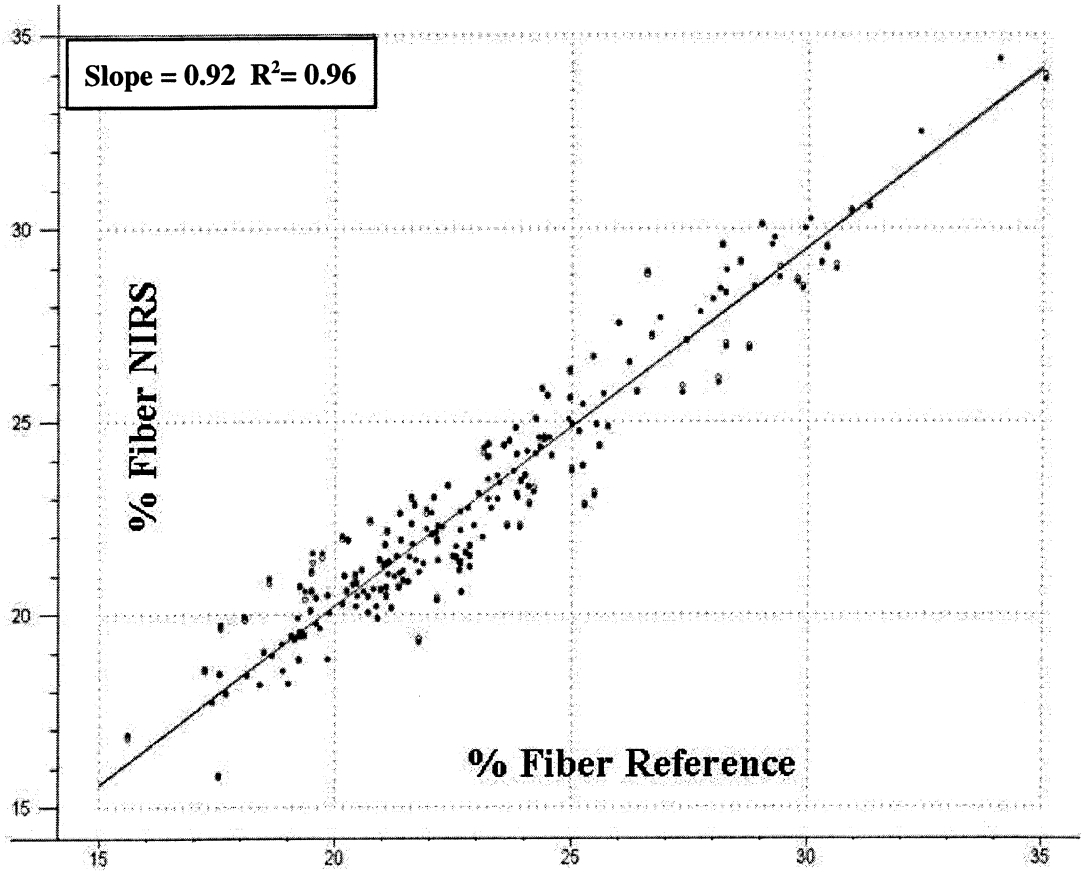

Figure 1. Scatter plot of percent fiber from water retting reference data vs percent fiber by NIRS using the Marten's regression equation.

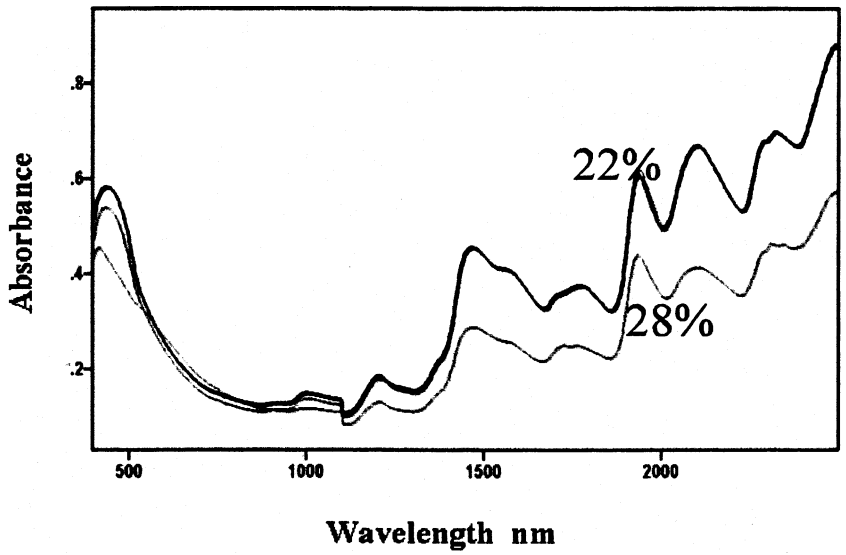

Figure 2. Near-infrared spectra of moderate fiber content.

Table 2. Standard Error of Calibration (RMSEC) and Cross Validation (RSEMP) for PLS1 Models Developed with the Unscrambler Chemometrics

model	RMSEC	R^2	RMSEP	R^2
400–2500 nm	1.31	0.91	1.40	0.90
1100–1700 nm	1.34	0.90	1.45	0.88
1100–2500 nm	1.56	0.91	1.44	0.89

Table 3. Standard Error of Calibration (SEC) and Cross Validation (SECv) for PLS1 Models Developed with Infrasoft International NIRS3 Chemometrics

model	SEC	R^2	SECv	R^2
400–2500 nm	1.23	0.84	1.40	0.84
1100–1700 nm	1.27	0.84	1.38	0.84
1100–2500 nm	1.18	0.85	1.36	0.85

ment methods and establishes guidelines for the judgment of a model's validity (Annual Book of ASTM Standards, 1994.) In the case of the models developed for this study, 6−8 PLS1 factors is very conservative. **Figure 2** is an overlaid plot of flax NIR spectra of three samples, two with moderate (22%) and one with high (28%) fiber contents.

Table 4. Standard Errors of Calibration for Advanced Chemometric Methods

model	RMSEP/ SEP	R^2	RMSEP/ SECv	R^2
Marten's regression	0.93	0.96	0.96	0.96
artificial neural network	1.29	0.88	1.31	0.88

The results in **Tables 2** and **3** contain the performance parameters for models developed with Unscrambler and ISI software packages, respectively, and show that the fiber content can be measured on intact flax stems with acceptable accuracy. The standard errors of calibration and cross validation/ performance (SEC and SECv, respectively) are comparable to those found by Barton and Windham (*5*) for fiber in forages that was used to establish NIRS as an Official Method for AOAC International. In that study, the SEP for acid detergent fiber (ADF) in forages was 1.14, with an R^2 of 0.94. The ADF procedure is a more robust and precise laboratory measurement than fiber from water retting. The principal difference between the two chemometric packages lies in the coefficient of determination (R^2). The Unscrambler software identifies more sample outliers in the calibration phase (Mahalanobis distance graphically plotted, 2.5 used as a criteria) than the ISI package; thus, more outliers are eliminated, and the R^2 for SEC is higher. The ISI software appears to be far more reluctant (in this case the global *H* was set to the default value of 2.5) to identify a sample as an outlier during calibration. The ISI package would only identify 5−7 samples, whereas Unscrambler identified 15−18. In all cases, the outliers were reference data (*T*) outliers; no spectral outliers were found. This difference is not surprising, due to the source of the reference data and the fact that the modification of the PLS1 algorithm in WINISI scales the reference data at each wavelength to a standard deviation of 1.0. Any gravimetric procedure is likely to contain several places in the method for systematic errors to occur. Usually these are due to weighing techniques, but in this case there is a systematic error which leads to a bias that would cause the reference method to consistently yield low values. When a stem is removed from a bundle to be used in the Fried test, there is no way to account for the loss of fiber within the total sample. Therefore, the measured bundle's fiber content will be reduced by that in the stem. If the same bundle is sampled twice during the retting process, the error is compounded. This error may be only 1−2% of the beginning dry weight of flax but could result in a 7−10% coefficient of variation in the fiber measurement.

There is no real difference between these models, regardless of software or instrument configuration. Thus, either a full VIS/ NIR (400−2500 nm), NIR (1100−2500 nm), or a diode array instrument (1100−1700 nm) would be suitable for the measurements. A second set of samples (34) was used as a validation set, and an SEP of 2.23 was obtained. When one is trying to decide if the flax is suitable for processing to obtain fiber, a value of 14.0 ± 2.23% would be sufficient. High-fiber textile-grade flax would be around 25−30% fiber. The bias on these samples was a little over 10% of the SEP and negative (−0.29). This result could be explained by spectral differences or the bias in the reference method discussed above.

An improvement in the SEP can be obtained by developing the models with either ISI's artificial neural network (ANN) algorithm or Unscrambler's Marten's regression (**Table 4**). The ANN result is somewhat improved over the PLS1 (SECv of 1.31) models in **Tables 2** and **3**. The ISI default options were used to develop the ANN model and were not optimized. The Marten's regression result is a substantial improvement, with an SECv of 0.96 and an R^2 of 0.96. Marten's regression leaves out any spectral data point which could hurt the model, thus eliminating those spectral parameters which only contribute noise. The resulting model is a PLS1 over only those spectral regions that contribute real information. In this example, 175 spectral data points out of 1039 were eliminated, leaving four segments. The resulting segments represented the portions of the spectra which measure C−H, C−O−H, and O−H stretch. This result can be approximated in WINISI by defining the spectral segments to be used in the model and generating the PLS1 model on a reduced spectral data set. If one continues to successively eliminate data points from the spectra until only those remain that have the highest correlation, the result would be similar to multiple linear regression (MLR) wavelength selection routines. An automated approach to select spectral windows was developed by Archibald and Akin (*10*). This procedure could be used to define those spectral elements essential to the measurement and, in essence, design a filter instrument capable of performing the measurement. In their study, Archibald and Akin found that only five narrow regions of the NIR spectrum were needed to measure the Fried test score (0−3) with an SECv of 0.3. This level of precision would equate to an SECv of 1.4 for samples containing 14% fiber. The range of fiber in this study from Table 1 was 15.6 to 35.0%. Thus it is reasonable to say that a simple filter instrument could be designed to measure the fiber content of flax straw.

CONCLUSION

The measurement of fiber content in flax straw stems can be accomplished with NIRS with an SEP of 2.2%. This level of precision is sufficient to allow fiber processors to determine the value of the standing fiber crop, the extent of processing needed to recover the fiber, and the suitability of the fiber for various uses. It has also been shown that the instrument type needed for the measurement covers the span of what is available in the marketplace. Further research will be conducted to transfer the calibrations to other instruments and to take the fiber measurement from the laboratory to the field.

ACKNOWLEDGMENT

The authors gratefully acknowledge the technical assistance of Ms. Judy Davis and Mr. Randy Cowan for the acquisition of the spectral data and laboratory retting of the flax.

LITERATURE CITED

(1) Allen, L. A. Recent developments in the microbiology of retting. Part II. The part played by bacteria in retting and purifying the liquor. *J. Soc. Chem. Ind*. **1946**, *65*, 373−377.
(2) Chesson, A. The maceration of linen flax under anaerobic conditions. *J. Appl. Bacteriol.* **1978**, *45*, 219−230.
(3) Sharma, H. S. S. Studies on Chemical and Enzyme Retting of Flax on a Semi-Industrial Scale and Analysis of the Effluents for their Physico-Chemical Components. *Int. Biodeterior*. **1987**, *23*, 329−342.
(4) Marten, G. C., Shenk, J. S., Barton, F. E., II, Eds. *NIRS analysis of forage quality*; USDA Handbook No. 643; U.S. Government Printing Office: Washington, D.C., 1985.
(5) Barton, F. E., II; Windham, W. R. Determination of acid detergent fiber and crude protein in forages by near infrared spectroscopy: Collaborative study. *J. Assoc. Off. Anal. Chem.* **1988**, *71*, 1162−1167.
(6) Kessler, R. W.; Kohler, R. New strategies for exploiting flax and hemp. *Chemtech* **1990**, December, 31−42.

(7) Faughey, G. J.; Sharma, H. S. S. A preliminary evaluation of near infrared spectroscopy for assessing physical and chemical properties of flax fiber. *J. Near Infrared Spectrosc.* **2000**, *8*, 61−69.
(8) Van Sumere, C. F. Retting of Flax with Special Reference to Enzyme-retting. In *The Biology and Processing of Flax*; Sharma, S. H. H., Van Sumere, C. F., Eds.; M Publications: Belfast, Northern Ireland, 1992; pp 157−198.
(9) Henriksson, G.; Akin, D. E.; Slomczynski, D.; Eriksson, K.-E. L. Production of Highly Efficient Enzymes for Flax Retting by Rhizomucor pusillus. *J. Biotechnol.* **1999**, *68*, 115−123.
(10) Archibald, D. D.; Akin, D. E. Use of Spectral Window Processing for selecting Near Infrared Reflectance Wavelengths for Determination of the degree of Enzymatic Retting of intact Flax Stems. *Vib. Spectrosc.* **2000**, *23*, 169−180.
(11) Morrison, W. H., III; Akin, D. E.; Himmelsbach, D. S.; Gamble, G. E. Chemical, Microscopic, and Instrumental Analysis of Graded Flax Fiber and Yarn. *J. Sci. Food Agric.* **1999**, *79*, 3−10.
(12) Archibald, D. D.; Henrikssen, G.; Akin D. E.; Barton, F. E., II. Spectroscopic Characterization of Enzymatic Flax Retting: Factor Analysis of FT-IR and FT-Raman Data. *Proceedings of the 11th International Conference on Fourier Transform Spectroscopy*; American Institute of Physics: Woodbury, NY, 1998; pp 653−656.
(13) Himmelsbach, D. S.; Khalili, S.; Gamble, G. R.; Akin D. E. Mid-Infrared, Raman and NMR Imaging of Flax. *Microsc. Microanal.* **1998**, *4* (Suppl. 2), 844−845.
(14) Himmelsbach, D. S.; Khalili, S.; Akin D. E. NIR-FT-Raman Microspectroscopic Imaging of Flax Stems. *Vib. Spectrosc.* **1999**, *19*, 361−367.
(15) Barton, F. E., II; Windham, W. R.; Champagne, E. T.; Lyon, B. G. Optimal Geometries For the Development of Rice Quality Spectroscopic Models. *Cereal Chem.* **1998**, *75*, 315−319.
(16) Shenk, J. S.; Westerhaus, M. O. Population, Definition, asample Selectionand Calibration Procedures For Near Infrared Reflectance Spectroscopy. *Crop Sci.* **1991**, 469−474.
(17) ASTM Designation E 1655-94. *Annual Book of ASTM Standards*; American Society for Testing and Materials: West Conshocken, PA, 1994.

Received for review March 26, 2002. Revised manuscript received August 2, 2002. Accepted August 5, 2002.

JF020367I