JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Near-Infrared Analysis of Soluble and Insoluble Dietary Fiber Fractions of Cereal Food Products

SANDRA E. KAYS* AND FRANKLIN E. BARTON, II

Richard B. Russell Agricultural Research Center, Agricultural Research Service, U.S. Department of Agriculture, P.O. Box 5677, Athens, Georgia 30604-5677

The use of near-infrared (NIR) reflectance spectroscopy for the rapid and accurate measurement of soluble and insoluble dietary fiber was explored in a diverse group of cereal products. Ground samples were analyzed for soluble and insoluble dietary fiber (AOAC Method 991.43) and scanned (NIRSystems 6500 monochromator) to obtain NIR spectra. Modified PLS models were developed to predict insoluble and soluble dietary fiber using data sets expanded to include products with high fat and high sugar contents. The models predicted insoluble dietary fiber accurately with an SECV of 1.54% and an R^2 of 0.98 (AOAC determined range of 0−48.77%) and soluble dietary fiber less accurately with an SECV of 1.15% and an R^2 of 0.82 (AOAC determined range of 0−13.84%). Prediction of independent validation samples by the soluble fiber model resulted in a bias that may be related to the way the reference method treats samples with different soluble fiber constituents. The insoluble fiber model can be used to rapidly monitor insoluble dietary fiber in cereal products for nutrition labeling.

KEYWORDS: Dietary fiber; near-infrared; NIR; cereal

INTRODUCTION

Cholesterol and glycemic index lowering properties of soluble dietary fiber have greatly stimulated interest in the soluble fiber content of cereal foods (*1*−*4*). These physiological effects of soluble fiber are associated with a decreased risk of cardiovascular disease. Several cereal products, including oat bran and barley flour, contain significant quantities of soluble dietary fiber, found predominantly as β-glucan. Other cereal products contain significant quantities of soluble hemicelluloses, such as the soluble arabinoxylans (*5*), which are likewise isolated in the soluble fiber fraction. Although not mandatory, the amount of soluble fiber present in a product can be included on the nutrition label in the United States as a subheading of total dietary fiber (*6*), an option often utilized by manufacturers of oat-based products.

Insoluble dietary fiber has beneficial effects in human physiology and is required for normal digestive function (*7*). Insoluble fiber is the predominant fiber fraction in most cereal products and consists of insoluble hemicelluloses, cellulose, resistant starch, and lignin. As with soluble dietary fiber, the amount of insoluble fiber may be listed on the nutrition label in the United States and may be used in a method for the determination of caloric content of foods, where insoluble fiber is subtracted from the total carbohydrate content before the calculation of calories contributed by the carbohydrate portion (*6*). The AOAC enzymatic−gravimetric method, used for the determination of soluble and insoluble dietary fiber, is very time-consuming, taking 2 days to perform an assay. The assay requires an additional day if sugar or fat extraction is required.

Near-infrared (NIR) spectroscopy is a rapid, accurate technique that is used for the analysis of protein, oil, and moisture in cereal grains (*8*) and for the analysis of certain components of foods (*9*). Few studies have addressed the potential of NIR spectroscopy for the determination of soluble dietary fiber. Horvath (*10*) reported prediction of soluble fiber and pectin in fabricated mixtures of whole wheat flour, wheat bran components, pectin, and lignin. Williams et al. (*11*) reported analysis of soluble fiber in oat bran products. These studies involved the NIR determination of soluble fiber in homogeneous populations. Most commercially available products are derived from a wide range of grain types, often with multiple grain types in a single product. NIR analysis is further complicated by the fact that a number of cereal products contain large amounts of crystalline sugar and/or fat, both of which have unique characteristics in the NIR region of the spectrum. They, likewise, may contain additives such as nuts, dried fruits, herbs, honey, cinnamon, and cocoa and are processed by a variety of methods. It would be highly advantageous to be able to predict both soluble and insoluble dietary fiber in heterogeneous samples and sample sets. Baker (*12*) developed an NIR spectroscopy technique for the prediction of neutral detergent fiber in diverse, ready-to-eat breakfast cereals, and we previously described the NIR analysis of insoluble dietary fiber in a wide cross-section of cereal products with low fat and sugar contents (*13*). The following study investigates the use of NIR spectroscopy for the prediction of soluble dietary fiber in a diverse group of cereal products and expands a calibration previously described for

* Author to whom correspondence should be addressed [telephone (706) 546-3338; fax (706) 546-3607; e-mail sekays@qaru.ars.usda.gov].

10.1021/jf0116552
Published on Web 04/09/2002

insoluble dietary fiber to include products with high fat and sugar contents.

MATERIALS AND METHODS

Instrumentation. An NIRSystems 6500 monochromator was used to obtain NIR spectra, as described previously (*14*).

Reagents. Heat stable α-amylase, A 3306; protease, P3910; amyloglucosidase, A 9913 (amyloglucosidase A); acid-washed Celite, C 8656; total dietary fiber control kit, TDF-C10; MES, 2-(*N*-morpholino)-ethanesulfonic acid, M-8250; and Tris, tris(hydroxymethyl)aminomethane, T1503, were purchased from Sigma Chemical Co., St. Louis, MO. An alternative source of amyloglucosidase (amyloglucosidase B) was obtained from Megazyme International Ireland Ltd. (Bray, Ireland).

Enzyme Purity and Activity. The purity and activity of α-amylase, protease, and amyloglucosidase used in the AOAC dietary fiber procedure were monitored using the Sigma total dietary fiber assay control kit. Briefly, α-amylase and amyloglucosidase activities were monitored by measuring the recovery of corn or wheat starch, and protease was monitored by measuring the recovery of casein in the total dietary fiber assay (AOAC Method 991.43) (*15*). Contamination of amyloglucosidase by *endo*-cellulase was monitored by measuring the recovery of *β*-glucan in the AOAC procedure using two sources of amyloglucosidase (amyloglucosidase A and amyloglucosidase B).

Samples and Sample Preparation. Cereal products were obtained from retail grocery stores and included breakfast cereals, crackers, flours, brans, pastas, cereal bars, and some unprocessed whole grain products. Grains represented in the cereal food products were wheat, oats, barley, rye, rice, millet, and multiple grains with similar distributions of grain types in the calibration and validation data sets. The distribution of grain types in the calibration and validation data sets was similar to that described in previous studies (*14*, *16*, *17*).

Samples were ground with a cyclone mill (Cyclotec 1093 sample mill, Perstorp Analytical, Silver Spring, MD) to $<500\ \mu$m. High-sugar samples (>20% sugar according to the product nutrition label) were mixed with liquid nitrogen to facilitate grinding. High-fat samples (>10% fat according to the product nutrition label) were ground in a coffee mill (Braun KSM-2, Braun Inc., Lynnfield, MA).

Due to the low number of products containing high soluble dietary fiber (7 of 143 samples), samples were fabricated to obtain high soluble fiber content by mixing powdered *β*-glucan (Sigma Chemical Co.) with ground cereal products that were new to the calibration. Twelve samples were fabricated in this way with final soluble dietary fiber values ranging from 3.8 to 13.8%, with a mean of 9.10%.

Spectroscopic Analysis. All ground cereal samples were scanned with a NIRSystems 6500 monochromator, which scans the range of 400−2498 nm in 2 nm intervals. The lower wavelength region (400−1098 nm) is observed by two silicon detectors and the upper wavelength region (1100−2498 nm) by two lead sulfide detectors. High-fat and high-sugar samples were scanned prior to fat and sugar extraction. Duplicates of each sample were scanned in cylindrical sample cells (internal diameter = 38 mm, depth = 9 mm) with an optical quartz surface and cardboard backing, scanning each duplicate 16 times. The data were averaged and transformed to log 1/*R*. The duplicate scans of each sample were examined visually and averaged. The wavelength range used for the present study was 1104−2494 nm.

Reference Method for Insoluble and Soluble Dietary Fiber Analysis. Insoluble and soluble dietary fiber were measured in all cereal product samples with AOAC Method 991.43 (*15*), using the Sigma Chemical Co. total dietary fiber assay kit. This is an enzymatic method that digests starch and protein enzymatically and separates the insoluble and soluble fractions of dietary fiber on the basis of their solubility in buffer at the assay temperature. Soluble dietary fiber, once isolated, is precipitated in 78% ethyl alcohol. Before the assay, high-fat samples were defatted by extracting three times with petroleum ether (25 mL/g of sample), for 15 min for each extraction step, with stirring, and evaporated overnight at room temperature in a fume hood. High-sugar samples were extracted four times with 85% ethyl alcohol (10 mL/g of sample), for 15 min for each extraction step, with stirring, and evaporated in a vacuum oven overnight at 30 °C. The percent fat or sugar extracted for each sample was calculated on the basis of the sample weight before and after extraction. Insoluble and soluble dietary fiber values were adjusted for the percent fat or sugar extracted. Insoluble and soluble dietary fiber values were calculated on a dry weight basis. Dry matter/moisture contents of milled cereal products were determined by using the AOAC air-oven method (AOAC Method 945.14) (*18*).

A separate study was conducted to determine whether the bias obtained in the prediction of soluble fiber in the independent validation samples was due to possible contamination of the amyloglucosidase A by *endo*-cellulase with consequent underestimation of *β*-glucan and, thus, soluble fiber (*19*). Cereal samples containing oats and barley and, thus, expected to contain significant *β*-glucan were reanalyzed for soluble dietary fiber using an alternative brand of amyloglucosidase, that is, amyloglucosidase B. Forty-nine samples from the calibration data set and 13 samples from the validation data set were reanalyzed.

Multivariate Calibrations. Multivariate analysis, to relate the spectral data to the reference data, was performed with a commercial spectral analysis program (NIRS3, version 4.1, Infrasoft International, Port Matilda, PA). Modified partial least-squares (PLS) regression (*20*) was the regression method selected. The modification to PLS scaled the reference method data and reflectance data, at each wavelength, to have a standard deviation of 1.0 before each PLS regression term. A model was obtained for insoluble dietary fiber using a data set with 90 low-fat, low-sugar samples, 34 high-sugar samples, and 20 high-fat samples. $\mathrm{Log}_{10}(1/R)$ spectra were transformed with standard normal variate and detrending procedures (*21*), to remove multiplicative interferences of scatter, and then transformed with second-derivative processing (gap = 16 nm, smoothing interval = 16 nm) prior to the application of the modified PLS procedure. The optimum number of PLS factors used for the insoluble dietary fiber calibration was determined by cross-validation (*22*). During cross-validation, one-sixth of the samples at a time was temporarily removed from the calibration data set and used for prediction. Performance statistics were accumulated for each group of removed samples. The optimum number of factors for insoluble dietary fiber was that which produced a minimum in overall error between modeled and reference values (standard error of cross-validation, SECV). One high-fat sample was removed from the calibration data set as a spectral outlier (Mahalanobis distance > 3). The model was validated with an independent set of 63 cereal samples. All validation samples were purchased and scanned at a different time from the calibration samples. The set included 32 low-fat, low-sugar samples plus 15 high-sugar samples, 10 high-fat samples, and 6 samples containing both high fat and high sugar. Model performance was measured as the coefficient of determination (r^2), the standard error of performance (SEP), the slope of the linear regression of NIR predicted versus AOAC determined insoluble fiber, and the bias or average difference between measured and modeled values (*23*).

Three NIR models were developed for soluble dietary fiber using modified PLS regression. The preprocessing was the same as that used for the insoluble dietary fiber model as this gave the optimum results. Initially, a calibration was developed for soluble dietary fiber using 90 low-fat, low-sugar cereal products plus 12 samples fabricated by mixing ground cereal products with *β*-glucan, as described above (SDF model 1). The model was validated using 32 low-fat, low-sugar cereal samples. Second, the calibration data set used for SDF model 1 was expanded to include 34 high-sugar and 20 high-fat samples and a new expanded model developed (SDF model 2). One high-fat sample was a spectral outlier (Mahalanobis distance > 3) and was removed. The new model was tested with a validation set of 32 low-fat, low-sugar samples plus 15 high-sugar samples, 10 high-fat samples, and 6 samples containing both high fat and high sugar. Third, as the vast majority of samples in the expanded data set have soluble fiber values below 6% (137 of 155 samples), a separate model was developed for samples with 0−6% soluble dietary fiber (SDF model 3). The model was developed using 84 low-fat, low-sugar samples, 34 high-sugar samples, and 20 high-fat samples. One high-fat sample was removed as a spectral outlier (Mahalanobis distance > 3). Samples fabricated by mixing ground cereal products with *β*-glucan were not used in this data set.

Equivalent modified PLS models were developed for soluble dietary fiber analyzed using amyloglucosidase B. Reference results obtained

Table 1. Statistics for the Prediction of Insoluble and Soluble Dietary Fiber in Cereal Products by Near-Infrared Spectroscopy[a]

		calibration					validation						
model	method	*n*	mean %	SD %	SECV %	R^2	*n*	mean %	SD %	SEP %	r^2	bias %	slope
IDF	AOAC	143	10.05	10.92			62	9.66	10.79				
	NIR	143	10.09	10.80	1.54	0.98	62	9.61	10.64	1.13	0.99	0.05	1.01
SDF model 1	AOAC	102	3.51	3.56			30	2.53	1.66				
	NIR	102	3.06	2.96	1.11	0.87	30	3.5	1.75	1.34	0.72	–0.98	0.81
SDF model 2	AOAC	155	2.85	2.68			62	1.80	1.46				
	NIR	155	2.79	2.45	1.15	0.82	62	2.3	2.08	1.33	0.66	–0.50	0.57
SDF model 3	AOAC	137	2.05	1.25			61	1.70	1.24				
	NIR	137	2.03	1.10	0.74	0.65	61	1.97	1.37	0.84	0.67	–0.26	0.74

[a] IDF, insoluble dietary fiber–high-sugar and -fat expanded model; SDF model 1, soluble dietary fiber–low-fat, low-sugar samples; SDF model 2, soluble dietary fiber–high-sugar and -fat expanded model; SDF model 3, soluble dietary fiber–high-sugar and -fat expanded model using samples with 0–6% soluble dietary fiber; NIRS, near-infrared spectroscopy; *n*, number of samples; SD, standard deviation; SECV, standard error of cross-validation; R^2, multiple coefficient of determination; SEP, standard error of performance; r^2, coefficient of determination.

using amyloglucosidase B, for the soluble fiber assay of samples containing oat and barley, were substituted in the calibration and validation data sets. Preprocessing for the models was identical to that used for soluble dietary fiber models developed from reference data using amyloglucosidase A.

RESULTS

Reference Method Analysis of Insoluble and Soluble Dietary Fiber. A total of 206 calibration and validation samples were analyzed for insoluble dietary fiber, and 218 samples were analyzed for soluble dietary fiber. Sixty-two samples were reanalyzed for soluble fiber using an alternative amyloglucosidase source. The range of insoluble dietary fiber measured by AOAC Method 991.43 was from 0 to 48.8%, and the standard error of the laboratory determinations (*24*) was 0.41%. The range of soluble dietary fiber in the low-fat, low-sugar samples was 0−13.8%, and the standard error of the laboratory determinations was 0.42%. The range of soluble dietary fiber for the high-fat, high-sugar samples was 0−4.3%, and the standard error of the laboratory determinations was 0.29%. The overall standard error of the laboratory determinations for soluble dietary fiber was 0.38%. The standard error of the laboratory determinations, across all types of samples, for the determination of soluble dietary fiber using amyloglucosidase B was 0.39%.

Tests for Enzyme Activity and Purity. Recoveries of wheat starch, corn starch, and casein measured at four time points over the duration of the reference analysis were 0.5 ± 0.2% (mean ± SE), 0.4 ± 0.1%, and 4.4 ± 1.4%, respectively, indicating normal enzyme activity for α-amylase, amyloglucosidase, and the proteases. Amyloglucosidase may often be contaminated with *endo*-cellulase, which causes hydrolysis of mixed linkage β-glucan from barley and oats, with a resultant underestimation of these components (*19*). When two sources of amyloglucosidase were compared, recovery of pure β-glucan was significantly less ($p < 0.01$) for amyloglucosidase A (90.5 ± 1.4%, $n = 3$), originally used in the AOAC procedure, than for amyloglucosidase B (96.8 ± 0.7%, $n = 3$), indicating that amyloglucosidase A may be contaminated with β-glucanase to a greater extent (**Table 2**). Cereal samples expected to contain β-glucan (oat and barley products) were reanalyzed using amyloglucosidase B, and it was found that 26 of 62 samples had higher values for soluble dietary fiber, 30 samples remained the same (i.e., the difference was within the standard error of the reference data, i.e., 0.39%), and six samples had lower values.

Calibration Model for Insoluble Dietary Fiber. An NIR model was obtained, using modified PLS regression, for the prediction of insoluble dietary fiber in cereal products ($n = 143$) containing a wide cross section of grains, additives, sugar, and fat contents. Using six cross-validation groups, the SECV for the insoluble dietary fiber model was 1.54% and the multiple coefficient of determination (R^2) 0.98 with an AOAC-determined sample mean of 10.05% (**Table 1**; **Figure 1**) and range of 0−48.77%. Linear regression of AOAC-determined insoluble fiber against NIR-predicted insoluble fiber ($Y = -0.05 + 1.00X$) gave an intercept and slope not significantly different from 0.0 and 1.0, respectively ($p > 0.5$). When the independent validation samples were predicted with the insoluble dietary fiber model, the SEP was 1.13%, the r^2 0.99, and the bias and slope 0.05% and 1.01, respectively (**Table 1**; **Figure 1**). The AOAC-determined sample mean and range were 9.66% and 0−39.48%, respectively. The intercept and slope of the linear regression line, plotting AOAC-determined versus NIR-predicted insoluble dietary fiber for the validation samples ($Y = -0.03 + 1.01X$), were not significantly different from 0.0 and 1.0, respectively ($p > 0.05$). One residual outlier was removed from the validation data set as the sample had a difference between AOAC-determined and NIR-predicted values of 2.5 times, or greater than, the standard error of the difference between the two values.

Calibration Models for Soluble Dietary Fiber. The statistics for the calibration models for soluble dietary fiber are presented

Table 2. Recovery of 150 mg of β-Glucan from the Soluble Dietary Fiber Assay (AOAC Method 991.43) Using Two Sources of Amyloglucosidase

	% recovery (dry wt basis)
amyloglucosidase A	90.52 ± 1.35[a]
amyloglucosidase B	96.75 ± 0.69[b]

[a] Mean ± SE, $n = 3$; numbers followed by different letters are significantly different from each other, Student's *t* test, $p < 0.01$.

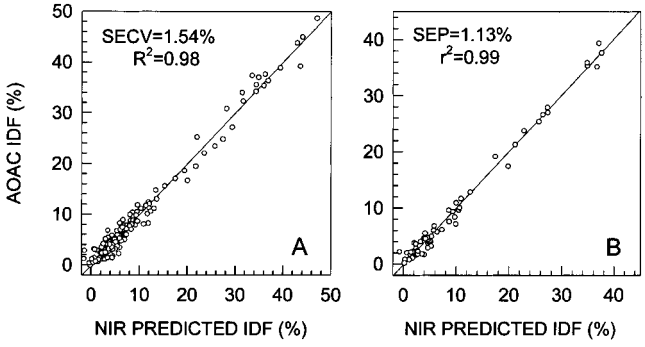

Figure 1. AOAC-determined versus NIR-predicted insoluble dietary fiber (IDF) for cereal products in the calibration (A) and validation (B) data sets.

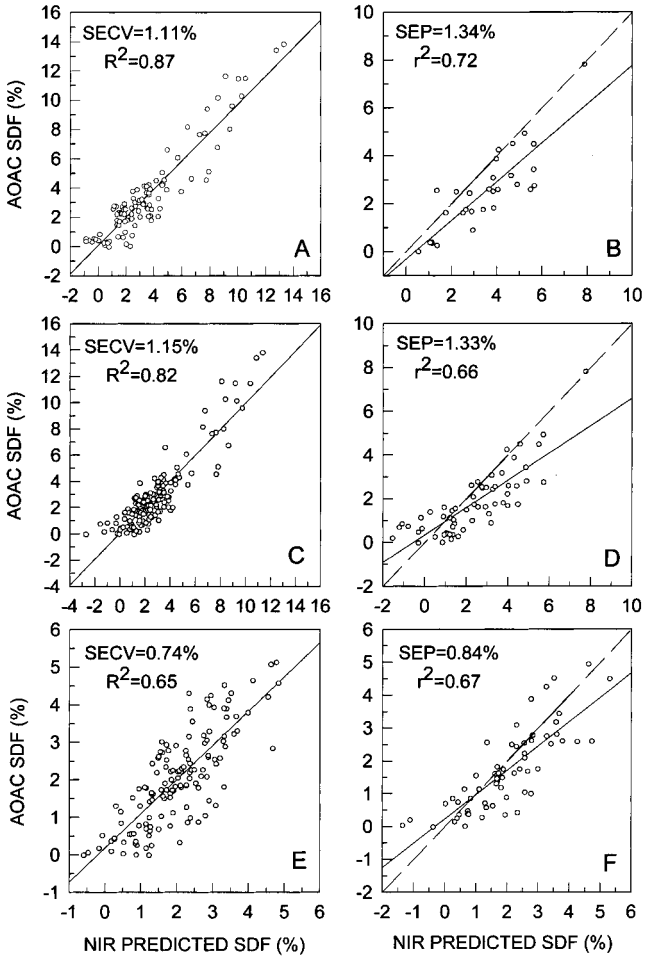

Figure 2. AOAC-determined versus NIR-predicted soluble dietary fiber (SDF) for cereal products in the calibration (A, C, and E) and validation (B, D, and F) data sets for three soluble fiber models: (1) model for products with low fat and low sugar contents (A and B); (2) model expanded to include products with high fat and high sugar contents (C and D); (3) model containing samples with <6% soluble dietary fiber (E and F).

in **Table 1** and **Figure 2**. The modified PLS model developed for soluble dietary fiber in low-fat and low-sugar samples, SDF model 1, had an SECV and R^2 of 1.11% and 0.87, respectively, with AOAC-determined sample mean and range of 3.51% and 0−13.84%, respectively. The modified PLS model was expanded to include high-fat and high-sugar calibration samples in addition to the low-fat, low-sugar samples (SDF model 2), and this new model had an SECV and R^2 of 1.15% and 0.82, respectively. The AOAC-determined sample mean and range were 2.85% and 0−13.84%, respectively. Linear regression of AOAC-determined soluble fiber against NIR-predicted soluble fiber for both SDF model 1 ($Y = 0.09 + 0.96X$) and SDF model 2 ($Y = 0.08 + 0.99X$) had intercepts and slopes not significantly different from 0.0 and 1.0, respectively ($p > 0.05$; **Figure 2**). The modified PLS model developed for the prediction of soluble fiber using low soluble fiber samples (0−6%, SDF model 3) had an SECV and R^2 of 0.75% and 0.64, respectively, with an AOAC-determined sample mean of 2.05% and a range of 0−5.14%. The intercept and slope of the linear regression of AOAC-determined versus NIR-predicted soluble dietary fiber ($Y = 0.20 + 0.91X$) were significantly different from 0.0 and 1.0 ($p < 0.05$), respectively.

Validation results for the three soluble dietary fiber models are presented in **Table 1** and **Figure 2**. The low-fat, low-sugar validation samples were predicted by SDF model 1 with an SEP and r^2 of 1.34% and 0.72, respectively (**Figure 2A**) and an AOAC-determined sample mean and range of 2.53% and 0.55−7.83%, respectively. (Two residual outliers were identified and removed from the validation data; i.e., the samples had a difference between AOAC-determined and NIR-predicted values of 2.5 times, or greater than, the standard error of the difference between the two values.) The slope and intercept of the linear regression of AOAC-determined versus NIR-predicted soluble dietary fiber ($Y = -0.30 + 0.81X$) values for the validation samples were significantly different from 0.0 and 1.0 ($p < 0.05$), respectively. A large mean bias was observed (−0.98) in the scatter plot (**Figure 2B**) with the model overpredicting soluble dietary fiber in 27 of the 30 samples.

Using SDF model 2 to predict soluble dietary fiber in the high-sugar and/or -fat expanded validation data set, the SEP and r^2 were 1.33% and 0.66, respectively, with an AOAC-determined sample mean of 1.80% and a range of 0−7.83%. The linear regression of AOAC-determined versus NIR-predicted values for soluble dietary fiber ($Y = 0.49 + 0.57X$) gave an intercept and slope that were significantly different from 0.0 and 1.0, respectively ($p < 0.05$). A systematic bias was observed with soluble dietary fiber being underestimated by SDF model 2 in samples below 1% soluble fiber and overestimated in samples above 1% (**Figure 2D**). One validation sample was identified as a spectral outlier (Mahalanobis distance > 3), using SDF model 2, and was removed from the validation data set. No residual outliers were present.

Using SDF model 3 with the validation samples containing both high- and low-fat and -sugar and <6% soluble fiber ($n = 61$), the resulting SEP and r^2 were 0.84% and 0.67, respectively, with an AOAC-determined sample mean of 1.70% and a range of 0−4.95%. As for SDF model 2, the slope and intercept of the linear regression of AOAC versus NIR values ($Y = 0.25 + 0.74X$) were significantly different from 0.0 and 1.0 ($p < 0.05$), and the scatter plot (**Figure 2F**) showed that soluble dietary fiber content in samples was, in general, underestimated in samples with <1% and overestimated in samples with >1% soluble dietary fiber. One spectral outlier (Mahalanobis distance > 3) was identified in the validation sample set and removed. No residual outliers were identified in the validation samples.

In an attempt to determine whether the reference method underestimated soluble fiber due to *endo*-cellulase contamination of amyloglucosidase, soluble dietary fiber was analyzed in oat- and barley-containing products using an alternative amyloglucosidase source (amyloglucosidase B) and PLS models developed for the prediction of soluble dietary fiber. The models had values for SECV and R^2 that were very similar to those for models developed with all of the reference data obtained using amyloglucosidase A. Furthermore, there was no improvement in SEP, r^2, slope, or bias when the models developed with reference data obtained using amyloglucosidase B were used to predict the independent validation samples.

PLS Loadings. The fat and sugar expanded model for insoluble dietary fiber used nine factors, which explained 97.6% of the spectral variation. Scores for factors 1, 2, and 3 had the highest correlation with insoluble dietary fiber and Pearson correlation coefficients of 0.46, 0.80, and 0.18, respectively. PLS loading plots show the regression coefficients of each wavelength to the constituent of interest and can indicate which wavelengths are important in developing a model. Wavelengths of high variation in the loading plots can be associated with areas of the spectrum of known chemical origin. Loading plots for factors 1 and 2 for the insoluble dietary fiber model indicate

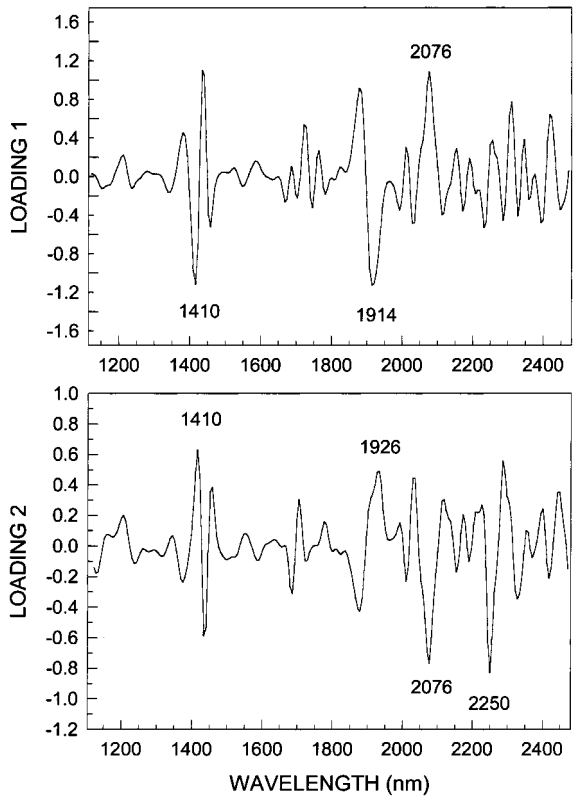

Figure 3. PLS loading spectra for insoluble dietary fiber in cereal products in a fat- and sugar-expanded model.

that vibrations due to OH groups in carbohydrate (2076 nm), CH groups in carbohydrate (2250 nm), and OH groups in water (1410 and 1914/1926 nm) are important in the model (**Figure 3**) (*25*). Hydroxyl groups in crystalline sugar may be responsible for strong absorbing in the 2076 nm region.

The fat and sugar expanded model (SDF model 2) for soluble dietary fiber used 12 PLS factors with scores for factors 1, 2, and 4 having the highest correlation with soluble dietary fiber and explaining 77.1% of the spectral variation. Pearson correlation coefficients for factors 1, 2, and 4 were 0.28, 0.47, and 0.28, respectively. Loading plots for the expanded, soluble dietary fiber model indicate that, as with the insoluble model, vibrations due to OH groups in carbohydrate (2076 nm) and water (1920/1926 nm) are important in the model (data not shown). Loadings for the low-fat, low-sugar model and the low soluble dietary fiber model were similar.

DISCUSSION

Accurate prediction of insoluble dietary fiber in cereal products has been achieved with an NIR model that encompasses a wide range of grain types, additives, such as dried fruits, herbs, nuts, honey, cinnamon, and cocoa, and processing techniques. The unique spectral characteristics of samples with high fat and high crystalline sugar content were incorporated into the model, resulting in similar accuracy to that of previous models for insoluble and total dietary fiber (*13*, *17*). The quantity of insoluble dietary fiber in packaged foods is an option that can be included on the product's nutrition label. In addition, the insoluble dietary fiber amount may be used in a method for the determination of caloric content of foods (*6*). Thus, rapid prediction of insoluble dietary fiber by NIR spectroscopy can aid the measurement of food components for nutrition labeling and monitoring.

Calibrations were developed for the prediction of soluble dietary fiber. The multiple coefficients of determination for soluble dietary fiber (0.64−0.87) are lower than those for insoluble dietary fiber in the current report (0.98) and lower than those for total dietary fiber in previous reports (0.98-0.99) (*14*, *16*, *17*). This may be due, partly, to a narrower range in values. The means and standard deviations for the AOAC-determined compared to the NIR-predicted values for the soluble fiber calibration are relatively similar. However, when the three soluble dietary fiber models are used to predict soluble fiber in the independent validation samples, a consistent bias was observed across the models. Validation samples with >1% soluble dietary fiber are, in general, overpredicted by the NIR models for soluble fiber. The overprediction increases as the values for soluble fiber increase. The statistics indicate a negative bias and a slope of prediction significantly different from 1.00 for each of the models.

The negative bias may be due to differences that exist between the calibration and validation data sets (*26*). However, the range in soluble dietary fiber in the validation data set is well within the calibration data set range and, although the distribution of values within the range is different when calibration and validation data sets are compared for SDF model 1 and SDF model 2, there is no difference for SDF model 3, where a bias also exists.

Errors in soluble dietary fiber reference analysis may also account for problems with NIR calibration for soluble dietary fiber. McCleary (*19*) reported a source of error may be the presence of a contaminating enzyme, *endo*-cellulase, in some sources of amyloglucosidase employed in the reference assay. *endo*-Cellulase activity in the assay can result in significant *endo*-depolymerization of mixed-linkage β-glucans from barley and oat. This would result in an underestimation of β-glucan during the dietary fiber procedure. In the present study soluble dietary fiber was underestimated in several cereal product samples expected to contain significant β-glucan. This emphasizes the need to monitor amyloglucosidase purity to obtain the best reference values for soluble dietary fiber. However, in the current study, no improvement in the calibration or validation of soluble dietary fiber was observed when soluble dietary fiber was determined using the alternative source of amyloglucosidase, even though the alternative source appeared to be of higher purity.

Manas et al. (*27*) also reported sources of error in the accepted methods of dietary fiber analysis indicating that the AOAC method for soluble dietary fiber is not as precise as the methods for total and insoluble dietary fiber. The distribution of soluble fiber between the soluble and insoluble fractions may vary depending on the assay conditions (e.g., temperature and pH) and composition of the individual samples. Soluble dietary fiber may be retained in the insoluble dietary fiber fraction to varying extents depending on the sample matrix; thus, the reference method may not be treating all samples alike. Errors due to soluble fiber partitioning into the insoluble fiber fraction would have a greater impact on soluble fiber accuracy as values are considerably lower than for insoluble fiber.

In summary, near-infrared spectroscopy can be used for the rapid and accurate determination of insoluble dietary fiber in a broad range of cereal products. Examination of the wavelengths of high variation in the PLS loadings suggests that analytically useful absorption is from OH groups in carbohydrate and water. It is concluded that the accuracy of NIR determination of soluble dietary fiber, in diverse cereal products, may be limited using currently available reference methods; however,

the NIR method may be useful for screening to determine whether individual products are a good or poor source of soluble dietary fiber.

ACKNOWLEDGMENT

Appreciation is expressed to Julie A. Auwater, R. Evan Miracle, and Laura L. Vines for excellent technical assistance.

LITERATURE CITED

(1) Anderson, J. W.; Hanna, T. J. Impact of nondigestible carbohydrates on serum lipoproteins and risk for cardiovascular disease. *J. Nutr.* **1999**, *129*, 1457S−1466S.

(2) Brown, L.; Rosner, B.; Willett, W. W.; Sacks, F. M. Cholesterol-lowering effects of dietary fiber: a meta analysis. *Am. J. Clin. Nutr.* **1999**, *69*, 30−42.

(3) Chandalia, M.; Abhimanyu, G.; Lutjohann, D.; von Bergmann, K.; Grundy, S. M.; Brinkley, L. J. Beneficial effects of high dietary fiber intake in patients with type 2 diabetes mellitus. *New Engl. J. Med.* **2000**, *342*, 1392−1398.

(4) Battilana, P.; Ornstein, K.; Minehira, K.; Schwarz, J. M.; Acheson, K.; Schneiter, P.; Burri, J.; Jéquier, E.; Tappy, L. Mechanisms of action of β-glucan in postprandial glucose metabolism in healthy men. *Eur. J. Clin. Nutr.* **2001**, *55*, 327−333.

(5) Izydorczyk, M. S.; Biliaderis, C. G. Cereal arabinoxylans: advances in structure and physiochemical properties. *Carbohydr. Polym.* **1995**, *28*, 33−48.

(6) *Code of Federal Regulations*. FDA, HHS; 21, part 101.9; 2001.

(7) Anderson, J. W.; Smith, B. M.; Gustafson, N. J. Health benefits and practical aspects of high-fiber diets. *Am. J. Clin. Nutr.* **1994**, *59*, 1242S−1247S.

(8) Williams, P. C.; Norris, K. H. *Near-infrared Technology in the Agricultural and Food Industries*; Williams, P. C., Norris, K. H., Eds.; American Association of Cereal Chemists: St. Paul, MN, 1987.

(9) Osborne, B. G.; Fearn, T.; Hindle, P. H. *Practical NIR Spectroscopy with Applications in Food and Beverage Analysis*; Longman Scientific and Technical: Harlow, U.K., 1993.

(10) Horvath, L.; Norris, K. H.; Horvath-Mosonyi, M.; Rigo, J.; Hegedus-Volgyesi, E. Study into determining dietary fiber of wheat bran by NIR-technique. *Acta Aliment.* **1984**, *13*, 355−382.

(11) Williams, P. C.; Cordeiro, H. M.; Harnden, M. F. T. Analysis of oat bran products by near-infrared reflectance spectroscopy. *Cereal Foods World* **1991**, *36*, 571−574.

(12) Baker, D. The determination of fiber in processed cereal foods by near-infrared reflectance spectroscopy. *Cereal Chem.* **1983**, *60*, 217−219.

(13) Kays, S. E.; Barton, F. E., II. The use of near-infrared reflectance spectroscopy to predict the insoluble dietary fibre fraction of cereal products. *J. Near Infrared Spectrosc.* **1998**, *6*, 221−227.

(14) Kays, S. E.; Windham, W. R.; Barton, F. E., II. Prediction of total dietary fiber in cereal products using near-infrared reflectance spectroscopy. *J. Agric. Food Chem.* **1996**, *44*, 2266−2271.

(15) *AOAC Official Methods of Analysis*, 15th ed., 3rd Suppl., Method 991.43. Total, Soluble, and Insoluble Dietary Fiber in Foods; AOAC: Arlington, VA, 1992.

(16) Kays, S. E.; Barton, F. E., II; Windham, W. R.; Himmelsbach, D. S. Prediction of total dietary fiber by near-infrared reflectance spectroscopy in cereal products containing high sugar and crystalline sugar. *J. Agric. Food Chem.* **1997**, *45*, 3944−3951.

(17) Kays, S. E.; Windham, W. R.; Barton, F. E., II. Prediction of total dietary fiber by near-infrared reflectance spectroscopy in high-fat- and high-sugar-containing cereal products. *J. Agric. Food Chem.* **1998**, *46*, 854−861.

(18) *AOAC Official Methods of Analysis*, 15th ed., Method 945.14. Moisture in Cereal Adjuncts: Air Oven Method (103−104°); AOAC: Arlington, VA, 1990.

(19) McCleary, B. V. Importance of enzyme purity and activity in the measurement of total dietary fiber and dietary fiber components. *J. AOAC Int.* **2000**, *83*, 997−1005.

(20) Shenk, J. S.; Westerhaus, M. O. Population definition, sample selection, and calibration procedures for near-infrared reflectance spectroscopy. *Crop Sci.* **1991**, *31*, 469−474.

(21) Barnes, R. J.; Dhanoa, M. S.; Lister, S. J. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Appl. Spectrosc.* **1989**, *43*, 772−777.

(22) Martens, H.; Naes, T. Assessment, validation and choice of calibration method. In *Multivariate Calibration*; Wiley: New York, 1989; pp 237−266.

(23) Hruschka, W. R. Data analysis: wavelength selection methods. In *Near-infrared Technology in the Agricultural and Food Industries*, 2nd ed.; Williams, P. C., Norris, K. H., Eds.; American Assocaition of Cereal Chemsits: St. Paul, MN, 2001; pp 39−58.

(24) Windham, W. R.; Mertens, D. R.; Barton, F. E., II. Protocol for NIRS calibration: sample selection and equation development and validation. In *Near Infrared Reflectance Spectroscopy (NIRS): Analysis of Forage Quality*; Marten, G. C., Shenk, J. S., Barton, F. E., II, Eds.; USDA Agriculture Handbook 643; U.S. Government Printing Office: Washington, DC, 1989; pp 96-103.

(25) Murray, I.; Williams, P. C. Chemical principles of near-infrared technology. In *Near-Infrared Technology in the Agricultural and Food Industries*; Williams, P. C., Norris, K. H., Eds.; American Association of Cereal Chemists: St. Paul, MN, 1987; pp 17−34.

(26) Dhanoa, M. S.; Lister, S. J.; France, J.; Barnes, R. J. Use of mean square prediction error analysis and reproducibility measures to study near infrared calibration equation performance. *J. Near Infrared Spectrosc.* **1999**, *7*, 133−143.

(27) Manas, E.; Bravo, L.; Saura-Calixto, F. Sources of error in dietary fibre analysis. *Food Chem.* **1994**, *50*, 331−342.

Received for review December 13, 2001. Revised manuscript received February 27, 2002. Accepted February 27, 2002.

JF0116552