Prediction of Total Dietary Fiber by Near-Infrared Reflectance Spectroscopy in High-Fat- and High-Sugar-Containing Cereal Products

Sandra E. Kays,* W. R. Windham, 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 near-infrared (NIR) spectral properties of cereal products containing high fat or high sugar can differ substantially from the spectral properties of other cereal products. An existing NIR model, using preprocessed reflectance spectra and partial least-squares analysis, for the prediction of total dietary fiber in cereal products was expanded to two new models called (1) the "fat-expanded" model and (2) the "fat- and sugar-expanded" model. The fat-expanded model enlarges the existing model with high-fat-content products as calibration samples, and the "fat- and sugar- expanded" model also includes products with high sugar and high crystalline sugar content. The dry milled cereal and grain products were analyzed in the laboratory according to AOAC method 991.43 for the determination of total dietary fiber, and NIR reflectance spectra were collected with a scanning monochromator. Data analysis and selection of representative high-fat and high-sugar samples were performed with a commercial analysis program. The two expanded models had standard errors of cross-validation and R^2 similar to those of the existing model, with acceptable standard error of performance and r^2 when tested with independent validation samples. The existing model was, thus, expanded to include high-fat, high-sugar, and high crystalline sugar cereal products while maintaining prediction accuracy.

Keywords: *Dietary fiber; near-infrared; chemometrics; cereal*

INTRODUCTION

Existing laboratory methods for the determination of dietary fiber provide precise and repeatable results for a wide range of foods; however, the analytical procedures are very time-consuming (Mongeau and Brassard, 1986; Lee et al., 1992; Englyst et al., 1994). The AOAC method (AOAC, 1992) used in the United States for nutrition labeling and monitoring requires 2 days to complete a determination, and extraction of fat or sugar can add a day to the time required for analysis. Near-infrared (NIR) spectroscopy has been used successfully in agriculture and the food industry for many years as an accurate and rapid method of analyzing certain constituents of products (Williams and Norris, 1987a; Osborne et al., 1993; Kays et al., 1998) and, thus, has been investigated as a method of dietary fiber analysis (Baker, 1983, 1985; Horvath et al., 1984; Williams et al., 1991; Kays et al., 1996, 1997).

A previous paper from this laboratory described the development of an NIR reflectance spectroscopy, partial least-squares (PLS) model for the prediction of dietary fiber in a broad range of cereal and grain products with a wide range of dietary fiber values (0−52% total dietary fiber; Kays et al., 1996). The model had a standard error of cross-validation (SECV) of 1.6% total dietary fiber and a multiple coefficient of determination (R^2) of 0.99 and was sufficiently accurate for nutrition labeling and monitoring purposes. The initial model was adapted for use with cereal samples that have a wide range of moisture contents (Windham et al., 1997). The SECV and R^2 of the moisture-adapted model were 1.8% total dietary fiber and 0.98, respectively. Both models were predominantly influenced by C−H and O−H groups in the carbohydrate- and water-absorbing regions of the spectrum, although minor influences from lipid and protein were also apparent.

The initial models did not include products with >10% fat and >20% sugar and did not predict total dietary fiber as well when tested with these products—unless fat and sugar were extracted prior to NIR scanning (Kays et al., 1996). Cereal products containing high fat and high sugar represent a significant portion of the cereal product market; however, they differ substantially in their spectral characteristics compared with other cereal products. The ability to incorporate the unique spectral characteristics of the high-fat and high-sugar products into a model for assessing dietary fiber would, therefore, be an advantage. A previous paper (Kays et al., 1997) described the unique spectral characteristics of sugars often found in cereals and of cereal products containing these sugars and described expansion of the initial calibration to include samples with high sugar and crystalline sugar content. The sugar-expanded model developed had an SECV and R^2 of 1.9% total dietary fiber and 0.98, respectively. Examination of PLS loadings suggested that the model was primarily influenced by carbohydrate with minor influences from water, lipid, and protein. In the current study, the initial model from Kays et al. (1997) is labeled the "original model" from which the two new expanded models, described in this paper, are derived.

* Author to whom correspondence should be addressed [telephone (706) 546-3338; fax (706) 546-3607].

S0021-8561(97)00648-1
Published on Web 02/10/1998

Cereal products containing significant amounts of fat (>10% fat) and having spectral properties that differ substantially from the spectral properties of other cereal products and high-sugar cereal products are, for example, the granola breakfast cereals, cereal-based snacks, and numerous types of crackers. The current work describes expansion of the original model for prediction of dietary fiber in cereal products to include (1) products with high fat content (fat-expanded model) and (2) products with high fat or high sugar content (fat- and sugar-expanded model) and testing of the expanded models with independent validation samples containing high fat or high sugar. In addition, the fat- and sugar-expanded model was tested on cereal samples containing both high fat and high sugar and on cereal samples having a wide range in moisture content. Observations were made on associations between wavelengths with high variation in the PLS loadings of the calibrations and known vibrations of interest in NIR spectra.

MATERIALS AND METHODS

Instrumentation, Reagents, and Enzyme Purity and Activity. Instrumentation, reagents, and testing of enzyme purity and activity were as previously described (Kays et al., 1996, 1997).

Samples and Sample Preparation. Cereal and grain products, including breakfast cereals, crackers, brans, flours, and cake and muffin mixes, were as described previously for the original calibration data set (Kays et al., 1997). Products used for the fat-expanded calibration and validation data sets contained >10% fat and included breakfast cereals, granolas, crackers, graham crackers, and cereal-based snacks. Twenty high-fat cereal products were available for expansion of the calibraton data set and 10 for expansion of the validation data set. High-fat samples were ground in a coffee mill (model KSM-2, Braun Inc., Lynnfield, MA). On the basis of the nutrition label value for products, the range, mean, and standard deviation of fat content of the high-fat samples used in the calibration were 10.7−30.0, 16.3, and 5.8%, respectively, and in the validation sample set were 11.1−23.3, 17.0, and 4.8%, respectively.

High-sugar samples were as described previously, and the types of sugars used by manufacturers to sweeten cereal products were described previously (Kays et al., 1997). Products were dry milled to <500 μm in a cyclone mill (Cyclotec 1093 sample mill, Perstorp Analytical, Silver Spring, MD). The high-sugar samples were mixed with liquid nitrogen to facilitate grinding. Thirty-seven high-sugar samples were available for expansion of the calibration data set and 14 high-sugar samples for validation data set expansion. The range, mean, and standard deviation of sugar content of the high-sugar samples used in the calibration were 21.8−53.3, 35.4, and 9.3%, respectively, and in the validation sample set were 22.2−55.6, 34.4, and 9.2%, respectively, as previously described (Kays et al., 1997).

Reference Laboratory Method for Total Dietary Fiber. Total dietary fiber in all samples was measured in the laboratory according to AOAC approved method 991.43 (AOAC, 1992; Lee et al., 1992), with modifications described by Kays et al. (1996). Before the AOAC procedure was performed, samples containing >10% fat 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. Samples containing >20% sugar were desugared by extracting four times with 85% ethyl alcohol (10 mL/g 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 was calculated for each sample on the basis of the sample weight before and after extraction. Total dietary fiber values for defatted or desugared samples were adjusted for the percent fat or sugar extracted. Total dietary fiber values for all samples were calculated on a dry weight basis (Kays et al.., 1996). Dry matter/moisture content of milled cereal products was determined according to AOAC air oven method 945.14 (AOAC, 1990).

Spectroscopic Analysis. All dry-milled cereal samples were scanned with the NIRSystems 6500 monochromator (NIRSystems, Silver Spring, MD). High-fat and high-sugar samples were scanned before defatting and desugaring. Duplicates of each sample were presented in cylindrical sample cells (internal diameter = 38 mm, depth = 9 mm) with optical quartz surface and cardboard backing. Each sample was scanned 16 times, and the data were averaged and transformed to log(1/*R*). The duplicate scans of each sample were examined visually for consistency and averaged. The wavelength range used for analysis was 1100−2500 nm.

NIR Calibraton on the Fat-Expanded Calibration Data Set. A commercial analysis program (NIRS 3 version 4.01, Infrasoft International Inc., Port Matilda, PA) was used to process data, select representative high-fat samples, and build chemometric models. As described for the sugar-expanded model (Kays et al., 1997), an algorithm called SELECT (Shenk and Westerhaus, 1991a,b) was used to select high-fat samples for calibration expansion. Eleven components (PLS factors) were used for the selection. Eighteen high-fat samples were chosen by the SELECT algorithm, of the 20 available, for calibration expansion. One of the 18 selected samples (Salted Sesame Sticks) was discarded as a spectral outlier (Mahalanobis distance > 3), based on PLS analysis. Seventeen high-fat samples were, thus, combined with the original 77 calibration samples to generate a 94-sample, fat-expanded calibration data set. A fat-expanded model was developed using the same preprocessing spectra transformations used for the original calibration (Kays et al., 1997). Briefly, log(1/*R*) spectra were transformed with standard normal variate and detrending procedures (Barnes et al., 1989), to remove multiplicative interferences of scatter, and then transformed with second-derivative processing (gap = 20 nm, smoothing interval = 10 nm). Data were subsequently centered using the CENTER program, available via NIRS 3 version 4.01, which allows centering of samples based on constituent values as well as spectral characteristics, i.e., partial least squares 1 (PLS1; Lindberg et al., 1983). Prior to calibration, log(1/*R*) spectra were mean centered, transformed with standard normal variate and detrending procedures (Barnes et al., 1989) to remove multiplicative interferences of scatter, and then transformed with second-derivative processing (gap = 8 nm, smoothing interval = 8 nm). Calibration was performed using modified PLS regression. 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 (Shenk and Westerhaus, 1991a). The optimum number of PLS factors used for total dietary fiber prediction was determined by cross-validation (Martens and Naes, 1989). During cross-validation one-sixth of the calibration samples at a time was temporarily removed from the calibration set and used for prediction, and performance statistics were accumulated for each group of removed samples. The optimal number of factors for total dietary fiber was that which produced a minimum in overall error between modeled and reference values (SECV). Preprocessing transformations were the optimum required to improve the SECV compared to PLS analysis with untransformed data.

NIR Calibration on the Fat- and Sugar-Expanded Calibration Data Set. A second model was developed using both the high-fat calibration expansion samples (n = 20) used above and the high-sugar calibration expansion samples from a previous paper (n = 37; Kays et al., 1997). Three spectral outliers were previously identified among the high-sugar samples, two in the calibration data set and one in the validation data set (Kays et al., 1997). These were not included in the present data set. As described for the sugar-expanded model (Kays et al., 1997) and the fat-expanded model above, a SELECT algorithm (Shenk and Westerhaus, 1991a,b) was used to select representative high-fat samples and high-sugar samples, from the pool of 57, for calibration expansion.

Table 1. High-Fat Cereal and Grain Products Used for the Fat-Expanded Calibration Data Set [Range, Mean, and Standard Deviation (SD) of Total Dietary Fiber (TDF) Percent]

sample	product type	no. of products	range in TDF %	mean TDF %	SD TDF %
high fat	wheat	8	2.0−9.6	5.1	3.3
	oats	3	4.8−7.9	6.3	1.5
	corn	1	6.1		
	multiple grain	5	4.0−10.9	6.9	3.0

[a] Grains represented in multiple-grain products in the high-fat sample set are wheat (4), oats (3), corn (2), rice (2), rye (1), and barley (1). The number of products of each grain type is in parentheses.

Table 2. High-Fat Cereal and Grain Products Used for the Fat-Expanded Validation Data Set [Range, Mean, and Standard Deviation (SD) of Total Dietary Fiber (TDF) Percent]

product type	no. of products	range in TDF %	mean TDF %	SD TDF %
wheat	6	2.1−7.4	4.5	1.8
oats	3	6.3−7.3	6.7	0.5
corn	1	5.2		

Eleven components were used for selection. Seventeen high-fat and 23 high-sugar samples were selected and were combined with the 77 samples from the original model (Kays et al., 1997) to generate a 117-sample fat- and sugar-expanded data set. Log(1/*R*) spectra were transformed and centered, as described for the fat-expanded data set. Finally, the fat- and sugar-expanded calibration model was developed using modified PLS (Shenk and Westerhaus, 1991a) with the same preprocessing spectra transformations used for the fat-expanded calibration model.

Model Validation. Each model was tested using the original 30 independent validation samples, plus 10 new independent high-fat samples for the fat-expanded model and plus 10 independent high-fat samples and 14 independent high-sugar samples for the fat- and sugar-expanded model. The fat- and sugar-expanded NIR model was also tested using independent cereal samples with both high-fat and high-sugar content ($n = 5$) and with milled cereal samples having a wide range in moisture content, i.e. milled cereal products subjected to four relative humidity treatments (20, 60, and 80% relative humidity and a vacuum oven), as described by Windham et al. (1997). Performance of models was reported as the standard error of performance (SEP), coefficient of determination (r^2), slope, and bias (Hruschka, 1987).

RESULTS

Total Dietary Fiber Measured by the Reference Method. The range for total dietary fiber in all cereal samples (including the original model), determined according to the AOAC enzymatic−gravimetric procedure, was from 0.6 to 52.1% ($N = 188$). For the high-fat samples the ranges in total dietary fiber for the calibration and validation data sets were from 2 to 10.9% and from 2.1 to 7.4%, respectively. The distribution of high-fat samples used for calibration expansion for each grain type is given in Table 1 and for the validation data set in Table 2. The range, mean, and standard deviation of total dietary fiber percent for each grain type is also presented. The standard error of the AOAC laboratory determinations (Windham et al., 1989) was 0.73% for the original samples, 0.36% for the high-fat samples, and 0.88% for the high-sugar samples.

Spectral Characteristics of Samples. NIR spectra (before preprocessing) of specific cereal samples are presented in Figure 1. The upper spectrum is of a high-fat cereal sample with a fat content of 14% and a sugar content of 2%. The middle spectrum is for a cereal that has low fat (0%) and low sugar (9%). The lower spectrum is for a cereal with a high sugar content of 47% and a low fat content of 3%. The spectrum of the high-fat sample differs from the other spectra in Figure 1 due to two sets of double bands, at 1728−1760 and 2310−2342 nm, characteristic of absorbance for C−H stretch groups present in lipids. The spectrum of the high-sugar sample differs from the other two spectra due to sharp absorbance peaks at 1434 and 2078 nm.

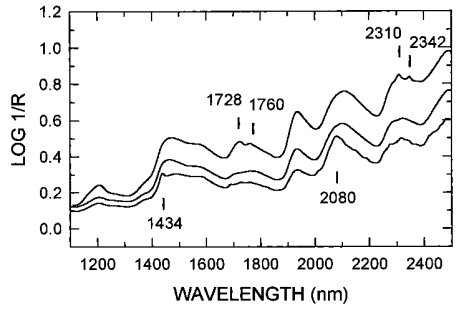

Figure 1. NIR spectra of a high-fat cereal product (14% fat, 2% sugar, upper plot), low-fat low-sugar cereal product (0% fat, 9% sugar, middle plot), and high-sugar cereal product (3% fat, 47% sugar, lower plot).

Fat-Expanded NIR Calibration Model for Total Dietary Fiber. The fat-expanded calibration data set contained cereal products with a range of fat content from high (>10% fat) to medium and low. An NIR calibration was obtained, using modified PLS, for the prediction of total dietary fiber in cereal products with a wide range in fat content. Using six cross-validation groups, the SECV for the fat-expanded model was 1.75% and the R^2 0.98 (Table 3). Linear regression of AOAC determined dietary fiber against NIR predicted dietary fiber ($Y = -0.14 + 1.00X$) gave an intercept and slope not significantly different from 0.0 and 1.0, respectively ($p > 0.05$). When 10 independent high-fat validation samples were combined with the original 30 independent validation samples and predicted with the high-fat model, the SEP was 1.77%, the r^2 0.98, and the bias 0.51 (Table 3). The intercept and slope of the linear regression line, plotting AOAC determined versus NIR predicted dietary fiber for the validation samples ($Y = 0.24 + 1.02X$), were not significantly different from 0.0 and 1.0, respectively ($p > 0.05$). A comparison of AOAC values versus NIR predicted values for total dietary fiber using the original model and the fat-expanded model is presented in Table 4. When the original model (Kays et al., 1997) was used to predict the high-fat validation samples, there was a marked overprediction of total dietary fiber and a large negative bias. The high-fat model corrected this deficiency. The high-fat model was used to predict the original validation samples ($n = 30$) alone with a resulting standard error of performance, r^2, bias, and slope of 1.87%, 0.98, 0.36%, and 1.04, respectively.

Fat- and Sugar-Expanded Calibration Model for Total Dietary Fiber. The fat- and sugar-expanded calibration data set contained cereal products with a broad range of fat content and a broad range of sugar and crystalline sugar content. Thus, an NIR model was obtained, using modified PLS, for prediction of total dietary fiber in cereal products with a wide range in fat and sugar contents. Using six cross-validation groups, the SECV for the fat- and sugar-expanded model

Table 3. Calibration and Validation Statistics for Dietary Fiber Prediction by the Fat-Expanded and Fat- and Sugar-Expanded NIR Models[a]

		calibration					validation[b]						
model	method	*n*	mean	SD	SECV	R^2	*n*	mean	SD	SEP	r^2	bias	slope
fat expanded	AOAC	94	14.02	13.03			40	12.93	12.53				
	NIRS	94	14.10	12.82	1.75	0.98	40	12.43	12.16	1.77	0.98	0.51	1.02
fat and sugar expanded	AOAC	117	12.59	12.13			54	10.98	11.48				
	NIRS	117	12.63	11.91	1.73	0.98	54	10.74	11.37	1.33	0.99	0.25	1.00

[a] Mean, standard deviation (SD), standard error of cross-validation (SECV), and multiple coefficient of determination (R^2). Mean, standard deviation, standard error of performance (SEP), coefficient of determination (r^2), bias, and slope for validation. [b] Independent validation samples ($n = 40$) used to test the fat-expanded model consisted of cereal samples with low fat and sugar ($n = 30$) plus samples with high fat (>10% fat, $n = 10$). Independent validation samples ($n = 54$) used to test the fat- and sugar-expanded model consisted of cereal samples with low fat and sugar ($n = 30$), plus cereal samples with high fat (>10% fat, $n = 10$), and cereal samples with high sugar (>20% sugar, $n = 14$).

Table 4. NIR Prediction of Total Dietary Fiber in Cereal Products

		NIR predicted TDF %[a]		
product	AOAC TDF %	original model[b]	fat-expanded model	fat- and sugar-expanded model
high fat				
Wild Blueberry Granola	7.32	15.80	5.48	5.67
Wheatables	5.50	12.28	5.67	5.54
Harvest Crisps	4.46	6.76	3.50	4.52
Cinnamon Grahams	3.07	3.00	1.47	1.78
Wheatsworth Crackers	4.52	11.25	5.29	6.01
Waverly Crackers	2.10	6.57	−0.33	1.50
Blue Corn Chips	5.23	19.81	4.06	4.68
Vegetable Crackers	7.36	10.55	8.21	9.11
Super Nutty Granola	6.28	16.62	4.69	5.72
Cashew Almond Granola	6.59	13.07	5.07	5.57
high sugar				
Fruit and Fiber	10.58	15.93		11.25
Healthy Choice	9.36	11.31		8.18
Golden Crisp	3.59	9.37		4.78
Honey Grahams	4.80	4.52		2.91
Oatmeal Crunch	4.52	4.18		3.67
Blueberry Morning	3.38	3.89		2.86
Golden Grahams	3.54	3.33		2.12
Honey Nut Cheerios[c]	7.29	0.24		6.35
Nut and Honey Crunch[c]	2.45	−3.97		1.12
Apple Cinnamon Toasted Oats[c]	6.87	−4.00		5.33
Honey Crunch Corn Flakes[c]	1.84	−7.09		0.49
Oat Bran Muffin Mix[c]	4.65	−10.62		4.28
Double Dip Crunch[c]	1.01	−9.97		−0.50
Cocoa Pebbles[c]	1.42	−17.53		0.24
high fat and high sugar				
Oats'n Honey Granola Bars	5.30	16.18		4.16
100% Natural Cereal	6.54	15.69		7.93
Cinnamon Toast Crunch	3.62	−1.43		1.48
Famous Amos Cookies	3.27	14.44		2.36
Banana Nut Crunch	5.84	11.79		5.14

[a] Total dietary fiber. [b] Original model (Kays et al., 1997). [c] Spectral evidence of crystalline sugar.

was 1.73% and the R^2 0.98 (Table 3). Linear regression of AOAC determined dietary fiber versus NIR predicted dietary fiber ($Y = -0.14 + 1.01X$) gave an intercept and slope not significantly different from 0.0 and 1.0, respectively ($p > 0.05$; Figure 2A). The combined model was tested using the original 30 validation samples plus the 10 high-fat and 14 high-sugar validation samples. The SEP, r^2, bias, and slope were 1.33%, 0.99, 0.25%, and 1.00, respectively. Linear regression of AOAC determined versus NIR predicted dietary fiber for the validation samples ($Y = 0.21 + 1.00X$) gave an intercept and slope not significantly different from 0.0 and 1.0, respectively ($p > 0.05$; Figure 2B). When the original model (Kays et al., 1997) was used for prediction of total dietary fiber in the original validation samples plus the 10 high-fat and 14 high-sugar validation samples, the SEP, r^2, bias, and slope were 5.59%, 0.81, 0.15%, and 0.8, respectively. The large SEP was due to marked overprediction of the high-fat samples and marked underprediction of samples containing crystalline sugar (Table 4). A comparison of AOAC values versus NIR predicted values using the original model (Kays et al., 1997) and the fat- and sugar-extended model for the corresponding validation samples is presented in Table 4. Data are also included in Table 4 showing prediction, by the original model and the fat- and sugar-expanded model, of total dietary fiber in cereal products ($n = 5$) containing both high fat and high sugar. The original model markedly overpredicted the total dietary fiber content of four of the five samples in this category, resulting in a large overall negative bias for these samples. The fat- and sugar-expanded model resulted in correct predictions. When the original validation samples alone ($n = 30$) are predicted by the combined model, the resulting SEP, r^2, bias, and slope are 1.40%, 0.98, −0.05%, and 1.03, respectively.

Performance of the fat- and sugar-expanded model was tested with spectra of independent validation samples of variable moisture content. Spectra of milled cereal products ($n = 27$) conditioned at 20, 60, and 80%

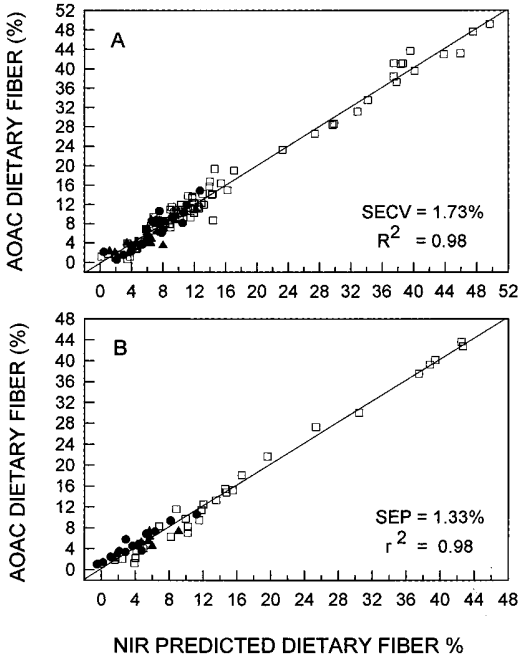

Figure 2. AOAC-determined total dietary fiber vs NIR predicted total dietary fiber for cereal products in the calibration data set (n = 117, A) and independent validation data set (n = 54, B) of the fat- and sugar-expanded model. Open squares represent original cereal samples, solid triangles represent high-fat samples (>10% fat), and solid circles represent high-sugar samples (>20% sugar).

relative humidity and in a vacuum oven, described by Windham et al. (1997), were utilized. The resulting SEP, r^2, bias, and slope were 2.09%, 0.98, −0.55%, and 1.12, respectively. The larger SEP (2.09% compared to 1.73%) was due to samples stored in the vacuum oven and some of the samples stored at 80% relative humidity. In general, these samples were overpredicted, resulting in a larger overall negative bias for each of these groups. The overall bias for samples conditioned in the vacuum oven was −1.23%. These samples had a residual moisture range of 1.19−2.85% with mean ± SD of 2.30 ± 0.34%. The overall bias was −0.77% for samples conditioned at 80% relative humidity. These samples had a residual moisture range of 12.86−16.21% and mean ± SD of 15.11 ± 1.13%. The residual moisture range and mean ± SD of milled cereal samples used in the fat- and sugar-expanded calibration data set (n = 117) were 3.03−12.89 and 7.34 ± 2.57%, respectively, and for the validation data set (n = 54) were 3.31−12.2 and 6.69 ± 2.32%, respectively. Thus, total dietary fiber is predicted well in cereal samples with moisture content within the moisture range of the samples in the calibration. In the authors' experience the moisture range of whole cereal samples immediately after removal from the product packaging and before milling is from 1.4 to 13.7%. The reason for a somewhat narrower moisture range in the milled cereals used for the calibration data set may be due to the effect of ambient conditions during milling and sampling with an accompanying uptake of moisture in some extremely dry samples or the loss of moisture from the wettest samples. Prediction of dietary fiber in milled cereal products with moisture content outside the calibration range may require a moisture-expanded equation [see Windham et al. (1997)].

PLS Loadings. The fat-expanded model employed eight factors that explained 98.4% of the spectral

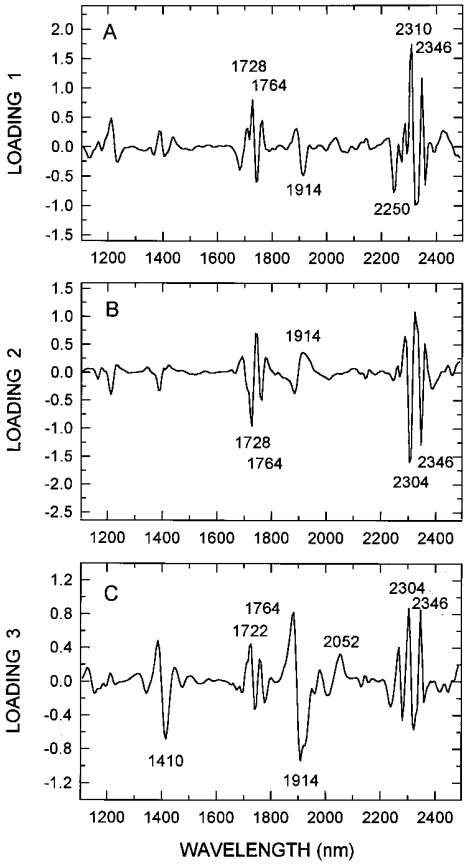

Figure 3. PLS loading spectra for total dietary fiber in cereal products in the fat-expanded model. Panels A, B, and C represent loadings for factors 1, 2, and 3, respectively.

variation. Sample scores having the highest correlation with dietary fiber were for factors 1, 2, and 3 with Pearson correlation coefficients of 0.56, 0.72, and 0.38, respectively.

PLS loading plots show the regression coefficients of each wavelength to dietary fiber for each factor and can indicate which wavelengths have the highest variation in a calibration set. Wavelengths of high variation may be associated with areas of the spectrum of known chemical origin. The calibration samples for the fat-expanded data set and the fat- and sugar-expanded data set appear to have loadings of high relative positive or negative values primarily at wavelengths that are associated with lipid and carbohydrate, respectively. Key values in loadings may be associated with known vibrations, and together with the correlation of individual loadings with fiber, fat, or sugar, an estimate may be obtained of which chemical component is the primary contributor for each loading. Factor 2 of the fat-expanded model was the most highly correlated to dietary fiber and had a loading with the greatest absorbances at 1728, 1764, 2304, and 2346 nm, associated with absorbance of C−H stretch groups in lipid (Figure 3B) (Murray and Williams, 1987; Williams and Norris, 1987b). The first loading was the second most highly correlated to dietary fiber, and its loading plot also has the most prominent absorption peaks in the spectral regions associated with C−H stretch groups of lipid (Figure 3A). The third PLS loading has absorption bands associated with C−H stretch groups of lipid in addition to O−H groups of the water band at 1410 and 1914 nm and amide groups in the protein region at 2052

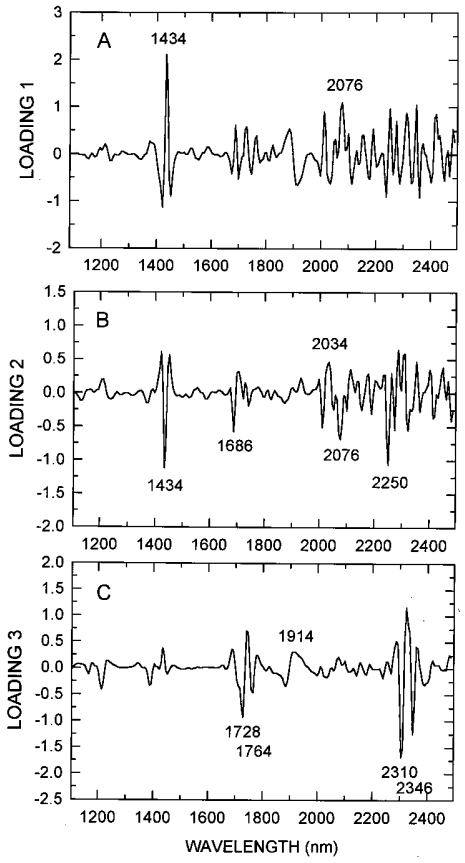

Figure 4. PLS loading spectra for total dietary fiber in cereal products in the fat- and sugar-expanded model. Panels A, B, and C represent loadings for factors 1, 2, and 3, respectively.

nm (Figure 3C). Overall, the constituent having the greatest influence in the high-fat model appears to be lipid via C−H stretch absorbance. Other influences appear to be from carbohydrate at 2250 nm and from water at 1914 and 1410 nm.

The fat- and sugar-expanded model utilized nine factors that accounted for 98.1% of the spectral variation. Sample scores having the highest correlation with dietary fiber were for factors 1, 2, and 3 with correlation coefficients of 0.46, 0.79, and 0.31, respectively. The second factor of the combined model was the most highly correlated to dietary fiber, and its loading had a sharp absorbance peak at 1434 nm (Figure 4B), similar to that in loadings 1 and 2 of the high-sugar model (Kays et al., 1997). Peaks also occurred in the second loading for the fat- and sugar-expanded model at 2076 and 2250 nm, suggesting absorption by groups in carbohydrate. The loading for the first factor (second most highly correlated to dietary fiber, Figure 4A) had a major absorbance peak at 1434 nm and smaller peaks at 1914 nm correlated to O−H absorbance in the water band, at 2076 nm correlated to O−H absorbance in mono- or disaccharides, and from 2200 to 2300 nm, a region often associated with C−H groups in carbohydrate. The loading for the third factor had principle areas of absorption at 1728, 1764, 2310, and 2346 nm typical of C−H stretch groups in lipid (Figure 4C). The fat- and sugar-expanded model was predominantly influenced by carbohydrate with minor influences due to lipid and water.

PLS Scores. Plots of factor 1 versus factor 2 scores for dietary fiber are represented in Figure 5. For the

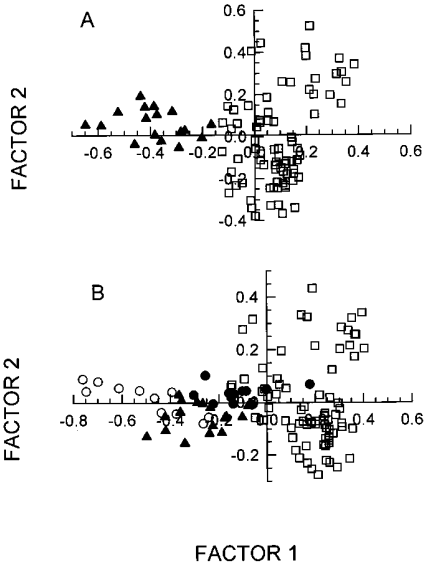

Figure 5. Plots of the PLS scores for factor 1 vs factor 2 for the fat-expanded model (A) and the fat- and sugar-expanded model (B). Open squares represent original cereal samples, solid triangles represent high-fat samples (>10% fat), solid circles represent high-sugar samples (>20% sugar), and open circles represent high-sugar samples (>20% sugar) having spectral evidence of crystalline sugar.

fat-expanded model, fat content, based on product nutrition label values, appears to be a significant contributor to factor 1 (Pearson correlation coefficient = −0.85) resulting in distribution of samples of high to low fat content along the factor 1 axis (Figure 5A). The predominant contribution to factor 2 appears to be dietary fiber (Pearson correlation coefficient = 0.72). For the fat- and sugar-expanded model the predominant contributor to factor 1 appears to be sugar content, based on product nutrient label values (Pearson correlation coefficient = −0.73). As in the sugar-expanded model (Kays et al., 1997) the crystalline sugar samples, which in general contain the largest amount of sugar, are visualized along the factor 1 axis (Figure 5B). For factor 2 the major contribution appears to be from dietary fiber (Pearson correlation coefficient = 0.79). Fat content did not appear to make a substantial contribution to any of the factors in the fat- and sugar-expanded model, based on Pearson correlation coefficients.

DISCUSSION

This study and a previous study have investigated the possibility of including the unique spectral characteristics of high-fat and high-crystalline-sugar cereal products into an NIR calibration for accurate and rapid determination of dietary fiber. First, the original NIR model was expanded to include cereal products with high sugar content, including products with large amounts of crystalline sugar (Kays et al., 1997). The SECV for the model was slightly higher than for the original model (1.88% compared to 1.64%); however, the model predicted samples with a much wider range of sugar content. Second, in the present study, the original model was expanded to include cereal products with high fat content. The fat-expanded model had a level of accuracy similar to that of the original model (SECV = 1.75%, R^2 = 0.98). Third, the original model was

expanded to include both high-fat and high-sugar products. As with the previous models, the fat- and sugar-expanded model encompasses a broad range of cereal products, such as breakfast cereals, crackers, brans, flours, pastas, cookies, and a broad range of grains, such as wheat, oats, barley, rye, corn, millet, amaranth, and products containing multiple grains. In addition, the fat- and sugar-expanded model encompasses products with a wide range in fat content, sugar content, and crystalline sugar content. The fat- and sugar-expanded model had an SECV similar to that of the original model and was found to accurately predict total dietary fiber in an independent group of samples also with a broad range of product types, grains, fat content, sugar content, and total dietary fiber content. The SECV, standard error of performance, bias, slope, and coefficients of determination observed indicate a high degree of precision and reliability in determining dietary fiber using the fat- and sugar-expanded calibration.

Examination of the loadings in the fat- and sugar-expanded calibration equation suggests that effects due to O−H and C−H groups in carbohydrate regions are most important in the model. The loadings for the two factors most highly correlated to dietary fiber have significant intensity in these regions (1434, 2076, and 2200−2300 nm). It is only in the loading plot for the third factor that significant intensity associated with absorbance by C−H stretch groups in lipid was observed. This loading is very similar to loading 2 of the fat-expanded model. Overall, the fat- and sugar-expanded model appears to be influenced by carbohydrate with minor influences from lipid, water, and protein. This is in contrast to the fat-expanded model, in which the major influence is from lipid. Loadings for factors 1 and 2 of the fat- and sugar-expanded model very closely resemble the loadings of factors 1 and 2 for the high-sugar model (Kays et al., 1997b). In both models the loadings are dominated by sharp bands at 1434 and 2076 nm. The shape and positioning of these bands are indicative of the importance of O−H groups in crystalline sugar in the development of these models.

CONCLUSIONS

The NIR spectroscopic model for prediction of total dietary fiber was expanded to include high-fat and high-sugar samples while maintaining the precision of the original calibration. The fat- and sugar-expanded model has the potential to reduce the time required for dietary fiber determination from several days to several minutes and has utility for agencies or industries needing to estimate total dietary fiber in a broad range of cereal products.

ACKNOWLEDGMENT

We thank Dr. Richard G. Leffler for valuable technical suggestions, Ms. Judy E. Davis and Mr. A. Terrell Boynton for technical assistance, and Ms. Lee Dexter of Canadian Harvest U.S.A. L.P. for supplying some of the samples used to develop the calibrations.

LITERATURE CITED

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

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

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

Baker, D. The determination of fiber, starch, and total carbohydrate in snack foods by near-infrared reflectance spectroscopy. *Cereal Foods World* **1985**, *30*, 389−392.

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

Englyst, H. N.; Quigley, M. E.; Hudson, G. J. Determination of dietary fiber as non- starch polysaccharides with gas liquid chromatographic, high-performance liquid chromatographic or spectrophotometric measurement of constituent sugars. *Analyst* **1994**, *119*, 1497−1509.

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.

Hruschka, W. R. Data analysis: wavelength selection methods. 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 53−54.

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.

Kays, S. E.; Barton, F. E., II; Windham, W. R.; Himmelsbach, D. S. The 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.

Kays, S. E.; Barton, F. E., II; Windham, W. R. NIR analysis of dietary fiber. In *Complex Carbohydrates: Definition, Analysis, and Applications*; Lee, S. C., Prosky, L., Eds.; Dekker: New York, 1998; in press.

Lee, S. C.; Prosky, L.; De Vries, J. W. Determination of total, soluble, and insoluble dietary fiber in foods—enzymatic-gravimetric method, MES-TRIS buffer: collaborative study. *J. AOAC Int*. **1992**, *75*, 395−416.

Lindberg, W.; Persson, J.-A.; Wold, S. Partial least-squares method for spectrofluorimetric analysis of mixtures of humic acid and ligninsulfonate. *Anal. Chem*. **1983**, *55*, 643−648.

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

Mongeau, R.; Brassard, R. A rapid method for the determination of soluble and insoluble dietary fiber: comparison with AOAC total dietary fiber procedure and Englyst's method. *J. Food Sci*. **1986**, *51*, 1333−1336.

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 Assocication of Cereal Chemists: St. Paul, MN, 1987; pp 17−34.

Osborne, B. G.; Fearn, T.; Hindle, P. H. *Practical NIR Spectroscopy with Applications in Food and Beverage Analysis*; Longman Scientific and Technical, Harlow, England, 1993.

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

Shenk, J. S.; Westerhaus, M. O. Population structuring of near-infrared spectra and modified partial least squares regression. *Crop Sci*. **1991b**, *31*, 1548−1555.

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

Williams, P. C.; Norris, K. H. Qualitative applications of near-infrared reflectance spectroscopy. 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, 1987b; pp 241−246.

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.

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. GPO: Washington, DC, 1989; pp 96-103.

Windham, W. R.; Kays, S. E.; Barton, F. E., II. Effect of cereal product residual moisture content on total dietary fiber determined by near-infrared reflectance spectroscopy. *J. Agric. Food Chem*. **1997**, *45*, 140−144.

Received for review July 29, 1997. Revised manuscript received December 16, 1997. Accepted December 17, 1997. Reference to company or trade names is for the purpose of description only and does not imply endorsement by the U.S. Department of Agriculture.

JF9706484