JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Exploring Authentic Skim and Nonfat Dry Milk Powder Variance for the Development of Nontargeted Adulterant Detection Methods Using Near-Infrared Spectroscopy and Chemometrics

Lucy L. Botros,*[,†] Joseph Jablonski,[‡] Claire Chang,[‡] Marti Mamula Bergana,[§] Paul Wehling,[∥] James M. Harnly,[⊥] Gerard Downey,[▽] Peter Harrington,[▲] Alan R. Potts,[†] and Jeffrey C. Moore[†]

[†]U.S. Pharmacopeial Convention, 12601 Twinbrook Parkway, Rockville, Maryland 20852-1790, United States

[‡]Center for Food Safety and Applied Nutrition, Division of Food Processing Science and Technology, U.S. Food and Drug Administration, 6502 South Archer Road, Bedford Park, Illinois 60501, United States

[§]Abbott Nutrition Division, Abbott Laboratories, 3300 Stelzer Road, Columbus, Ohio 43219, United States

[∥]General Mills, Inc., 330 University Avenue SE, Minneapolis, Minnesota 55414, United States

[⊥]Food Composition and Methods Development Laboratory, Beltsville Human Nutrition Research Center, Agricultural Research Service, U.S. Department of Agriculture, Building 161, BARC-East, Beltsville, Maryland 20817, United States

[▽]Teagasc Food Research Centre Ashtown, Ashtown, Dublin 15, Ireland

[▲]Center for Intelligent Chemical Instrumentation, Department of Chemistry and Biochemistry, Clippinger Laboratories, Ohio University, Athens, Ohio 45701-2979, United States

ABSTRACT: A multinational collaborative team led by the U.S. Pharmacopeial Convention is currently investigating the potential of near-infrared (NIR) spectroscopy for nontargeted detection of adulterants in skim and nonfat dry milk powder. The development of a compendial method is challenged by the range of authentic or nonadulterated milk powders available worldwide. This paper investigates the sources of variance in 41 authentic bovine skim and nonfat milk powders as detected by NIR diffuse reflectance spectroscopy and chemometrics. Exploratory analysis by principal component analysis and varimax factor rotation revealed significant variance in authentic samples and highlighted outliers from a single manufacturer. Spectral preprocessing and outlier removal methods reduced ambient and measurement sources of variance, most likely linked to changes in moisture together with sampling, preparation, and presentation factors. Results indicate that significant chemical variance exists in different skim and nonfat milk powders that will likely affect the performance of adulterant detection methods by NIR spectroscopy.

KEYWORDS: *skim milk powder, nonfat dry milk, melamine, NIR spectroscopy, chemometrics, PCA, varimax, compendial, diffuse reflectance, variance*

INTRODUCTION

Skim milk powder (SMP) and nonfat dry milk (NFDM) are important food ingredients and sources of nutrition, with more than 9 billion pounds estimated to be produced globally in 2011.[1] Numerous testing standards exist for both of these ingredients and other milk derivatives, but no authoritative testing standards currently exist for verifying the identities and integrities of these ingredients. This was underscored by the tragic 2008 melamine adulteration incident involving milk powders, which highlighted vulnerabilities in existing food safety and quality assurance systems that were not capable of guarding against the possibility of unknown adulterants.[2−4]

A workshop on this topic was convened by the United States Pharmacopeia (USP) in 2009 entitled "Food Protein Workshop—Developing a Toolbox of Analytical Solutions to Address Adulteration."[5] One of the key outcomes from the meeting was a need for standardized and reliable nontargeted screening procedures combined with multivariate statistical analysis tools to assess food ingredients rich in protein, such as milk and plant protein-derived ingredients, in quality assurance (QA) and quality control (QC) settings. Such procedures would become useful tools to allow authentication of ingredients based on a qualitative comparison with a library of milk powders, with the expectation that adulterated samples would classify as outliers and as such be considered nonauthentic. This nontargeted approach has the potential to significantly advance a solution to the age-old problem of using targeted methods to detect adulteration—as those responsible for adulteration are constantly evolving and engineering new, previously unknown adulterants to circumvent existing targeted QC methods.

Several promising analytical methods, including near-infrared (NIR) spectroscopy, are currently being investigated by a USP-led collaborative research project aimed at developing and validating a toolbox of methods to detect adulteration in SMP and NFDM.[6] Benefits of NIR spectroscopy compared to other technologies include its ready availability, low cost, high

Received: May 28, 2013
Revised: September 9, 2013
Accepted: September 16, 2013
Published: September 16, 2013

Table 1. Certificates of Analysis Data from 41 Milk Powder Samples Acquired from Eight Suppliers, Produced between August 2008 and May 2012

	particle size[a] (μm)						production			content (%)		
sample	$d(0.1)$	$d(0.5)$	$d(0.9)$	supplier	class[b]	process type[c]	location	country	date	moisture	fat	protein
S021	9.6	36.6	86.8	A	NFDM	LH	A-1	USA	7/12/2010	3.6	0.65	35.67
S022	15.7	60.7	133.4	A	SMP		A-1	USA	2/27/2010	3.8	0.4	33.71
S023	16.2	51.2	112	A	NFDM	MH	A-1	USA	5/5/2010	3.3	0.67	35.4
S024	7.3	32.6	87.7	A	NFDM	MH	A-1	USA	5/5/2010	3.1	0.66	35.56
S030	14.5	42.2	86.2	A	NFDM	HH	A-2	USA	7/18/2010	3.87	1.05	
S031				A	NFDM	HH	A-2	USA	11/16/2009	3.49	0.69	
S032				A	NFDM	LH	A-2	USA	6/19/2010	3.92	0.99	
S033	11.1	36.6	77.2	A	NFDM	LH	A-2	USA	2/26/2010	3.71	0.95	
S047	10.7	40.2	92.8	A	NFDM	LH	A-1	USA	6/7/2010	3.68	0.59	35.5
S051	13.2	45.9	105.6	A	NFDM	LH	A-1	USA				
S053	18.2	54.6	124	A	NFDM	LH	A-1	USA				
S054	16.7	68.1	152.8	A	NFDM	LH	A-1	USA				
S055	17.6	54.9	124.1	A	NFDM	LH	A-1	USA	8/26/2008	3.63	0.83	35.69
S061	16.4	57.8	124.9	H	NFDM	LH	H-1	USA	3/8/2011	3.4	0.54	
S068	17.2	58.6	124.7	H	NFDM	LH	H-1	USA	2/21/2011	3.46	0.573	
S070	15.5	57.2	122	H	NFDM	LH	H-1	USA	2/7/2011	3.294	0.62	
S076	13.2	43.5	94.4	A	NFDM	HH	A-2	USA	1/14/2011	3.5	0.61	
S077	20.5	65.5	139.5	A	SMP	LH	A-1	USA	2/21/2011	4	0.58	33.4
S080	11.5	42.6	110.8	B	SMP	LH	B-1	USA	3/27/2011	4	0.65	34.29
S081[d]	21.5	69.4	196.9	B	NFDM	HH	B-2	USA	3/9/2011	3.17	0.75	35.44
S082[d]	29.8	123.8	432.9	B	NFDM	LH	B-2	USA	2/27/2011	3.59	0.66	36.09
S084	10.6	37.3	89.2	B	SMP	MH	B-1	USA	1/30/2011	3.66	0.6	34.06
S085	11.1	40.5	99	B	SMP	MH	B-1	USA	3/8/2011	3.85	0.69	34.22
S086[d]	23.6	80.9	245.8	B	NFDM	HH	B-2	USA	1/15/2011	3.58	0.7	35.64
S087[d]	19.8	77.9	193.2	B	NFDM	LH	B-2	USA	3/7/2011	3.29	0.62	35.84
S089	15.3	53.2	143.6	B	NFDM	MH	B-3	USA	3/12/2011	3.6	0.78	35.73
S091[d]	16	68	231.4	B	NFDM	MH	B-3	USA	12/26/2010	3.78	0.95	36.31
S093	20.1	64.2	146.3	A	NFDM	LH	A-1	USA	2/1/2011	3.8	0.76	36.04
S094	10.1	36.3	89.8	A	NFDM	MH	A-1	USA	2/13/2011	3.8	0.77	35.9
S095	23.4	82.3	187.6	D	SMP	MH	D-1	New Zealand	10/20/2010	3.9	1	32.7
S096[d]	17.3	57.8	125	B	SMP	MH	B-4	USA	2/8/2011	3.96	0.67	34.12
S097	15.3	50.7	110.8	B	SMP	LH	B-4	USA	3/12/2011	3.78	0.69	34.3
S098	13.7	41.6	89.9	B	SMP	LH	B-4	USA	8/29/2010	3.92	0.75	34.4
S106	21.8	66	148.8	E	SMP	MH	E-1	Ireland	8/17/2010	3.82	0.95	37
S107	22.8	65.3	135.1	E	SMP	MH	E-1	Ireland	5/15/2010	4.49	0.95	35.7
S108	14.9	51.5	116.2	G	NFDM							
S110	12.7	37.6	77.3	G	NFDM							
S116	13.1	39	83.2	C	SMP	MH	C-1	Denmark	4/2/2011	4	0.5	
S117	15.2	39.7	89.8	C	SMP	MH	C-1	Denmark	3/15/2011	4	0.07	
S145				F	NFDM	LH	F-1	USA	5/12/2012	1.8	0.01	
S149				F	NFDM	HH	F-1	USA	5/15/2012	2.37	0.02	

[a] $d(0.5)$, median diameter; $d(0.9)$, 90% of volume distribution below the given diameter; $d(0.1)$, 10% of volume distribution below the given diameter. [b] NFDM, nonfat dry milk; SMP, skim milk powder. [c] LH, low heat; MH, medium heat; HH, high heat. [d] Samples characterized by HPSEC and LC-UV.

throughput, and robust and rapid analytical measurements. However, developing these nontargeted classification methods is complicated by the potential physicochemical variability of pure, nonadulterated milk powder ingredients in commerce worldwide. Such variations are well-known to broaden the range and classification boundaries of authentic ingredients, thereby decreasing the method's sensitivity for detecting lower concentration adulterants. This problem is especially true with NIR diffuse reflectance spectroscopy, which already exhibits typical detection limits on the order of 0.1%, where physical properties influence the resulting spectra and chemical signatures are not well-resolved.

The basic compositional variability of SMP and NFDM (e.g., total protein, lactose, water, fat, and ash) is thought to be somewhat limited by the standardization of raw milk used to produce these powders and international standardization efforts for product compositions. Little is known, however, of the variability of minor chemical constituents, such as milk metabolites, small-molecule additives, and protein composition in commercial SMP and NFDM, and their influence on NIR spectra. For raw fluid milk, factors reported to influence these minor constituents include raw milk geographic origin, animal origin (e.g., bovine versus water buffalo) and breed, season, and animal diet.[7] For further processed ingredients like milk powders, processing parameters such as preheating temper-

dx.doi.org/10.1021/jf4023433 | J. Agric. Food Chem. 2013, 61, 9810−9818

atures, concentrate heating temperatures, drying temperatures, and drying equipment (e.g., spray versus drum driers), may also introduce additional chemical and physical differences that are measurable by NIR spectroscopy. This was confirmed by a study that reported that heat treatment type (low, medium, or high heat) could be discriminated by NIR spectroscopy and chemometrics.[8] More research is therefore needed to better characterize the NIR variance of commercial SMP and NFDM and determine how this variance may affect the performance of nontargeted NIR analysis methods for detecting adulteration.

Understanding the repeatability and reproducibility of a NIR measurement is an important consideration when developing classification methods for detecting adulteration. Advanced NIR platforms for solid-phase reflectance spectroscopy are available and have been designed to reduce the effects of instrumental variance. The use of standard materials to monitor and verify instrumental calibration, like wavelength accuracy, photometric linearity and accuracy, and noise, is also common practice to ensure performance. However, extraneous features can still be manifest in NIR spectra from other sources of measurement variance, including ambient conditions and sample presentation parameters. For example, ambient temperature changes can have significant effects on NIR spectra for materials involving hydrogen bonding or containing water. A difference of a few degrees may result in significant spectral changes such as peak intensities and absorbance shifts. Hygroscopic materials are also sensitive to humidity, as the NIR spectrum is known to have broad intense bands related to water absorption. Presentation of the sample to the measurement interface can also introduce variability. The material particle size and diameter of the sample cup can alter the scattering effects on the spectra, while the homogeneity and measured surface area of the sample can also influence the accuracy of the measurement.

In this study, variance of NIR spectra from 41 different bovine skim milk powders and nonfat dry milk powders was explored by use of principal component analysis and varimax rotation methods. Experimental design was controlled in such a way as to either reduce the influence of NIR measurement variance or monitor well-known sources of variance. Resulting spectral data were then interpreted for influential sources of variance by use of principal component score trends and spectral signatures in rotated principal component loadings. Chemical analysis of samples of interest is also reported to support the interpretations of the rotated principal components.

MATERIALS AND METHODS

Milk Powder Samples. A total of 41 milk powders, including 19 skim milk and 22 nonfat dry milk, were acquired from eight suppliers produced between August 2008 and May 2012. Certificates of analysis indicated product origin details (including production sites and lot numbers), and processing conditions (condensing temperatures labeled as high, medium, and low heat). Proximate chemical composition was also indicated on the certificates including levels of moisture (%), fat (%), and protein (%). A detailed summary of all milk powders studied and their supplied attributes and properties is provided in Table 1.

NIR Spectral Measurement. Fourier transform (FT) near-infrared spectra were acquired at the U.S. Food and Drug Administration, Center for Food Safety and Applied Nutrition, Division of Food Processing Science and Technology, with a PerkinElmer Frontier FT-NIR system (Waltham, MA) fitted with the NIRA reflectance accessory (diffuse reflectance). A 12 mm diameter spot was illuminated on the sampling interface, while the spinning cup feature of the reflectance accessory was enabled during acquisition. Each resulting percent reflectance (% R) NIR spectrum was an average of 32 scans at 4 cm^{-1} resolution, over a spectral range between 1000 and 2500 nm (4000 and 10 000 cm^{-1}).

Instrument performance was internally verified daily by vendor-specific tests in transmittance (T) mode, including the "abscissa check" (wavelength accuracy) and the "ordinate check" (photometric response); both checks used an internal polystyrene standard for comparison against spectra acquired at calibration. Photometric noise was also verified daily to be within specification by use of the "noise check", which calculated root-mean-square noise (RMS, % T), peak-to-peak noise (% T), and baseline trending over a specified range.

A background scan (99% Spectralon diffuse reflectance standard) was acquired at the beginning of the experiment, per software requirements, and all automatic prompts for additional background scans were disabled for the remainder of the experiment. However, extra reflectance standard measurements were incorporated into the experimental design as independent samples. Spectra of a USP NIR suitability reference standard (USP, catalog no. 1457844, lot no. G0K264, Rockville, MD) and a 99% Spectralon diffuse reflectance standard (Labsphere, catalog no. AS-01160-060, North Sutton, NH) were acquired at specified intervals on each day of analysis. These spectra were used to monitor the drift in wavelength accuracy and correct for drift in photometric intensity, independent of the system's internal requirements. A tolerance for agreement for wavelength accuracy per USP general chapter ⟨1119⟩[9] is ±1 nm for peaks between 70 and 2000 nm and ±1.5 nm for peaks between 2000 and 2500 nm. Four wavelength peaks were measured across the spectral range (1261.1, 1536.2, 1971.2, and 2313.1 nm), and wavelength peak maxima were determined via a center of gravity algorithm.[10]

Sample Analysis. NIR spectra of six subsamples for each milk powder sample were acquired in a randomized order on three consecutive days of analysis with two subsamples per milk powder being acquired on each day of analysis. Stock samples were stored in sealed glass jars and remixed by multiple inversions between subsampling. For each milk powder subsample, a 1.0−1.5 cm thick (about 25 g) portion was evenly distributed into a 100-mm dish (PerkinElmer, catalog no. L1181257, Oakbrook, IL) by gently swirling, taking care not to impact any surfaces, so as to not alter the natural particle size distribution. The dish was placed on the sampling interface of the reflectance accessory and covered for each measurement. Since the same dish was used for each subsample, it was thoroughly cleaned between measurements by pouring out the milk powder and removing the excess particles with a vacuum and Kimwipe tissue.

In addition, on each day of analysis, six replicate measurements were acquired for randomly selected subsamples with the sample remaining on the sampling interface between replicates. While variance contributions from subsampling and instrumental repeatability are expected to be relatively small in comparison to the variance associated with the chemical and physical differences between samples, it is nonetheless necessary to characterize these contributions. Reflectance standards were acquired at intervals of every six milk powder subsamples. Spectral acquisition included 246 unique subsample spectra, 105 additional replicate spectra (not included in exploratory principal component analysis described in the following section), and 84 reference standard spectra, for a total of 435 spectra.

Principal Component Analysis and Varimax Rotation. Principal component analysis (PCA) was used to explore the variance in the repeatability of the NIR measurement, milk powder subsamples, and NIR spectra of the 41 commercial milk powder samples (MATLAB 2012, The Mathworks Inc., Natick, MA, and PLS_toolbox 5.2, Eigenvector Research Inc., Wenatchee, WA). PCA is an exploratory chemometrics method that aims to reduce the dimensionality of data from a large number of original measurements (e.g., 6000 variables in an NIR spectrum) to a small number of principal components (typically, the first 3−5 components), with the remaining, higher-order components typically reflecting measurement noise. The reduction is calculated such that each principal component

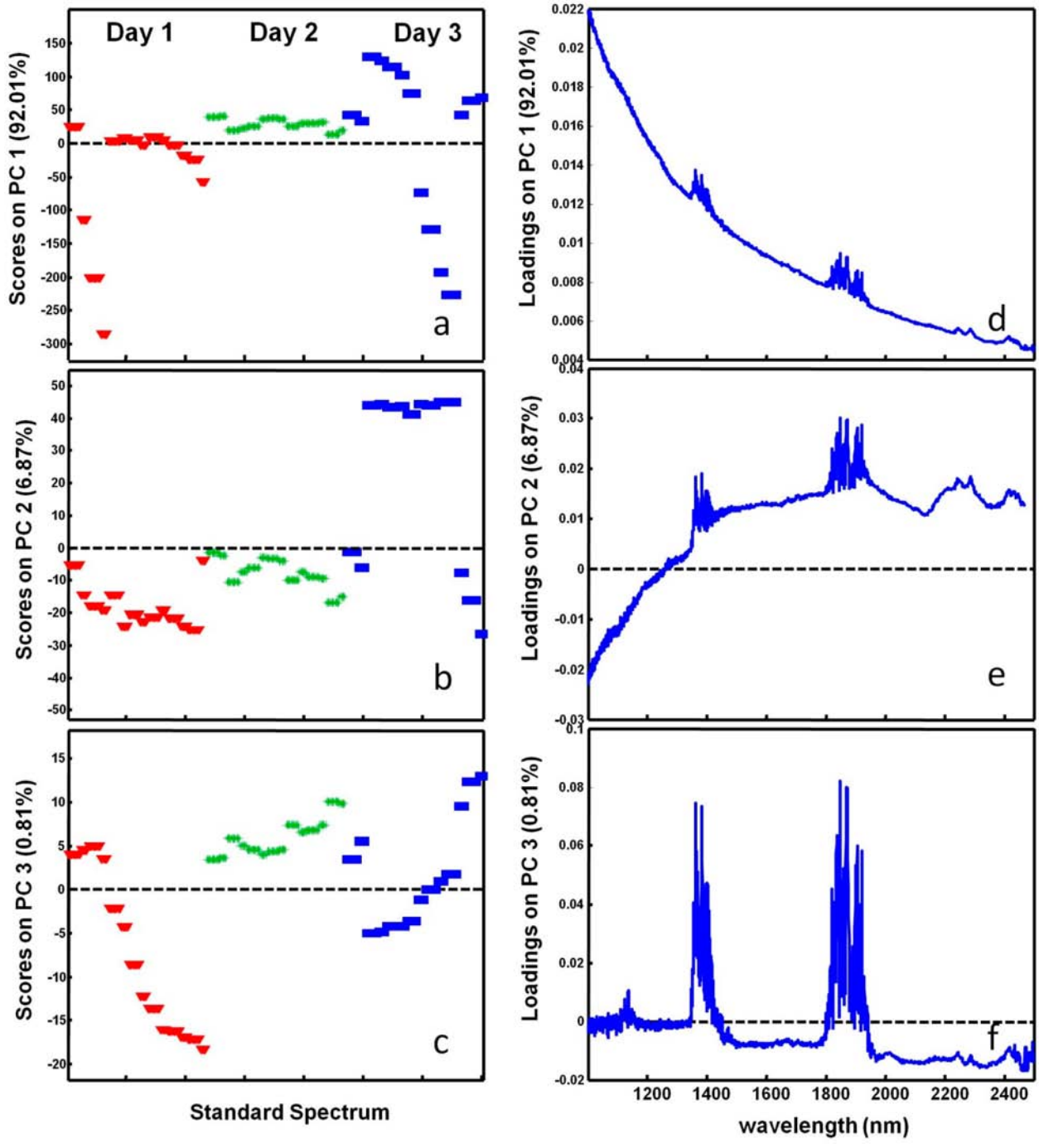

Figure 1. Principal component analysis of 99% reflectance standard acquired on three consecutive days of analysis (N = 42). Scores plots of (a) PC 1, (b) PC2, and (c) PC3 and their respective loadings plots (in panels d–f) are shown.

(PC) is orthogonal to its preceding component and explains the largest percentage of the total variance in the remaining data set. For example, the first PC accounts for the largest percentage of total variance; the second PC explains the largest percentage of the remaining variance, and so on. Principal components can be expressed as a linear combination of the original spectral variables, where each variable is weighted on the basis of its variance contribution for that PC and can be plotted graphically as the variable loading plot. Similarly, each sample can be projected onto each PC loading and can be plotted by its score or projection onto the principal component.

Varimax rotation, an orthogonal rotation method, is used to rotate the principal components so that groups of variables will load onto a single rotated component instead of being distributed across several principal components. The rotated component is referred to as a factor and may correspond to a factor in the experimental design or property of the data; this may aid the spectral interpretation to chemical or physical sources of variance.[11] The interpretation is simplified because, after a varimax rotation, original variables that contribute variance in multiple relevant (or retained) PCs tend to be expressed in a single rotated component. Generally, the varimax solution means that each component has a small number of heavily weighted spectral variables and a large number of insignificant spectral variables.

Exploratory Chemical Analysis. Six of the aforementioned milk powders (highlighted in Table 1) were selected for further characterization and chemical analysis by high-pressure size-exclusion chromatography (HPSEC) for estimation of denatured protein. The

dx.doi.org/10.1021/jf4023433 | J. Agric. Food Chem. 2013, 61, 9810–9818

HPSEC method utilized a Shodex protein column KW-803 (8 × 300 mm, maintained at 25 °C), with a mobile phase of 0.05 M NaH_2PO_4 and 0.15 M NaCl at pH 7.0 (flow rate = 0.3 mL/min). Separated analytes were detected at 214 nm with a total run time of 75 min. Samples were also analyzed for levels of *ε-N*-(furoylmethyl}-L-lysine (furosine), as an early-stage marker for Maillard browning, by liquid chromatography (LC-UV)[12] following acid hydrolysis.

RESULTS AND DISCUSSION

Wavelength Accuracy. Spectral peak positions of the USP NIR system suitability standard were determined by use of a custom-written center of gravity script in MATLAB. Deviations from the expected wavelength positions of 1261.1, 1536.2, 1971.2, and 2313.1 nm, as provided by the USP system suitability standard certificate at a bandwidth of ±2 nm, were calculated. All peaks were demonstrated to be within tolerance of ±1 nm for peaks below 2000 nm and ±1.5 nm for peaks above 2000 nm, and no distinct trends between days of analysis were observed.

Photometric Intensity. Deviations in spectral profiles of the 99% reflectance standard were observed over the consecutive days of analysis. Trends of this spectral variance were explored by PCA, where the data were mean-centered prior to analysis. Figure 1a−c contains reflectance standard scores for PCs 1−3 plotted against their sequential acquisition in time. Respective loading plots of these PCs (Figure 1d−f) explain 99.7% of the spectral variance and demonstrate contributions from a sloping baseline and broad spectral features centered at ∼1400 and ∼1930 nm. While the cause of the baseline slope is uncertain, the latter features are typical of ambient moisture, which has known absorbance bands in those regions. Principal components 4 and 5, which accounted for less than 0.3% of the total variance, possessed some features between 2200 and 2400 nm (data not shown). These absorbance features can be attributed to artifacts present on the standard or sampling interfaces and were only observed for three of the 99% reflectance standard measurements acquired on day 1.

Reflectance standard measurements are typically used in calculating double-beam "pseudo-absorbance" spectra and are intended to correct for instrumental drift and ambient variance contributions in sample spectra. PCA of single-beam milk powder spectra (N = 351, %R mean-centered spectra) showed clear moisture band contributions in PC5−PC7 loading plots (i.e., greater than 0.1% variance), similar to those observed in the reference standard. Baseline sloping effects were also observed in three of the four first principal components calculated from this data set. As a result, milk powder absorbance spectra (A) were calculated by use of eq 1, where the reflectance standard spectrum (R_{RS}) used for the correction was that which was acquired just prior to the reflectance milk powder spectrum (R_{milk}).

$$A = -\log_{10}\left(\frac{R_{milk}}{R_{RS}}\right) \quad (1)$$

While the relationship between absorbance and diffuse reflectance is not accurately defined by eq 1 (for reasons not discussed in this report),[9,13] for the purposes of this application, the estimate or "pseudo-absorbance" will be considered sufficient. After this conversion, principal component contributions of ambient moisture bands and sloping baselines were no longer observed in any of the first seven PCs of the milk powder absorbance spectra.

Note, subsample measurements (N = 18) that were corrected with the three outlier 99% reflectance standard spectra (as described previously) had also exhibited extraneous features between 2200 and 2400 nm, which were not present before the double-beam absorbance calculation. These resulting subsample absorbance spectra were removed from all subsequent data analyses.

Milk Powder Variance by Chemometrics. *Preprocessing.* Resulting "pseudoabsorbance" spectra (6001 variables/spectrum) were further corrected by use of standard preprocessing algorithms applied to NIR spectral data, including standard normal variate (SNV) correction and first-derivative transformation with a Savitzky−Golay algorithm (window size = 35 points, third-order polynomial fit). End points of all spectra were subsequently removed from the spectral data set (20 points from both higher and lower wavelength ends). Preprocessing methods employed are used to correct for any potential physical phenomena or interferences that result in unwanted signal variability that may not be corrected by instrument calibration methods. For example, diffuse reflectance spectra of powdered samples often contain effects due to light scatter from particles within the sample; these effects are manifest as a multiplicative interference across the NIR spectrum. The magnitude of the multiplicative scatter is a function of particle size and its distribution. Typical preprocessing techniques used to correct this include multiplicative scatter correction (MSC) or SNV transformation. SNV generally provides the same results as the more commonly used MSC method, without the need for a reference spectrum. For each spectrum, the mean value of all variables (e.g., absorbance values) is subtracted from each variable. Each mean-subtracted variable is then divided by the standard deviation of all variables for that spectrum.

Particle size can also influence the spectral path length (or light beam penetration) as a result of variations in sample packing, bulk density, and sample thickness; this is manifested as a constant background in the NIR spectrum. Derivatives are often used to reduce this effect, where the background of first-derivative spectra is converted to a constant level, correcting constant baseline offsets. The additional benefit of derivative preprocessing is its ability to emphasize small shoulders and peaks so that the resulting spectra have more pronounced features. These attributes may be useful when targeting small changes in intensity. Savitzky−Golay convolutions are often used to calculate derivative spectra,[14] where at each variable in a spectrum, a polynomial of specified order is fit to the number of points (window) surrounding the variable. An estimate for the value of that variable is calculated from the derivative of the fitted function. The algorithm moves to the next point along the spectrum and performs the same calculation with the same window size and polynomial order. Since fewer data points are fitted near the end-points of a spectrum, the approximation of the polynomial fit and subsequent derivative can introduce unusual features in this region, and are often removed from the spectral data set. However, the challenge of applying a derivative is the interpretation of the resulting spectrum, because peaks and features are no longer visually intuitive. It is helpful to remember that first-derivative spectra have peaks at regions of maximum slope in the original spectrum and cross the zero line at locations of peak maxima/minima in the original spectrum.

The additional advantage of using these preprocessing methods is that both SNV and first-derivative transformation

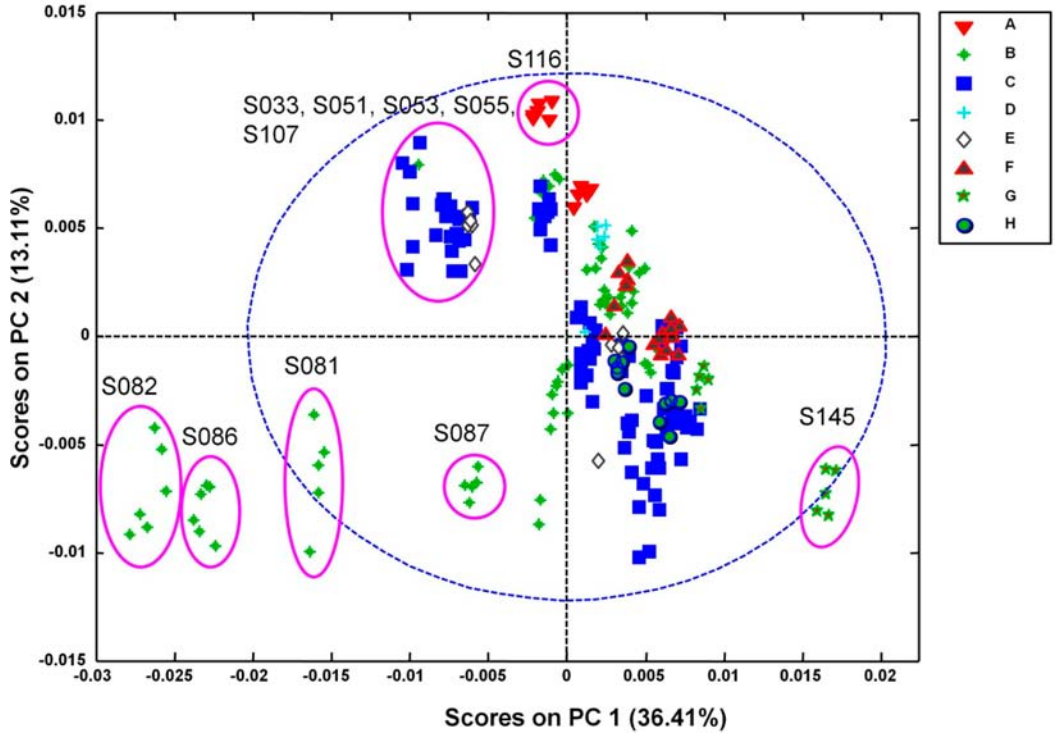

Figure 2. Score plot of PC1 versus PC2 from PCA of NIR spectra of 41 varying milk powders from eight different suppliers (A–H) and their subsample measurements (total = 228 spectra). Spectra were preprocessed by use of standard normal variate (SNV) correction and first-derivative transformation using a Savitzky–Golay algorithm (window size = 35 points, third-order polynomial fit).

were shown to be effective in removing variability introduced between replicate and subsample measurements of the same milk powder material (as evaluated by PCA), indicating that the major source of variance between replicate measurements is from light-scattering and path-length effects, while minor sources were attributed to random noise contributions. The NIR spectra are also mean-centered so that absolute magnitudes are not considered in a multivariate analysis but only relative changes to the mean. This preprocessing step is often used prior to PCA.

PCA and Varimax Rotation. Principal component analysis was applied to the 228 preprocessed milk powder spectra from 41 unique milk powder samples with either five or six subsamples each (18 subsamples had been previously removed from the data set; see description under Photometric Intensity). Score plots were explored for unique clustering patterns for various classification categories, including day of analysis, SMP versus NFDM, supplier, and condensing temperatures (high, medium, and low heat). No clear trends were observed in many of these categories, except for the resolved clustering of particular samples: S081, S082, S086, and S145 along PC 1 and S116 along PC 2 (Figure 2). Interestingly, samples S081, S082, and S086 were manufactured by the same supplier, while S145 exhibited a lower moisture content than the majority of the milk powder samples (mean ± std = 3.61% ± 0.47%, S0145 = 1.80%). An additional cluster of samples, S033, S051, S053, S055, and S107, was observed in covariance of PC1 and PC2; however, no single sample property could be attributed to this cluster, even though the majority of these samples were low heat processed samples.

Five principal components were retained from the PCA, capturing 60.60% of the total variance, and were rotated with the varimax factor rotation algorithm (Figure 3a–e). Interpretations of the rotated components revealed features related to chemical sources of variance, including water and R–OH combination band contributions for PC 1 (1450, 1940 nm), a distinguishing lactose spectrum for PC 2, other sugar contributions for PC 3, lipids (fats) and protein contributions for PC 4, and additional C–H combination band contributions in PC 5.[15−18] Few signal contributions from below 1400 nm were observed in these principal components, demonstrating the limited sensitivity in the third overtone region of the NIR spectrum. While some contributions were observed at ∼1400 nm, these small features can generally be attributed to moisture.

On the basis of these interpretations, principal component analysis of targeted spectral regions on the spectral data set (228 spectra) was performed to confirm the chemical sources of variance for the resolved samples in Figure 2. Score and loading plots (Figure 4a–c) from PCA of the NIR spectra between 2200 and 2500 nm, the C–H combination band region, demonstrated significant discrimination of supplier B samples S081, S082, S086, and even S087, based on the covariance structure of PC 1 and PC 2 (not varimax-rotated). Absorbance bands in this region are most likely correlated to lactose, fat, and protein content and are typically used for quantitative determination of these constituents. Score and loading plots from PCA of NIR spectra between 1700 and 2200 nm resolved similar sample clusters as observed for the full spectral window, again emphasizing the major contributions of both moisture (∼1930 nm) and R–OH (∼2000 nm) combination bands in discriminating the same samples, S081, S082, and S086. The spectral band for R–OH stretch (2000 nm) is most likely associated with functional groups in sugars (lactose, etc.) and may also suggest that the source of variance in the spectral bands above 2200 nm is also correlated to this

dx.doi.org/10.1021/jf4023433 | J. Agric. Food Chem. 2013, 61, 9810–9818

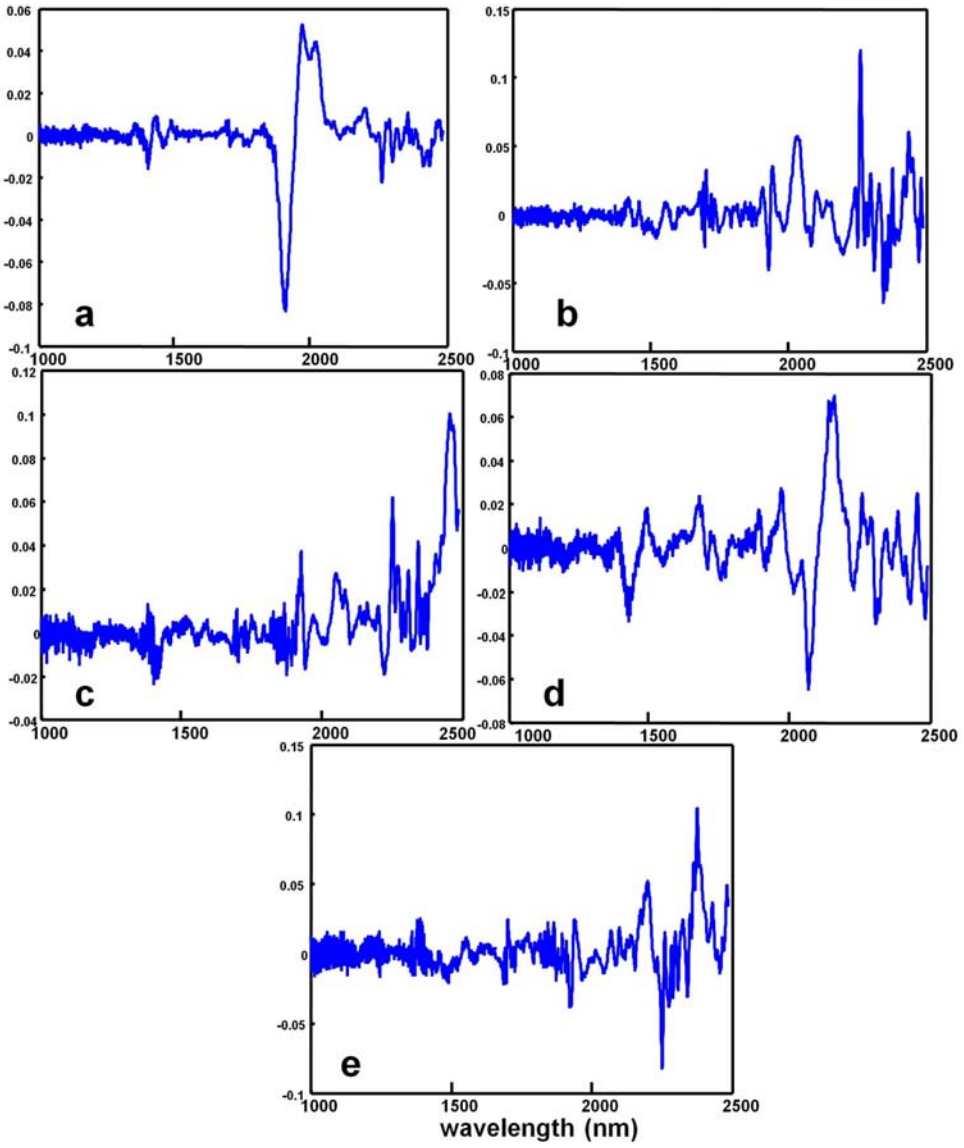

Figure 3. Varimax rotated loading plots of (a) PC 1, (b) PC 2, (c) PC 3, (d) PC 4, and (e) PC 5, from PCA of NIR spectra of 41 varying milk powders from eight different suppliers and their subsample measurements (total = 228 spectra). Spectra were preprocessed by use of standard normal variate (SNV) correction and first-derivative transformation with a Savitzky–Golay algorithm (window size = 35 points, third-order polynomial fit).

chemical source of variance, since similar milk powder samples are discriminated in both regions.

Milk Powder Variance by Exploratory Chemical Analysis. Milk powders S081, S082, S086, S087, S091, and S096 were selected for further characterization by chemical analysis; four of these were discriminated by PCA, while the other two samples clustered near the center of the PCA space (along the first five PCs). Basic compositional analysis showed no difference between these samples for total protein (total nitrogen content), total fat, total lactose, total ash, and total sugars (data not shown). Additional chemical analysis for aggregated protein and furosine levels (Table 2) suggested a correlation with condensing temperature, where a direct relationship was observed between the heat level and aggregated protein, and between the heat level and furosine concentration. Both correlations are theoretically expected since an increase in condensing temperatures can cause changes in the tertiary structure of milk proteins, leading to denaturation and aggregation.[19] The extent of the Maillard reaction can also be catalyzed by heat and an increase in furosine, a byproduct of this reaction, is expected.[20] While these results are expected, they do not support the clustering patterns observed in the PCA space. Additional sources of variance are thought to contribute to the separation of these samples, and further characterization of these milk powders is required. One possibility that should be explored is the presence at low levels of chemical food additives that are authorized in international standards for addition to milk powders.[21] Exploring these and other unknown sources of variance could be investigated by use of targeted assays for specific chemical additives and multivariate approaches by Raman and NMR spectroscopy.

This study has demonstrated that appropriate experimental design and spectral preprocessing can reduce the instrumental and measurement sources of variance in NIR spectra of skim and nonfat dry milk powders, thus providing the basis for a robust compendial method for authentication. However,

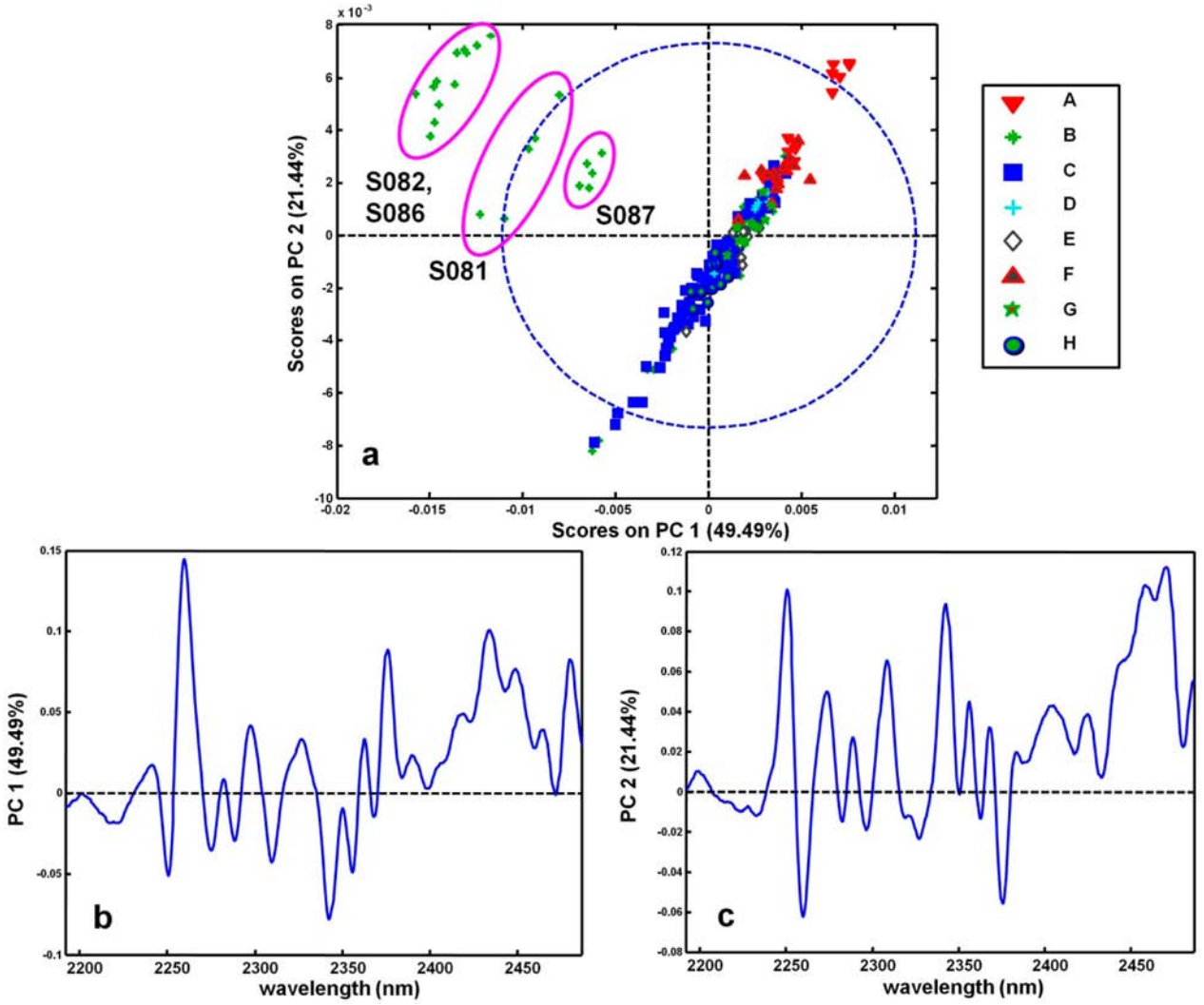

Figure 4. PCA of NIR spectra between 2200 and 2500 nm of 41 varying milk powders from eight different suppliers (A–H) and their subsample measurements (total = 228 spectra). Score plots of (a) PC 2 versus PC 1 and their loading plots for (b) PC 1 and (c) PC 2 are shown. Spectra were preprocessed with standard normal variate (SNV) correction and first-derivative transformation by use of a Savitzky–Golay algorithm (window size = 35 points, third-order polynomial fit).

Table 2. HPSEC Data for Approximation of Protein Aggregation and LC–UV Data for Determination of Furosine of Six Selected Samples from Table 1

sample (process)[a]	aggregated protein (% of total protein)	furosine[b] (mg/100 g)
S081 (HH)	27	242
S082 (LH)	12	163
S086 (HH)	28	215
S087 (LH)	11	152
S091 (MH)	23	137
S096 (MH)	19	105

[a]LH, low heat; MH, medium heat; HH, high heat. [b]Furosine is an early-stage marker for Maillard browning.

defining boundary conditions for classifying authentic milk powder is still challenged by the unknown chemical sources of variance that discriminate between authentic milk powders. In addition, the development of specifications is limited by the number and source of authentic milk powders, as the 41 samples analyzed here do not necessarily represent the population of commercially available milk powders in the United States and other countries. Finally, the sensitivity in detecting adulterants present in samples is still unknown; potentially broad specifications may reduce the capability of such methods to detect any low-level adulterants present in skim and nonfat dry milk powder.

AUTHOR INFORMATION

Corresponding Author

*Phone +1 301 230 7421; fax +1 301 816 8142; e-mail lyb@usp.org.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We acknowledge Paden Skinner of CFSAN, FDA, and Paul Johns of Abbott Nutrition for their technical assistance in acquiring spectral and chemical data in the study. We also thank Stefan Schuber of USP for editorial assistance.

ABBREVIATIONS

A, absorbance; FT, Fourier transform; HPSEC, high-pressure size-exclusion chromatography; LC-UV, liquid chromatography−ultraviolet (detection method); MSC, multiplicative scatter correction; NFDM, nonfat dry milk; NIRA, near-infrared reflectance accessory; NIR, near-infrared spectroscopy; NMR, nuclear magnetic resonance; PCA, principal component analysis; PC, principal component; QA, quality assurance; QC, quality control; R, reflectance; RMS, root-mean-square; SMP, skim milk powder; SNV, standard nominal variate; T, transmittance; USP, U.S. Pharmacopeia

REFERENCES

(1) International Dairy Federation.. The World Dairy Situation 2012. *Bull. Int. Dairy Fed.* **2012**, *458*, 1−235.
(2) Moore, J. C.; DeVries, J. W.; Lipp, M.; Griffiths, J. C.; Abernethy, D. R. Total protein methods and their potential utility to reduce the risk of food protein adulteration. *Compr. Rev. Food Sci. Food Safety* **2010**, *9*, 330−357.
(3) Moore, J. C.; Lipp, M.; Griffiths, J. C. Preventing the adulteration of food protein. *Food Technol.* **2011**, *65* (2), 46−50.
(4) Moore, J. C.; Spink, J.; Lipp, M. Development and application of a database of food ingredient fraud and economically motivated adulteration from 1980−2010. *J. Food Sci.* **2012**, *77* (4), R118−R126.
(5) U.S. Pharmacopeia. U.S. Pharmacopeia's Food Protein Workshop, June 16 and 17, 2009. http://www.usp.org/meetings-courses/workshops/food-protein-workshop. (Accessed Apr 12, 2013).
(6) Moore, J. C.; Ganguly, A.; Smeller, J.; Botros, L.; Mossoba, M.; Bergana, M. Standardisation of non-targeted screening tools to detect adulterations in skim milk powder using NIR spectroscopy and chemometrics. *NIRnews* **2012**, *23*, 9−11.
(7) Frankhuizen, R. NIR analysis of dairy products. In *Handbook of NIR Analysis*, 3rd ed.; Burns, D. A., Ciurczak, E. W., Eds.; CRC Press: Boca Raton, FL, 2008; pp 415−436.
(8) Downey, G.; Robert, P.; Betrand, D.; Kelly, P. M. Classification of commercial skim milk powders according to heat treatment using factorial discriminant analysis of near-infrared reflectance spectra. *Appl. Spectrosc.* **1990**, *44* (1), 150−155.
(9) U.S. Pharmacopiea. Near-Infrared Spectroscopy ⟨1119⟩. In *USP 35−NF 30*; U.S. Pharmacopeia: Rockville, MD, 2013; pp 714−720.
(10) Cameron, D. G.; Kauppinen, J. K.; Moffatt, D. J.; Mantsch, H. H. Precision in condensed phase vibrational spectroscopy. *Appl. Spectrosc.* **1982**, *36* (3), 245−250.
(11) Abdi, H. Factor rotations in factor analysis. *Encycl. Soc. Sci. Res. Methods* **2003**, 792−795.
(12) Ewen, J.; McKenna, R.; O'Kane, K.; Phillips, R.; Johns, P. Effect of carbohydrate DE on blocked lysine and furosine in a liquid nutritional product. *Food Chem.* **2010**, *119*, 323−327.
(13) Griffiths, P. R.; Dahm, D. J. Continuum and discontinuum theories of diffuse reflectation. In *Handbook of NIR Analysis*, 3rd ed.; Burns, D. A., Ciurczak, E. W., Eds.; CRC Press: Boca Raton, FL, 2008; pp 22−64.
(14) Savitzky, A.; Golay, M. J. E. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **1964**, *36* (8), 1627−1639.
(15) Workman, J.; Weyer, L. *Practical Guide and Spectral Atlas for Interpretive Near-Infrared Spectroscopy*, 2nd Ed.; Taylor & Francis: New York, NY, 2012.
(16) O'Neil, A.; Edwards, H. Spectroscopic characterization. In *Solid State Characterization of Pharmaceuticals*; Story, R. A., Ymen, I., Eds.; John Wiley & Sons: New York, 2011.
(17) Sasic, S.; Ozaki, Y. Band assignment of near-infrared spectra of milk by use of partial least-squares regression. *Appl. Spectrosc.* **2000**, *54* (9), 1327−1338.
(18) Rodriguez-Otero, J. L.; Hermida, M.; Centeno, J. Analysis of dairy products by near-infrared spectroscopy: a review. *J. Agric. Food Chem.* **1997**, *45* (8), 2815−2819.
(19) Anema, S. G.; McKenna, A. B. Reaction kinetics of thermal denaturation of whey proteins in heat reconstituted whole milk. *J. Agric. Food Chem.* **1996**, *44*, 422−428.
(20) Vallejo-Cordoba, B.; Mazorra-Manzano, M.; Gonzalez-Cordova, A. F. New capillary electrophoresis method for the determination of furosine in dairy products. *J. Agric. Food Chem.* **2004**, *52*, 5787−5790.
(21) WHO−FAO. Codex Standard 207−1999, Codex standard for milk powders and cream powder. *Codex Alimentarius* 1999. www.codexalimentarius.org/input/download/standards/333/CXS_207e.pdf (Accessed Apr 15, 2013).