JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Protein and Oil Composition Predictions of Single Soybeans by Transmission Raman Spectroscopy

Matthew V. Schulmerich,[†,‡,§] Michael J. Walsh,[†,‡] Matthew K. Gelber,[†,‡] Rong Kong,[†,‡] Matthew R. Kole,[†,‡] Sandra K. Harrison,[∥] John McKinney,[∥] Dennis Thompson,[∥] Linda S. Kull,[⊥,▽] and Rohit Bhargava*[,†,‡,§,#]

[†]The Beckman Institute for Advanced Science and Engineering, University of Illinois at Urbana−Champaign, 405 North Mathews Avenue, Urbana, Illinois 61801, United States

[‡]Department of Bioengineering, 1304 West Springfield Avenue, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States

[§]Micro and Nanotechnology Laboratory, 208 North Wright Street, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States

[∥]Illinois Crop Improvement Association, 3105 Research Road, Champaign, Illinois 61822, United States

[⊥]National Soybean Research Laboratory, University of Illinois at Urbana−Champaign, 1101 West Peabody Drive, Urbana, Illinois 61801, United States

[▽]College of Agriculture, Consumer, and Environmental Sciences, University of Illinois at Urbana−Champaign, 1304 West Pennsylvania Avenue, Urbana, Illinois 61801, United States

[#]Department of Electrical and Computer Engineering, Department of Mechanical Science and Engineering, and University of Illinois Cancer Center, University of Illinois at Urbana−Champaign, 1206 West Green Street, Urbana, Illinois 61801-2906, United States

ABSTRACT: The soybean industry requires rapid, accurate, and precise technologies for the analyses of seed/grain constituents. While the current gold standard for nondestructive quantification of economically and nutritionally important soybean components is near-infrared spectroscopy (NIRS), emerging technology may provide viable alternatives and lead to next generation instrumentation for grain compositional analysis. In principle, Raman spectroscopy provides the necessary chemical information to generate models for predicting the concentration of soybean constituents. In this communication, we explore the use of transmission Raman spectroscopy (TRS) for nondestructive soybean measurements. We show that TRS uses the light scattering properties of soybeans to effectively homogenize the heterogeneous bulk of a soybean for representative sampling. Working with over 1000 individual intact soybean seeds, we developed a simple partial least-squares model for predicting oil and protein content nondestructively. We find TRS to have a root-mean-standard error of prediction (RMSEP) of 0.89% for oil measurements and 0.92% for protein measurements. In both calibration and validation sets, the predicative capabilities of the model were similar to the error in the reference methods.

KEYWORDS: *soybean, nondestructive analysis, protein, oil, Raman spectroscopy, transmission, light scattering*

INTRODUCTION

A reliable measure of economically important soybean grain components is important to the soybean industry. Soybean producers, grain elevators, co-ops, processors, and distributors desire reliable grain composition data for a variety of purposes. In commercial trade, uncertainty in the analysis of a grain component can cause major losses to grain elevators in recouping premiums paid for that component. Discrepancies between grain buyers and sellers can bring export shipments to a halt or require major concessions by one or both parties, potentially blemishing future opportunities. Commodity merchandisers and grain processors depend on valid analyses of product value. Additionally, soybean breeders attempting to develop and improve grain quality require reliable information about grain composition. Laboratories providing grain industry services and maintaining laboratory accreditation rely on the latest, industry-approved technologies to ensure quality, precision, and accuracy.

Researchers are continually striving for more precision, accuracy, and breadth in scope for the prediction of grain components. Therefore, grain analytics remains an active area of research. Emerging science and technology may provide viable alternatives and lead to next generation instrumentation for grain compositional analysis. One such technology is transmission Raman spectroscopy (TRS). Raman spectroscopy uses only one wavelength of light to generate a signal response. This wavelength can be chosen from the NIR spectrum and thus can maintain the necessary penetration of light that is achieved with NIR spectroscopy. During Raman scattering, light traveling through the whole grain transfers energy to specific, fundamental molecular vibrations. As a result, some of

Received: March 22, 2012
Revised: July 2, 2012
Accepted: July 3, 2012
Published: July 3, 2012

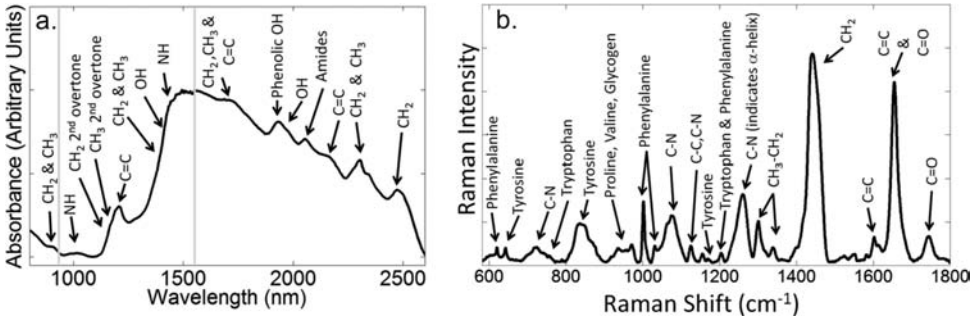

Figure 1. Soybean: (a) Representative NIR spectrum with assigned chemistry. Gray breaks indicate regions collected employing different detectors. (b) Representative Raman spectrum with assigned chemical functional groups.

the light emitted from the grain is at a lower energy than the light that was incident on the whole grain sample. These measurable differences in energy arise as a result of specific chemical groups in the soybean. The biochemical information that is provided through a single Raman spectrum includes distinct vibrational bands originating from proteins, amino acids, lipids, fatty acids, and mineral.[1−3]

TRS is an expanding field that can provide quantitative analysis of a sample's chemistry. The method is providing solutions in both pharmaceutical quality assurances[4−6] and biomedical sciences.[7,8] The pass-through geometry, as opposed to a backscattered configuration, maximizes the illuminated sample volume and ensures that the largest contribution of Raman photons originates from the bulk of the sample, thus rejecting fluorescence that might arise from the sample surface.[9−11]

The need for nondestructive grain analysis in the soybean industry is currently met by near-infrared (NIRS) spectroscopy, which provides nondestructive, quick, and easy grain analysis. NIRS is a widely practiced and well accepted analytical technique that has undergone extensive development.[12] Figure 1 illustrates a NIR and a Raman spectrum of a soybean. In Figure 1a, the primary absorption bands are highlighted and labeled with arrows.[13,14] The *x*-axis represents the spectral wavelength of light detected, and the *y*-axis illustrates a measure of light absorbance. In Figure 1b, the *x*-axis represents the Raman shift from a 785 nm excitation frequency, and the *y*-axis represents photon counts illustrating the relative contribution of signal from specific functional groups present in the soybean. The Raman spectrum clearly shows narrower spectral bands that can be directly assigned to specific chemistry (functional groups) including bands correlated with specific amino acids.[15] High chemical specificity is important in grain analytics, and therefore, Raman spectroscopy could have much potential in quantifying economically important grain components.

The Raman spectral response is a result of the concentration of different types of constituent chemical bonds. The resulting spectrum of a given sample arises as a linear combination of the chemical components in accordance with Beer's law. Because the spectral response is linear with analyte concentration, it is possible to develop a calibration (mathematical model) to predict the concentration of an analyte; however, this mathematical model must first be developed using a primary/reference analytical method for the range of concentrations that are of interest for a particular sample. The mathematical model used is usually some form of a linear regression (PCR or PLS); however, neural network algorithms also have been used.[16−18] There are only a few examples[19−22] of using a Raman based approach for seeds/grain analysis and no reports to apply the methodology to whole intact soybeans. Despite the potential for higher chemical specificity and the availability of data analysis methods, the use of Raman spectroscopy for seed/grain analysis has not been widespread, largely due to a relatively weak Raman signal. However, advances in lasers, filters, spectrographs, and detectors are making applications such as seed analysis more practical. In this communication we illustrate the use of a standard PLS calibration approach to explore the potential of using TRS in nondestructively predicting the compositional analysis of protein and oil contained in intact single soybeans.

MATERIALS AND METHODS

Soybean Samples. Soybean samples anticipated to span the range of potential values for oil and protein were collected and analyzed using the Foss Infratec 1229 NIR spectrometer. Twenty samples best representing the diversity of commercially available soybeans were selected. Factors considered for representative sampling included oil, protein, and moisture concentrations, seed size, seed shape, year grown, seed coat sheen, and hilum color. Protein concentrations of the 20 selected samples ranged from 35.8 to 47.8% dry basis; oil concentrations ranged from 18.3 to 23.8% dry basis. Subsamples of these soybeans were submitted for Raman spectral analysis. Individual soybeans were randomly selected, identified by a unique number, and stored in 48-well tissue culture plates. Each plate held 20 soybeans, one from each of the selected bulk samples. The randomly selected soybeans were inspected to ensure that each was intact and free from debris. Eight soybeans were removed from the sample set after visual inspection (i.e. the soybean was cracked or damaged in some way).

Transmission Raman measurements were collected on 1072 soybeans. These were subdivided into 536 soybeans set aside for oil measurements and 536 for protein measurements. Two of the oil measurements were excluded from the data set due to obvious signs of burning from the laser. This burning was likely the result of dark colored dirt or debris on the seed coat. The remaining set of 1070 soybeans showed no sign of burning or damage. 236 soybeans were used to develop a calibration model for predicting oil content, and 237 soybeans were used to develop a calibration model for predicting protein content. The calibration model was then validated on an independent set of 298 soybeans for oil prediction and 299 soybeans for protein predictions.

After TRS Raman measurements, the samples were reduced to a fine powder and "destructive chemistry" methods were used for quantifying reference values of oil and protein. Individual soybeans were initially crushed to approximately one-half to two-thirds original diameter to rupture the seed coat and breach the cotyledons. Without this initial damage to the soybeans, the time and degree of pulverization was inconsistent. The damage was accomplished by wrapping the seed in weighing paper and squeezing the seed gently between the jaws of household pliers. It was found that a laboratory ore crusher was effective but more difficult to use than the pliers. The

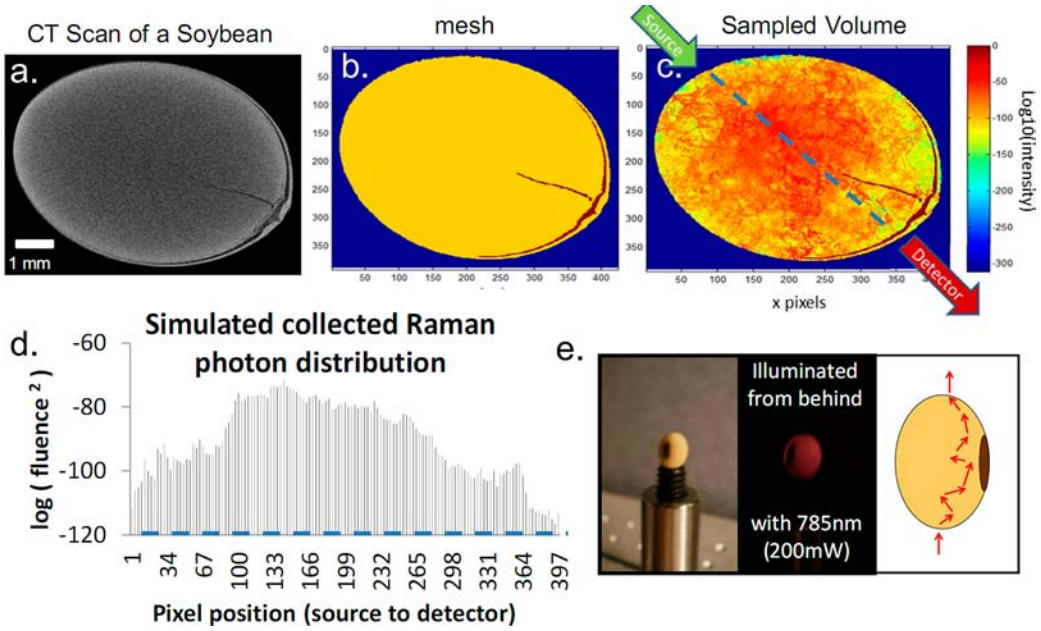

Figure 2. (a) A computed tomography image of a soybean. (b) A computed mesh for modeling the propagation of photons. (c) Random walk generation of Raman photons through the mesh. (d) Modeled Raman photon distribution. (e) Photograph of a soybean illuminated with 200 mW of 785 nm light and schematic of NIR light scattering through the soybean.

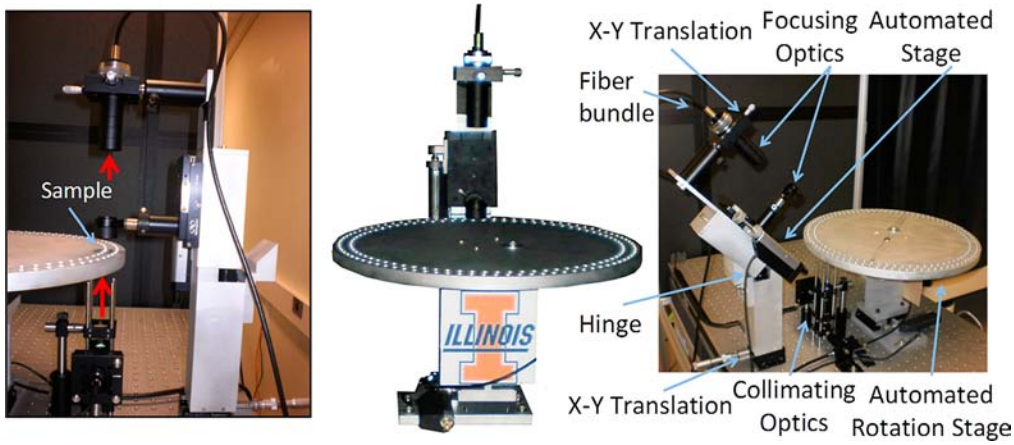

Figure 3. Transmission Raman instrument for single grain soybean analysis. Excitation to collection is highlighted by red arrows.

crushed soybean was then placed in a stainless steel vial (0.5 in. diameter × 1.0 in. long) along with a stainless steel ball (0.25 in. diameter) and pulverized by shaking using a Wig-L-Bug model MSD amalgamator (Dentsply Rinn, Elgin, IL) set at 3800 rpm and 14 s. Ground samples were placed into 2.0 mL microcentrifuge tubes (Fisher Scientific Cat. No. 02-681-258) until analysis could begin. Oil concentration was determined using AOCS Official Procedure Am 5-04 using an Ankom XT15 extractor (Ankom Technology, Macedon, NY) with the entire amount of ground soybean (approximately 100–200 μg) used for the analysis. Petroleum ether was used as the extraction solvent with an extraction time of 60 min. Protein concentrations were determined in accordance with AOCS Official Method Ba 4e-93 using a Thermo Finnigan Flash EA1112 Nitrogen/Protein analyzer (CE Elantech, Lakewood, NJ) and duplicate quantities of approximately 200–300 μg.

Transmission Raman Spectroscopy. Transmission Raman spectroscopy offers an approach for optically homogenizing a soybean by taking advantage of the light scattering properties of the soybean, allowing for the sampling volume to be larger than the spatial heterogeneity of the grain. The ability for TRS to sample from the bulk of the soybean is illustrated in Figure 2. Figure 2a illustrates a computed tomography (CT) image of a soybean showing the internal structure. CT measurements were collected at a 24.1 μm voxel resolution with a SkyScan MicroCT scanner (model 1172, SkyScan, Belgium) operating at 61 kV. A 2D pixel mesh that defines the geometry of the soybean and its internal structures, Figure 2b, was generated based on the CT data. We then modeled the propagation of 785 nm light through the mesh using a random walk Monte Carlo simulation that generated Raman photons along the path the incident photons traveled. The Monte Carlo modeling software,[23] written in-house, simulates the travel of hundreds of thousands of individual photons through the mesh in order to determine the origin of the Raman signal. After optical properties have been assigned, each photon takes a "random walk" through the soybean mesh. Accurate photon behavior at region boundaries follows Snell's Law and the Fresnel equations, while scattering and absorption events follow a Poisson distribution and Beer's Law, respectively. Using the ad-joint source approximation,[24] the sampling region of the detector can be modeled in an analogous fashion. The intersection of the illuminated volume and the sampled volume indicates the origin of the collected Raman signal. Figure 2c illustrates the origin of Raman photons collected from a collimated source at a detector placed in a 180° transmission

dx.doi.org/10.1021/jf301247w | J. Agric. Food Chem. 2012, 60, 8097–8102

geometry. Figure 2d depicts a cross section of the detected Raman photon distribution through a soybean illustrating that the bulk soybean material is well sampled with a higher weighting of Raman photons originating from the middle region of the soybean. Figure 2e illustrates a photograph of a soybean illuminated with 200 mW of 785 nm light focused to a 3 mm spot on the back side of the soybean. The center pane of Figure 2e shows the soybean "glowing", illustrating that the soybean is highly light scattering and that the 785 nm photons are able to penetrate through the entire thickness of the soybean. Based on this transmission approach, we built an instrument to collect Raman spectra from a whole individual soybean. A labeled photograph of the instrument is depicted in Figure 3.

In our experiments, 785 nm (Invictus, Kaiser Optical Systems Inc., Ann Arbor, MI) light was focused from a 100 μm core multimode fiber to a 3.5 mm spot size projected onto the surface of a single intact soybean. Illumination of the soybean occurred from the bottom of the stage as illustrated by the red arrow in Figure 3. The power at the sample was 300 mW. Upon 785 nm light illumination, the photons are preferentially scattered as opposed to being absorbed. Consequently, they propagate through the soybean diffusively and generate Raman photons from multiple locations in the soybean. We position the collection optics (labeled focusing optics in Figure 3) on the opposite side of the soybean so that the illumination and collection are in a 180° transmission geometry. The focusing optics consisted of a 30 mm focal length achromatic lens to collect light emitted from the soybean. A second 60 mm focal length lens was utilized to focus the collected light onto 47 collection fibers (Fiber Tech Optica, Kitchener, Ontario) labeled as "Fiber bundle" in Figure 3. The collection fibers were arranged in a rectangular array at the sample and a linear array at the spectrograph. The spectrograph (RXN-1, Kaiser Optical Systems Inc., Ann Arbor, MI) was optimized for 785 nm excitation and included a prestage notch filter for rejecting Raleigh (elastic) scattered light. For all measurements, the spectral region from 400 to 1800 cm^{-1} was collected with 6−8 cm^{-1} spectral resolution. Soybeans were manually loaded into individual wells of the automated rotation stage. The soybeans naturally fell onto their side, and as a result, laser illumination was not over the hilum of the seed. The instrumentation was combined with sampling and focusing stages controlled by LabView (National Instuments, Austin, TX) to allow for automated data collection once the stage was loaded.

Data Processing. All processing was performed in Matlab2008b (The Mathworks, Nantucket, MA). The instrument's output was a frame from the charged coupled device (CCD) that consisted of 1024 × 255 pixels. The dispersive grating in the spectrograph projected the wavelength axis along the width of the chip (1024 pixels) while the chip height (255 pixels) contained spectra projected from each of the collection fibers. The acquisition time per frame was 60 s with 10 sequential image frames collected per sample. Cosmic rays were removed in an automated manner by comparing sequential frames and applying a mean filter to remove pixels that had a value greater than 8 times the standard deviation of neighboring pixels. Frames from a given sample were then averaged and pixels marked as cosmic rays did not contribute to the average, resulting in a single CCD frame. A CCD frame acquired without a sample present and processed in the same manner was then subtracted from the sample frame to correct for the CCD's dark current. A "pin-cushion" and rotation correction was applied to the sample frame to correct for slit image curvature and a slight rotation of the CCD relative to the spectrograph's slit.[25] The wavelength dependent response of the CCD was corrected by focusing the collection optics onto a NIST traceable white light source emitted from a calibration accessory (HCA, Kaiser Optical Systems Inc., Ann Arbor, MI) and dividing the sample frame by the CCD's white light response. The neon channel of the same calibration accessory was then used to collect the atomic emission spectrum of neon and convert the wavelength axis of the CCD from pixels to wavenumbers. The Raman shift was then calculated with the Raman spectra collected from a standard Teflon sample. Baseline correction was then accomplished by fitting the background signal with a fifth order polynomial and then subtracting that background from the spectra. The median spectrum from each frame was calculated. Each spectrum was then normalized by area and then mean centered. This processing resulted in a single Raman spectrum per soybean which was used as an input for PLS model development.

For smaller data sets, a leave-one-out cross-validation model was generated to predict the percent protein and percent oil from a Raman spectrum using the wet chemistry results and the 20 spectra acquired from each of the soybean varieties. The model was built in Matlab2008b by using 19 of the 20 Raman spectra to generate a PLS regression model with the "plsregress" function. The number of components (latent variables) used to generate the model for cross-validation was typically between three and six components. This was manually determined by observing the minimum for the root-mean-standard error of calibration validation in the leave-one-out cross-validation model. The model was then used to predict the percent protein and percent oil for the left-out soybean measurement. This was repeated for each of the soybeans. The calibration model was evaluated by comparing the predicted protein and oil concentrations to those determined by wet chemistry methods.

For larger data sets it is known that a leave-one-out cross-validation can lead to overfit calibration models.[26] Therefore, we utilized a leave-one-batch-out cross-validation. In this case, all the single soybeans originated from 20 batches. The model was built in Matlab2008b with the PLStoolbox (eigenvector Research Inc., Wenatchee, WA). A custom cross-validation was developed in the software where one batch was left-out to develop the calibration model and then the calibration was utilized to predict the left-out batch. The number of components (latent variables) used to generate the model for cross-validation was typically between 3 and 6 components. This was manually determined by observing the minimum for the root-mean-standard error of calibration validation in the cross-validation model. This was repeated for each of the batches. The calibration model was evaluated by comparing the predicted protein and oil concentrations to those determined by wet chemistry methods.

RESULTS AND DISCUSSION

To determine optimal acquisition times, the transmission Raman instrument was loaded with 40 soybeans and a signal-to-noise comparison was made by coadding multiple frames. The results are illustrated in Figure 4. To quantify the signal-to-noise, we compared the signal of the 1658 cm^{-1} band (C═O and C═C) to the peak-to-peak signal in the 1770−1885 cm^{-1} region where no signal was present. The signal-to-noise maximizes at a 5 min acquisition. This occurs because after 5 min the signal that is contributed by additional time is degraded by the read-out noise of the CCD. Representative Raman spectra acquired at 2, 5, and 10 min are illustrated in Figure 4b. The signal-to-noise is very good for all three measurements; however, it is optimal at a 5 min acquisition time. Minimizing

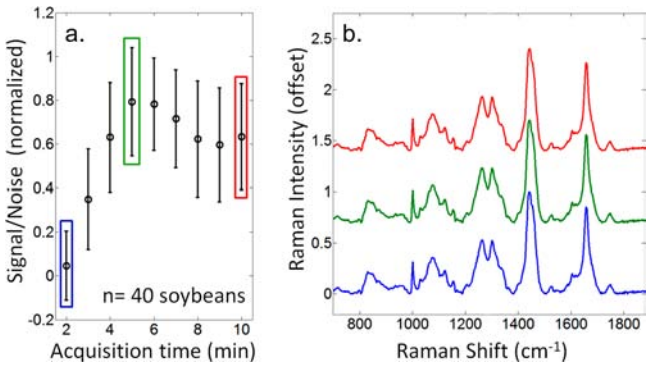

Figure 4. (a) Average signal-to-noise as a function of acquisition time. (b) Representative Raman spectrum collected at 2 min (blue), 5 min (green), and 10 min (red).

dx.doi.org/10.1021/jf301247w | J. Agric. Food Chem. 2012, 60, 8097−8102

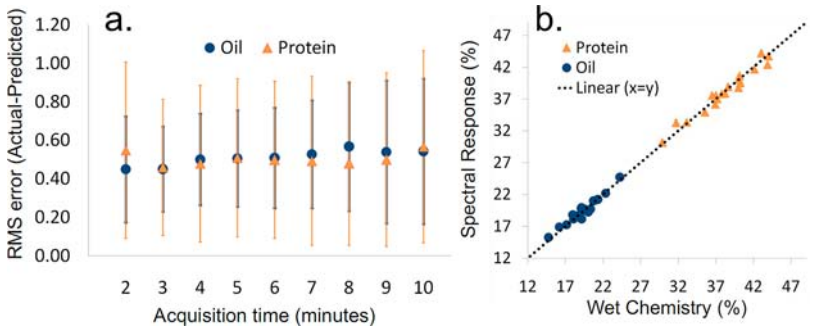

Figure 5. (a) Residuals illustrating the rms error and standard deviation of oil and protein content predicted using transmission Raman spectra of 40 soybeans (20 protein and 20 oil) for different acquisition times. (b) Correlation plot illustrating the Raman predicted values compared to wet chemistry values for a 2 min acquisition time.

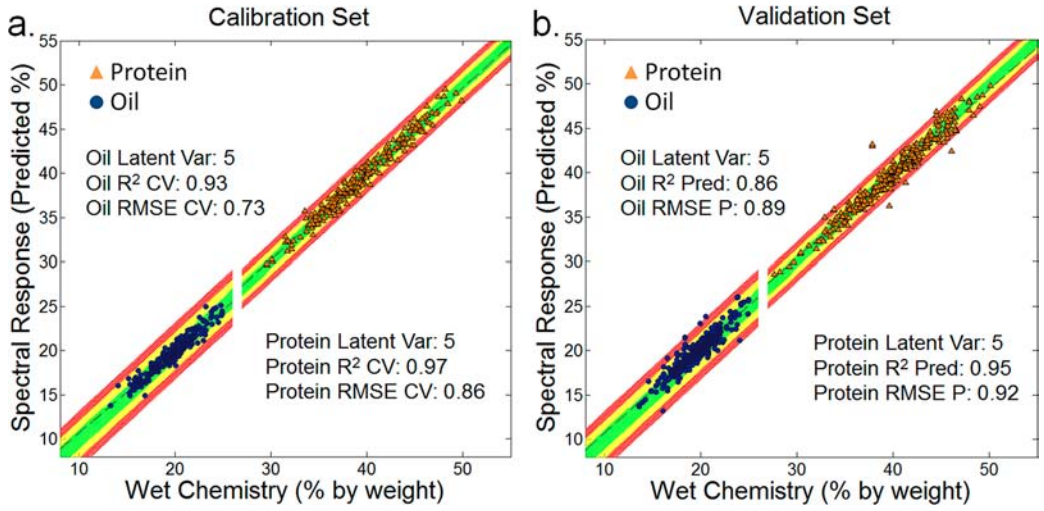

Figure 6. (a) Correlation plot consisting of 236 soybeans for oil calibration and 237 soybeans for protein calibration. (b) Correlation plot consisting of 298 soybeans for oil validation and 299 soybeans for protein validation.

the acquisition time would improve the instrument's throughput. We then used these different acquisition times to generate a leave-one-out cross-validation model in order to determine if the difference in signal-to-noise would affect the predictions.

Next, we compare leave-one-out cross-validation models developed for acquisition times ranging from 2 to 10 min in Figure 5a. Despite the improved signal-to-noise at acquisition times longer than 2 min, we do not observe a significant improvement in the ability to predict the protein or oil content as the acquisition time is increased, possibly due to the multivariate nature of the algorithm. The predicted concentration of protein and oil compared to the concentration obtained by wet chemistry methods is illustrated in Figure 5b. Here, five replicates of 2 min measurements per soybean were used to acquire the spectral data. Error bars for replicates (one standard deviation from the mean) are plotted; however, they are smaller than the marker size. The relative standard error (RMSECV) of the TRS method is less than 1% for both protein and oil, as determined by the leave-one-out cross validation model with a 2 min acquisition time.

The 2 min acquisition time was carried forward to a larger calibration and validation set, though it should be noted that acquisition times as low as 30 s are likely feasible with this instrument. Larger calibration and validation sets are shown in Figure 6. Figure 6a shows the predictions from a model developed using a leave-one-batch-out cross-validation approach. The standard deviation in the wet chemistry for replicate oil measurements on single soybeans was 1.07%, and the standard deviation in wet chemistry for replicate protein measurements was 0.75%. The errors in the reference method are illustrated in Figure 6 as green, yellow, and red lines representing 1, 2, and 3 standard deviations from a perfect correlation. The root-mean-standard error of cross validation for the oil calibration (RMSECV) set was 0.73%, and the root-mean-standard error of prediction (RMSEP) was 0.89%. Prediction capabilities were similar for protein content with a root-mean-standard error of cross validation for the protein calibration (RMSECV) set of 0.86% and the root-mean-standard error of prediction (RMSEP) of 0.92%. In both calibration and validation sets, the predicative capabilities of the model were similar to the error in the reference method.

Here the TRS methodology was applied along with the light scattering properties of soybeans to effectively homogenize the heterogeneous bulk of individual intact soybeans. This approach allows the nondestructive quantification of both protein and oil content in individual soybeans. While NIRS is a more practical approach to quantifying protein and oil content as it is currently slightly cheaper and is faster, the work demonstrated here suggests that there are alternative nondestructive approaches capable of obtaining the same information. Moreover, by using Raman spectroscopy, we gain the benefit of chemically specific vibrational modes which have the potential to improve quantification of additional

economically important grain components, including amino acids, fatty acids, and sugars, among others. Studies to investigate this potential are currently underway in our laboratories.

AUTHOR INFORMATION

Corresponding Author

*E-mail: rxb@illinois.edu.

Funding

We gratefully acknowledge support for this work through the United Soybean Board grant code 9282. Support from the National Science Foundation Industry/University Cooperative Research Center, the Center for Agricultural, Biomedical, and Pharmaceutical Nanotechnology, and CHE 0957849 is gratefully acknowledged.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Penel, G.; Delfosse, C.; Descamps, M.; Leroy, G. Composition of bone and apatitic biomaterials as revealed by intravital Raman microspectroscopy. *Bone* **2005**, *36*, 893−901.

(2) Chrit, L.; Hadjur, C.; Morel, S.; Sockalingum, G.; Lebourdon, G.; Leroy, F.; Manfait, M. In vivo chemical investigation of human skin using a confocal Raman fiber optic microprobe. *J. Biomed. Opt.* **2005**, *10*, 44007.

(3) Overman, S. A.; Thomas, G. J., Jr. Raman markers of nonaromatic side chains in an α-helix assembly: Ala, Asp, Glu, Gly, Ile, Leu, Lys, Ser, and Val residues of phage fd subunits. *Biochemistry (N. Y.)* **1999**, *38*, 4018−4027.

(4) Matousek, P.; Parker, A. W. Non-invasive probing of pharmaceutical capsules using transmission Raman spectroscopy. *J. Raman Spectrosc.* **2007**, *38*, 563−567.

(5) Buckley, K.; Matousek, P. Recent advances in the application of transmission Raman spectroscopy to pharmaceutical analysis. *J. Pharm. Biomed. Anal.* **2011**, *55*, 645−652.

(6) Hargreaves, M. D.; Macleod, N. A.; Smith, M. R.; Andrews, D.; Hammond, S. V.; Matousek, P. Characterisation of transmission Raman spectroscopy for rapid quantitative analysis of intact multi-component pharmaceutical capsules. *J. Pharm. Biomed. Anal.* **2011**, *54*, 463−468.

(7) Schulmerich, M. V.; Cole, J. H.; Dooley, K. A.; Morris, M. D.; Kreider, J. M.; Goldstein, S. A.; Srinivasan, S.; Pogue, B. W. Non-invasive Raman Tomographic Imaging of Canine Cortical Bone Tissue. *J. Biomed. Opt.* **2008**, *13*, 020506.

(8) Stone, N.; Matousek, P. Advanced Transmission Raman Spectroscopy: A Promising Tool for Breast Disease Diagnosis. *Cancer Res.* **2008**, *68*, 4424−4430.

(9) Everall, N.; Matousek, P.; Macleod, N.; Ronayne, K. L.; Clark, I. P. Temporal and spatial resolution in transmission raman spectroscopy. *Appl. Spectrosc.* **2010**, *64*, 52−60.

(10) Everall, N.; Priestnall, I. A. N.; Dallin, P.; Andrews, J.; Lewis, I. A. N.; Davis, K.; Owen, H.; George, M. W. Measurement of spatial resolution and sensitivity in transmission and backscattering raman spectroscopy of opaque samples: Impact on pharmaceutical quality control and raman tomography. *Appl. Spectrosc.* **2010**, *64*, 476−484.

(11) Matousek, P.; Everall, N.; Littlejohn, D.; Nordon, A.; Bloomfield, M. Dependence of signal on depth in transmission Raman spectroscopy. *Appl. Spectrosc.* **2011**, *65*, 724−733.

(12) McClure, W. F. 204 Years of near infrared technology: 1800−2003. *J. Near Infrared Spectrosc.* **2003**, *11*, 487−518.

(13) Wang, D.; Ram, M. S.; Dowell, F. E. Classification of Damaged Soybean Seeds Using Near-Infrared Spectroscopy. *Trans. Am. Soc. Agric. Eng.* **2002**, *45*, 1943−1948.

(14) Westad, F.; Schmidt, A.; Kermit, M. Incorporating chemical band-assignment in near infrared spectroscopy regression models. *J. Near Infrared Spectrosc.* **2008**, *16*, 265−273.

(15) Jenkins, A. L.; Larsen, R. A.; Williams, T. B. Characterization of amino acids using Raman spectroscopy. *Spectrochim. Acta, Part A: Mol. Biomol. Spectrosc.* **2005**, *61*, 1585−1594.

(16) Brereton, R. G. Introduction to multivariate calibration in analytical chemistry. *Analyst* **2000**, *125*, 2125−2154.

(17) Thomas, E. V.; Haaland, D. M. Comparison of multivariate calibration methods for quantitative spectral analysis. *Anal. Chem.* **1990**, *62*, 1091−1099.

(18) Karnes, H. T.; Shiu, G.; Shah, V. P. Validation of bioanalytical methods. *Pharm. Res.* **1991**, *8*, 421−426.

(19) Kim, Y.; Lee, S.; Chung, H.; Choi, H.; Cha, K. Improving Raman spectroscopic differentiation of the geographical origin of rice by simultaneous illumination over a wide sample area. *J. Raman Spectrosc.* **2009**, *40*, 191−196.

(20) Barua, A. G.; Hazarika, S.; Pathak, J. S.; Kalita, C. Spectroscopic investigation of the seeds of chilli (Capsicum annum L.). *Int. J. Food Sci. Nutr.* **2008**, *59*, 671−678.

(21) Liu, D.; Cheng, F. Nondestructive prediction of vigour of maize seeds by Raman spectroscopy. *American Society of Agricultural and Biological Engineers Annual International Meeting 2011*; 2011; Vol. *2*, pp 1012−1018.

(22) Sohn, M.; Himmelsbach, D. S.; Barton, F. E., II A comparative study of fourier transform Raman and NIR spectroscopic methods for assessment of protein and apparent amylose in rice. *Cereal Chem.* **2004**, *81*, 429−433.

(23) Wang, L.; Jacques, S. L.; Zheng, L. MCML - Monte Carlo modeling of light transport in multi-layered tissues. *Comput. Methods Programs Biomed.* **1995**, *47*, 131−146.

(24) Arridge, S.; Schweiger, M. Photon-measurement density functions. Part 2: Finite-element-method calculations. *Appl. Opt.* **1995**, *34*, 8026−8037.

(25) Esmonde-White, F. W. L.; Esmonde-White, K. A.; Morris, M. D. Minor distortions with major consequences: Correcting distortions in imaging spectrographs. *Appl. Spectrosc.* **2011**, *65*, 85−98.

(26) Martens, H. A.; Dardenne, P. Validation and verification of regression in small data sets. *Chemom. Intell. Lab. Syst.* **1998**, *44*, 99−121.