Determination of Moisture in a Protein Sample Using a Portable NIR Instrument

Suzanne K. Schreyer, Ph.D. and Michelle A. Pressler, Ph.D. Thermo Fisher Scientific, Tewksbury, MA

ABSTRACT

Near infrared (NIR) spectroscopy has been used effectively and reliably to measure moisture in a wide variety of samples. Due to the large overtone band for water in NIR, levels of accuracy tend to be high and limits of detection are lower than for most materials measured by NIR. In this feasibility study the performance of a portable NIR instrument was evaluated to monitor moisture levels within a protein sample, as proof of concept to monitor lyophilized samples in a pharmaceutical process. Specifically, a multivariate Partial Least Squares (PLS) model was developed and evaluated for the ability to predict the levels of moisture in samples of casein protein. The results generated with the portable NIR instrument are also evaluated and compared with a laboratory based NIR instrument in order to demonstrate the comparable performance and improved utility of the portable instrument.

INTRODUCTION

Lyophilization is a well-established procedure for the preservation and storage of proteins and other biopharmaceutical materials in both the food and the pharmaceutical industries. Monitoring protein moisture content is an important process as moisture affects both the physical and the chemical state of the materials. In particular, lyophilization of proteins in pharmaceutical processes requires monitoring of these samples as any excess moisture, or moisture lower than optimum, will impact the usability and the shelf life of the samples.

Current methods of monitoring the samples rely on destructive techniques whereby a subset of samples are pulled, opened and analyzed for moisture. There are a number of issues inherent in this type of testing. This method assumes that the selected samples are actually representative of the larger population and renders sampling data vulnerable to potential inconsistency. The method is destructive so samples cannot be used again causing unnecessary loss of product. The process is expensive and time-consuming due to both the materials involved and the specialized training required to perform the relevant protocols.

In contrast the use of NIR technology is fast (seconds), easy to use, and allows for accurate determination of the moisture levels in a closed sample vial. Scanning of the sample can be performed through the glass vial without compromising the material; the vial can then be returned to the sample set as viable product improving operational efficiency. It is possible to perform 100% inspection of samples to comply with regulatory requirements and ensure that anomalous samples will be identified. The deployment of a portable NIR analyzer allows the instrument to be taken directly to the sample source—eliminating the time and protocols required to move samples to the lab.

Thermo Scientific microPHAZIR RX Analyzer can provide raw material identification in seconds.

EXPERIMENTAL DESIGN

Replicate casein samples were made using casein from bovine milk (Sigma Aldrich). Twenty standard samples were placed in glass vials in a saturated NaCl solution within a humidity chamber set to produce 75% humidity. Samples were sealed in the humidity chamber for varying periods of time to produce a range of moisture content. At various time intervals, duplicate samples were taken, mixed thoroughly, and capped to equilibrate moisture within the sample.

Once mixed and equilibrated, samples were scanned on the portable Thermo Scientific™ microPHAZIR™ RX NIR instrument and the lab-based Thermo Scientific Antaris™ NIR instrument. The microPHAZIR RX has effective wavelength range from 1595-2395 nm, at a spectral resolution of (the default) 8 nm. Five sample scans were co-averaged and triplicate scans were taken of the sample. Spectral data was collected in diffuse reflectance mode and evaluated in the proprietary software Method Generator (MG).

Spectra collected on the Antaris FT-NIR had 32 scans co-averaged and triplicate scans were generated with each sample. The wavelength scanned was from 4000cm -10,000cm^{-1} (1000nm -2500nm) at a resolution of 32 cm^{-1}. TQ Analyst was used to evaluate and process the spectra.

Once the spectra were collected, the samples were measured using a Mettler Toledo LOD balance to determine the moisture content. These reference values were then combined with spectral data and PLS models were created to predict moisture in the casein protein. Results from the Antaris were used for comparison to those generated by the microPHAZIR RX; model development will focus solely on the microPHAZIR RX in the following sections.

RESULTS AND DISCUSSION

Data preparation and generation

Spectral data from the microPHAZIR RX was combined with the reference moisture values from the LOD balance for all samples. The moisture values ranged from approximately 9 to 13 wt/wt%.

The collected spectra of the moisture in casein samples are shown in Figure 1. Color coding is by sample number with each sample containing different moisture levels.

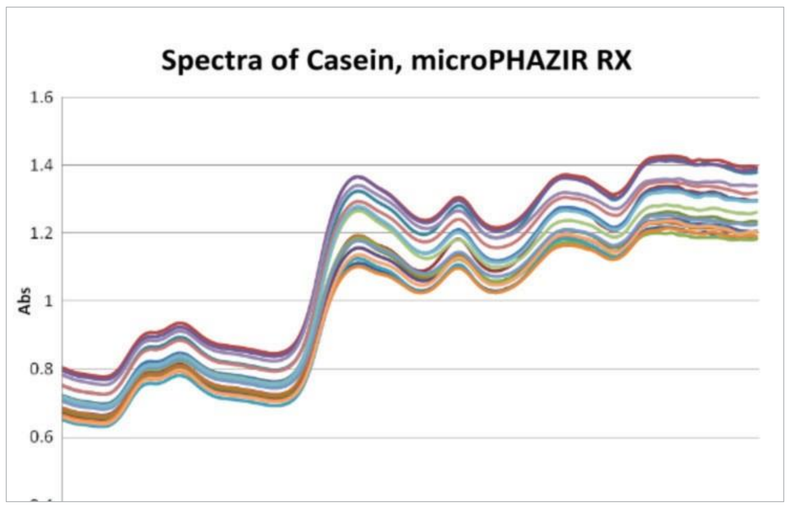

Figure 1. The collected spectra of the moisture in casein samples

Spectra preprocessing was done to optimize the robustness of subsequent models. In this case preprocessing consisted of Standard Normal Variates (SNV) to compensate for packing or density differences, and Savitsky-Golay (1st derivative-5 point smooth-2nd order polynomial). Preprocessing was kept to a minimum due to the strong water absorbance peaks commonly seen in NIR. The wavelength was selected for the strong water overtone region (1800-2100 nm) as shown in Figure 2.

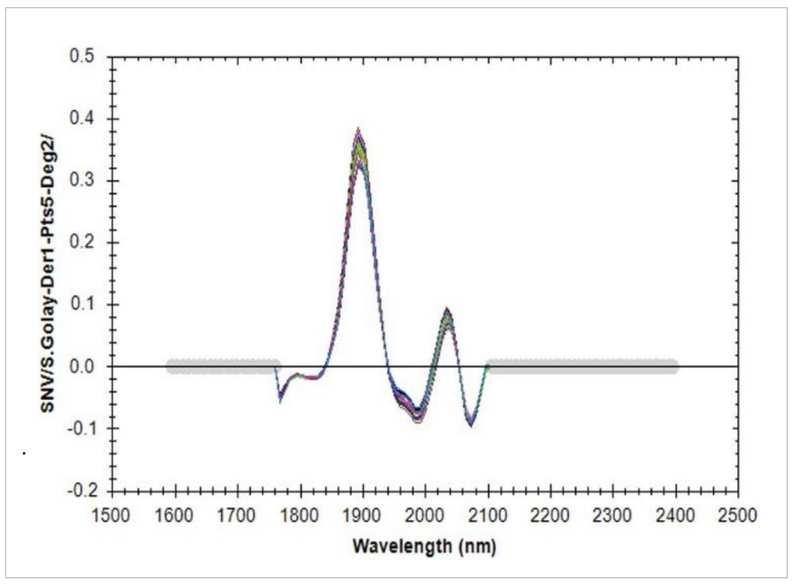

Figure 2. Wavelength range was selected to observe the strong water overtone

Model development

Initially the data set was split into a calibration and a validation data set. A PLS model was used to generate the quantitative model for moisture prediction using the calibration data only. Based on the conditions described above, the PLS model was developed in MG software. The model was then used to predict results on the validation data. Cross validation results on the calibration data were recorded.

Initially the data was evaluated to determine the required number of factors to use for the PLS model. A plot of the number of factors and their contribution to the model (Screen plot, Figure 3) indicates that 3 loadings would provide useful information to the model and limit the noise contribution.

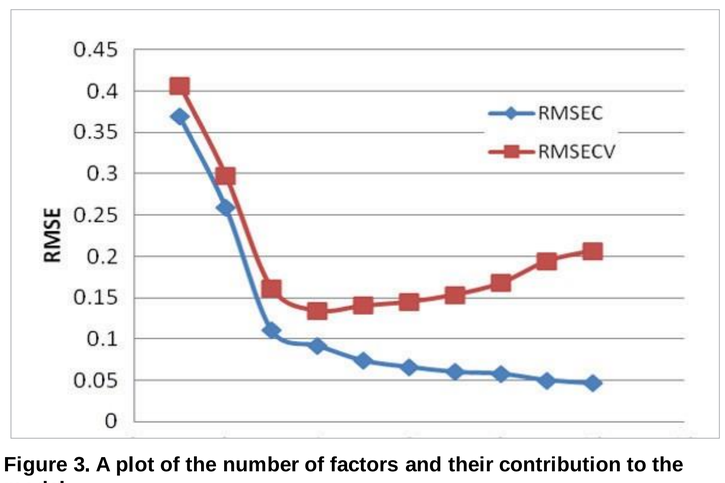

Figure 3. A plot of the number of factors and their contribution to the model

The resulting PLS model was built to utilize these 3 factors. The subsequent correlation plot of reference moisture values and the predicted moisture values is shown in Figure 4. The plot shows the calibration (blue) and the validation data results (red). As clearly seen, both calibration and validation results are linear, indicating good agreement between predicted and actual results. The results from the Antaris data set are also shown for comparison. Again, a high degree of agreement is seen with the predicted and reference values. Both plots also show a similar response to expected model predictive behavior as they correspond to a very similar linear pattern.

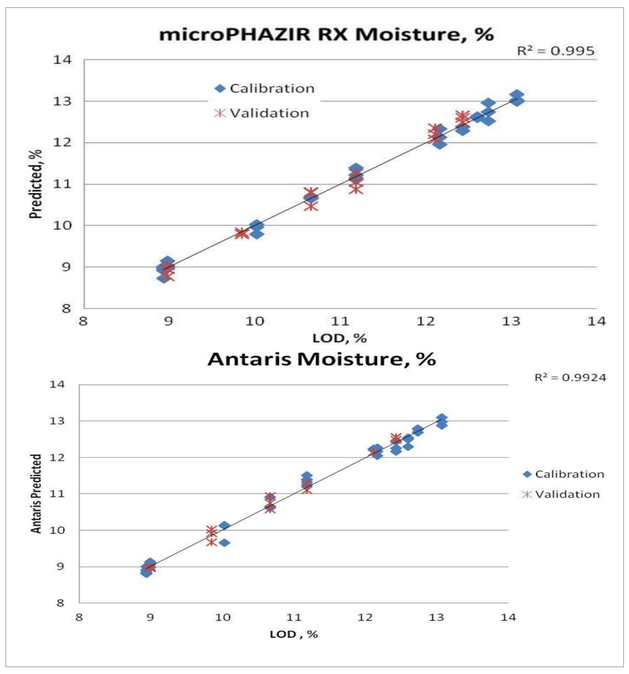

Figure 4. Correlation plots between calibration and validation values for microPHAZIR RX and Antaris, respectively

The similarity between the Antaris and the microPHAZIR RX are further corroborated in the RMSE results. RMSE, which incorporates standard deviation of prediction errors, is a consistent indicator of model performance and robustness, and as Table 2 shows, both instruments are expected to have excellent accuracy predicting moisture in the casein protein samples. Accuracy with respect to the RMSE for calibration and cross validation are similar for the Antaris and the microPHAZIR RX. Note the low and consistent standard deviations indicating both precision in the obtained results and consistency between devices observed.

3 Factor	Antaris	microPHAZIR RX
RMSEC	0.14	0.11
RMSEV	0.12	0.15
RMSECV	0.20	0.16
R2	0.99	1.00
StDev, Residual	0.11	0.11

Table 2. RMSE comparison of Antaris and MicroPHAZIR RX methodologies

CONCLUSION

Predictive models for moisture levels in casein protein were built on a benchtop laboratory and handheld NIR spectrometer. Both systems allowed for the development of an effective, robust PLS model that provided accurate results with consistency. In terms of practical implementation, the slight decrease in accuracy of the handheld system (RMSEV = 0.12 [Antaris] vs. 0.15 [microPHAZIR]) is overcome by the advantages of performing the analysis at the point of sampling, allowing a reduction in testing time and expense and lost product and enabling 100% material inspection.

Based on this feasibility study, it is expected that moisture analysis comparable to those currently obtained via benchtop systems should be readily producible with a handheld platform allowing for rapid and robust monitoring of a substantial range of protein samples. Deployment of the microPHAZIR RX can provide a valuable tool for the monitoring of moisture in raw materials, intermediate production stages, or final products at a degree of convenience that is unmatched by conventional laboratory based systems.

TRADEMARKS/LICENSING