

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 spectroscopy (NIRS) 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 PLS model was developed and evaluated for the ability to predict the levels of moisture in a casein protein. The spectra were collected in diffuse reflectance mode using a portable NIR instrument. The use of a portable instrument is also evaluated and compared with a lab based NIR instrument in order to show the 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. The monitoring of these samples for their moisture content is an important maintenance 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 on 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 assumes that the subset of samples are actually representative of the whole sample population; the method is destructive so samples cannot be used again, and; the process is expensive and time-consuming depending on the testing done.

In contrast, the use of NIR technology is fast (seconds), easy to use and allows for rapid determination of the moisture levels in a closed sample vial. Scanning of the sample can be done through the glass vial without damaging the sample which can then be returned to the sample set. 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.

Experimental

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

Once mixed and equilibrated, samples were scanned on the lab based Antaris NIR instrument and on the portable Thermo Scientific microPHAZIR NIR instrument. The microPHAZIR has effective wavelength range from 1595-2395 nm. Five sample scans are co-averaged, and a total of 3 scans were taken of the sample. Spectral resolution was by default 8 nm. Spectral data was evaluated in the proprietary software Method Generator (MG).



The Thermo Scientific MicroPHAZIR can provide raw material identification in seconds.

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

Sample	LOD, %
1	8.93
2	8.98
3	9.85
4	10.02
5	10.66
6	11.18
7	12.16
8	12.11
9	12.43
10	12.6
11	8.93
12	8.98
13	9.85
14	10.02
15	10.66
16	11.18
17	12.73
18	13.07
19	12.43
dry	8.92

Table 1.

Once the spectra were collected, the samples were measured using Mettler Toledo LOD balance to determine the moisture content. These reference values were then combined with spectral data, and Partial Least Squares (PLS) models were made for prediction of moisture in the casein protein. The models were made and evaluated in either the MG software for the microPHAZIR or TQ Analyst for the Antaris data. Results from the Antaris were used for comparison to the microPHAZIR so model development will focus on the microPHAZIR in the following sections.

Results and Discussion

Data preparation and generation

Spectra data from the microPHAZIR was combined with the reference moisture values from the LOD balance for the 20 samples. The moisture values ranged from 9 to 13 wt/wt% as shown in the Table 1.

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

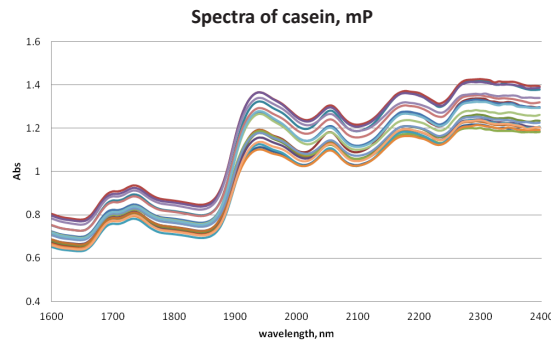


Figure 1. The collected spectra of the moisture in casein samples

Spectra preprocessing was done to optimize the robustness of the 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 well, as shown in the Figure 2.

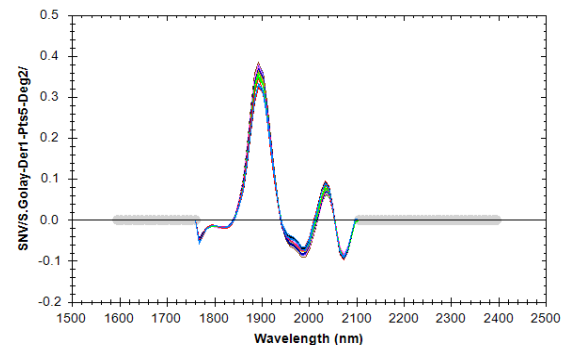


Figure 2. Wavelength was selected for the strong water overtone region

Model Development

Initially the data set was split into a calibration and a validation data set. Partial Least Squares (PLS) was used to generate the quantitative model for moisture prediction using the calibration data only. Based on the conditions

described previously, the PLS model was made 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 (Scree plot, Figure 3) indicates that 3 loadings would provide useful information to the model – and limit the noise contribution.

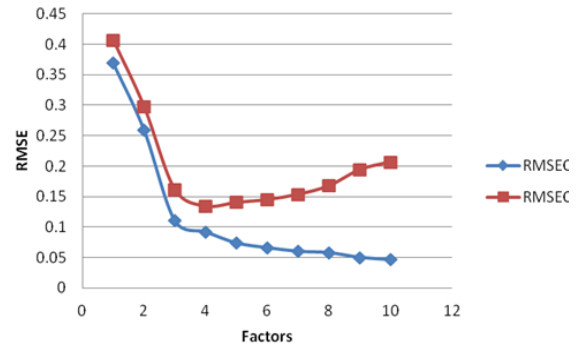
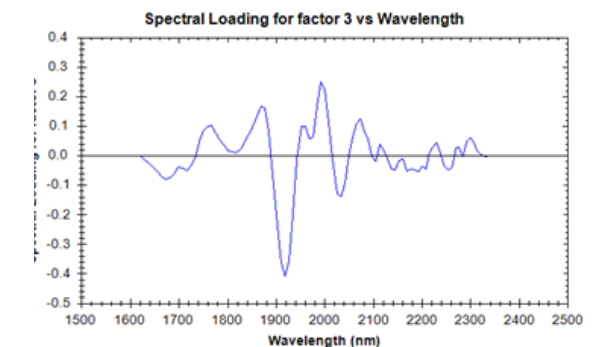
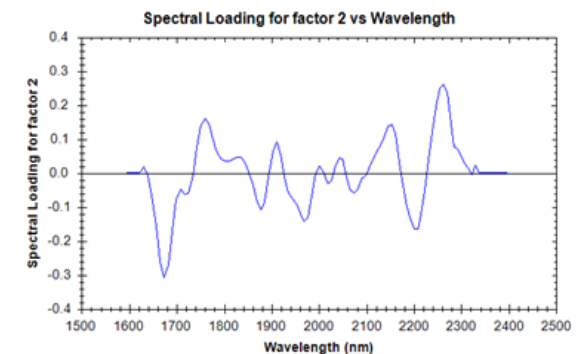
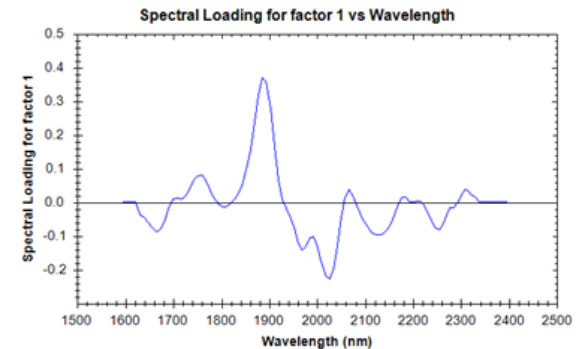


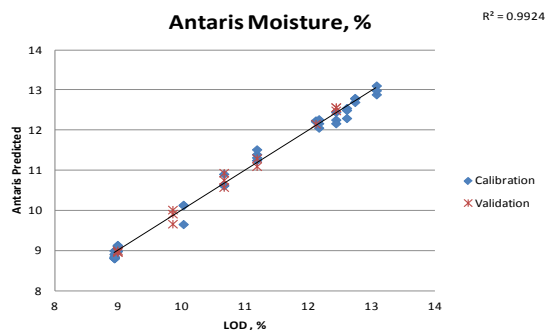
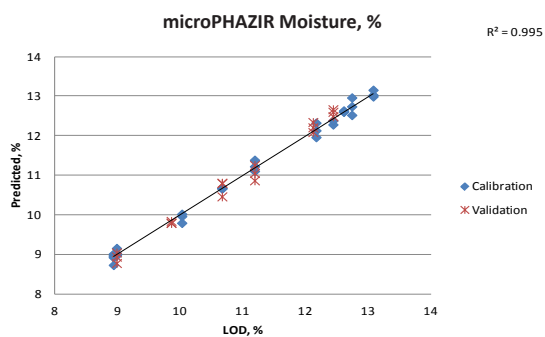
Figure 3. A plot of the number of factors and their contribution to the model

Further examination of the information provided by the 3 factors corroborates the information content. In this case, the first and third factors contain information about the

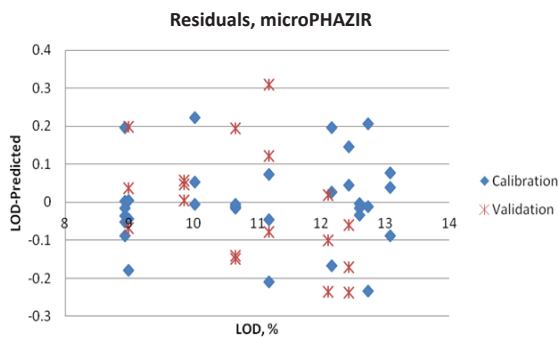


water peaks in the spectra. Interestingly, the second loading has information mostly from the CH region of the spectra indicating possible secondary structural changes associated with the protein due to differences in moisture.

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 the Figures, below. 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. In this case, 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 are very similar in terms of the linearity.



Analysis of the residuals from the microPHAZIR PLS model also indicate the model only has random variations in the residuals, as seen in the plot below for both the calibration and the validation results. Residuals are a way to evaluate if the model has any systematic variations that may cause issues with the PLS model. In this case, only random variations are observed, so this is positive for model robustness. In all these cases it is further noted that calibration and validation data are performing similarly. Validation results do have greater residuals than calibration – as is expected due to slightly higher RMSEs.



The similarity between the Antaris and the microPHAZIR are further corroborated in the root mean square (RMSE) results. RMSE is a good indicator of model performance and robustness, and as the table below shows, both instruments are expected to have good accuracy to predict moisture in the casein protein samples. Accuracy in terms of the RMSE for calibration and cross validation are similar for the Antaris and the microPHAZIR. Precision of results is indicated by the standard deviation associated with the residuals and it is also similar between instruments.

3 Factor	Antaris	microPHAZIR
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

Conclusion

Predictive models for moisture levels in casein protein were built on both a benchtop laboratory and handheld NIR spectrometer. Both systems allowed for the development of an effective, robust PLS model. The argument could be made that the benchtop NIR is comparatively better due to higher resolution. In many implementation instances, the slight decrease in accuracy of the handheld system (RMSEV = 0.12 [Antaris] vs. 0.15 [handheld]) is overcome by the advantage of doing the analysis at the point of sampling.

Based on this feasibility study, it is expected that moisture analysis of other protein samples should be equally straightforward and allow for rapid and robust prediction of lyophilized materials with the handheld technology. Deployment of the microPHAZIR can provide a valuable tool for the prediction of moisture in any stage of the pharmaceutical drying process or for final release by analyzing the moisture levels in the final product.



© 2014 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific Inc. and its subsidiaries. Specifications, terms and pricing are subject to change. Not all products are available in all countries. Please consult your local sales representative for details.

Portable Analytical
Instruments

Americas
Boston, USA
+1 978 642 1132

Europe, Middle East, Africa
Munich, Germany
+49 89 3681 380

India
Mumbai, India
+91 22 6680 3000

Asia Pacific
New Territories, Hong Kong
+852 2885 4613

www.thermoscientific.com/quality
sales.chemid@thermofisher.com

Thermo
SCIENTIFIC

Part of Thermo Fisher Scientific