Available online at www.sciencedirect.com
ScienceDirect
Journal of Pharmaceutical and Biomedical Analysis 46 (2008) 592–596
JOURNAL OF PHARMACEUTICAL AND BIOMEDICAL ANALYSIS
www.elsevier.com/locate/jpba

Short communication

Determination of moisture content of lyophilized allergen vaccines by NIR spectroscopy

Yiwu Zheng [a], Xuxin Lai [a], Susanne Wrang Bruun [b], Henrik Ipsen [a], Jørgen Nedergaard Larsen [a], Henning Løwenstein [a], Ib Søndergaard [c], Susanne Jacobsen [b,*]

[a] *ALK-ABELLO A/S, Hørsholm, Denmark*
[b] *Enzyme and Protein Chemistry, BioCentrum-DTU, Technical University of Denmark, Kgs. Lyngby, Denmark*
[c] *Center for Microbial Biotechnology, BioCentrum-DTU, Technical University of Denmark, Kgs. Lyngby, Denmark*

Received 8 June 2007; received in revised form 23 October 2007; accepted 9 November 2007
Available online 21 November 2007

Abstract

Moisture content is an important parameter for lyophilized vaccines. Currently, Karl Fischer titration is widely used for moisture determination in routine analysis. However, this method is time-consuming, sample destructive and requires environment polluting reagents, as well as the results rely on the random samplings. In this study, near infrared spectroscopy was used as a fast, non-invasive and non-destructive method to determine the moisture content in lyophilized allergy vaccines. Five different vaccine products were investigated, which contained water in the range of 0.17–1.51% (w/w, KF). Different data pre-treatments, wavelength selection and partial least squares regression were applied to construct calibration models. Multi-products model and product-specific models were obtained, which show the possibility of NIR as a rapid method to discriminate whether moisture content fit into the specifications of a pharmaceutical company.

Keywords: Moisture; NIR; Vaccine; Partial least squares regression

1. Introduction

Lyophilization (freeze-drying) is widely applied for the preservation and storage of many vaccines. Lyophilized vaccines always contain some water, which usually constitutes between 1% and 5% (w/w) of the final product after the bulk of the aqueous solvent has been removed during the freeze-drying process. Moisture content (MC) is a critical quality parameter influencing not only the vaccine appearance and quality but also its shelf life [1–3]. Currently, Karl Fischer (KF) titration is widely used for determination of MC in lyophilized products [4]. This method is time-consuming, destructive and requires environment polluting reagents. Further ambient moisture may affect the results for low moisture level samples since the vials have to be opened before analysis. In addition, many random samples have to be analysed for each batch in the production line, which is a resource consuming procedure, and random sampling has inherent disadvantages. Clusters of products deviating in water content caused by momentary production problems may not always be monitored and entire batches of the product may have to be rejected due to the results from the outliers that have been sampled. It would therefore be advantageous to replace these procedures with an automated online method.

Near infrared spectroscopy (NIR) is a physical, non-invasive and non-destructive method. Analysis can be performed through sealed glass vials, directly on the final product without opening the vials. Water is the strongest absorbing compound in the NIR region, exhibiting four absorption maxima around 970, 1190, 1450, and 1940 nm. The NIR absorption of water has been well investigated because of its importance in all applications of the technique. Strong absorption bands of water appear around 1400–1450 nm and 1900–1950 nm, which are due to the first overtone of OH stretching and combination of the OH stretching band and OH bending, respectively [5]. Variations in the water

* Corresponding author at: Enzyme and Protein Chemistry, BioCentrum-DTU, Technical University of Denmark, Søltofts Plads, Building 224, 2800 Kgs. Lyngby, Denmark. Tel.: +45 4525 2741; fax: +45 4588 6307.
E-mail address: sja@biocentrum.dtu.dk (S. Jacobsen).

0731-7085/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2007.11.011

band around the 1400 nm region are dependent on the state of water in the sample. In wet samples, i.e. water content reaching 70–90%, the adsorption band appears at similar wavelengths to those of pure water, around 1400–1410 nm. In dry samples this band associated with free water seems to disappear and a new trace of absorption appears near 1430 nm due to bound water [6]. These two bands have often been used to quantify water content in various samples. Other water bands in the NIR region are located at 970 and 1190 nm. The intensities of these bands are not strong, since they arise from the second overtone of the OH stretching band, as well as the combination of the first overtone of the OH stretching and the OH bending [5].

The determination of MC by NIR spectroscopy in both transmittance and reflectance modes have been described extensively in early literatures [7–14], covering applications in agriculture, forestry, textiles, chemicals, and pharmaceuticals. The early works have been summarized and discussed by Blanco et al. [11], and recent applications have been reviewed by Reich [7]. With regard to protein pharmaceuticals, however, NIR has not been widely adopted [7,9]. Lyophilized pharmaceutical products typically contain 1% to 2% residual water, which is a narrow range to construct a calibration model. To compensate for the narrow MC, previous studies have extended the range by adding artificial samples, which were obtained by adsorbing additional water and/or further drying [9,15]. However, the artificial calibration samples did not improve the prediction results significantly.

The objective of this study was not to determine the MC of the freeze-dried vaccine products quantitatively, but to demonstrate the feasibility of using NIR as a non-invasive, non-destructive method to rapidly discriminate whether MC fits into the specification of a pharmaceutical company. The major challenges in this work are the small amount of sample, low MC compared to Jones' work [16] and variances of sample cakes and vials. As the samples were of very low weight and MC, this work challenges NIR to be used as an alternative measurement for moisture analysis.

2. Materials and methods

2.1. Samples

Five different lyophilized allergen vaccine products, a total of 95 vials with unbroken sample cake were collected by random sampling. The sample amount is approximately 10 mg/vial.

2.2. NIR measurement

NIR measurements were carried out using a FT-NIR spectrometer (PerkinElmer Spectrum One, PerkinElmer) with an InGaAs detector. The samples were directly measured, i.e. through the bottom of the intact glass vials by diffuse reflectance without any extra preparation. NIR spectra were collected in reflectance mode at ambient temperature, although products were stored refrigerated. Each spectrum was the average of 80 scans at 16 cm^{-1} resolution over the range 700–2500 nm. Four spectra for each sample were collected by rotating the vial in 4 different angles. Therefore, a total of 380 spectra were obtained.

2.3. KF titration analysis

The same samples as mentioned above were subsequently analyzed for MC using a 737 KF Coulometer (Metrohm, Switzerland) in ALK-ABELLO A/S according to standard procedure. The overall range of MC (as analyzed by KF titration and given in Table 1) is from 16.8 to 151.0 μg/vial, which corresponds to 0.17–1.51% (w/w).

2.4. Calibration model construction

The KF data were assigned to corresponding NIR spectra. The aim of this work was not to quantitatively determine the MC. Besides partial least squares (PLS) regression, soft independent modelling of class analogy (SIMCA), as well as principal component analysis (PCA) were also performed to build models using the Unscrambler software, version 9.2 (Camo process AS). Different data pre-treatments were applied to remove baseline offset and other physical information prior to building models.

3. Results and discussion

3.1. Relationship between MC and NIR data

Raw NIR spectra of the 95 samples are shown in Fig. 1A. Because of the small amount of sample, a part of the NIR radiation passed through the sample cake and did not reflect back to the detector, which resulted in low NIR absorbance peaks and high baseline offsets. Light scattering due to the physical conditions of the sample such as particle size, density and bottle variation affected the raw spectra as well. Due to aforementioned reasons and the very low water content, the two water bands at 1400–1450 nm and 1900–1950 nm were not obvious to use.

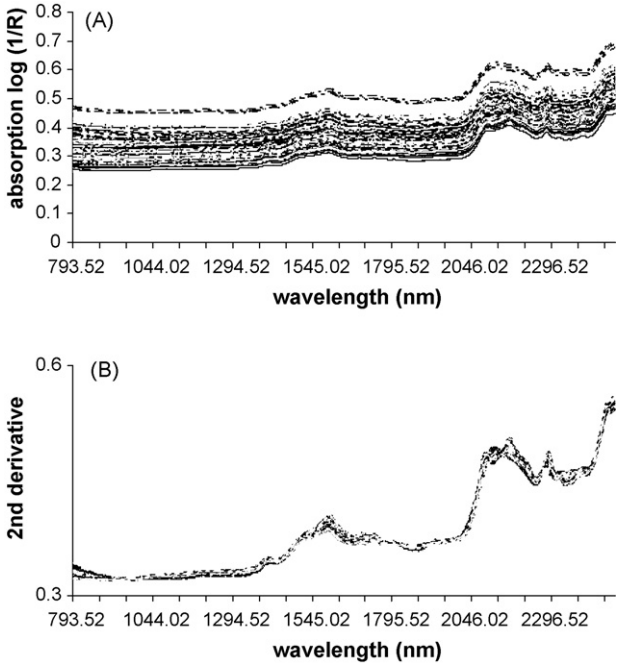

Fig. 1. NIR spectra of calibration samples: (A) raw data; and (B) EMSC data.

Table 1
Detailed information of the samples

Name of sample	Concentration of antigens (SQ-U)	Number of batches	Number of vials	Moisture content (% w/w by KF)	
				Mean	Range
Bee venom	45	3	16	0.39	0.23–0.71
	450	3			
	4500	3			
	45000	3			
	450000	3			
Cat hair and dander	45	2	33	0.44	0.17–1.23
	450	2			
	4500	2			
	45000	2			
	450000	2			
House dust mites	45	1	7	0.48	0.25–0.74
	450	1			
	4500	1			
	45000	1			
	450000	1			
Grass pollen	45	2	16	0.46	0.22–0.66
	450	2			
	4500	2			
	45000	2			
	450000	2			
Wasp venom	45	3	23	0.60	0.25–1.51
	450	3			
	4500	3			
	45000	3			
	450000	3			

Different data pre-treatments were tested to remove the effect of light scattering. Fig. 1B showed the spectra subjected to the extended multiplicative scatter correction (EMSC) pre-treatment, which is a method well suited for removal of physical effects from the chemical information [17]. Visually inspecting the EMSC corrected spectra, a clear quantitative relationship between the NIR adsorption and the MC levels could not be seen, although most of the light scattering interferences in the raw spectra were drastically reduced. However, these lyophilized samples did contain, besides residual moisture, also other components, such as antigens, adjuvant, inorganic salts, and other chemicals to facilitate the lyophilizing process and cake formation. These compounds may interact with water and cause shift of the water bands. Therefore, water information may also be derived from these spectral variations.

3.2. Multi-products calibration model for MC

PLS, SIMCA and PCA were performed to build calibration or classification models. PLS regression showed the best correlation and gave usable MC values. Thus, PLS was applied in this work, while the results of SIMCA and PCA are not shown. The general models involving five vaccine products were constructed using PLS regression with different data pre-treatments and different wavelength windows, e.g. 1900–1950 nm and the significant *X* variables/wavelength found after Jack-knifing ($p < 0.05$) over the whole NIR wavelength. Since the sample set covered 5 products with 3 batches at 5 different protein concentrations, and the data set was small compared to the product and batch variance, all NIR spectra were used for the calibration instead of dividing them into calibration and prediction data sets. The models and their performances were assessed by segment cross validation (numbers of segment is four).

Different spectral pre-treatments were used to remove the spectral disturbances and to improve the performance of the calibration models. In this study, most of the used pre-treatment methods improved the performance of the calibration models. The NIR region 1900–1960 nm is a good wavelength window for water determination; even though there were no very clear water band alterations in the spectra by visual inspection. The model based on this region showed better results than the model using the significant *X* variables in the whole wavelength window because fewer PLS factors were needed. This may be due to less noise in the region of 1900–1960 nm. The spectra contain more information when spanning the entire NIR window, but also contain more noise. Thus, more PLS factors are required. Detailed information of the corresponding models is not given here.

The best calibration model was obtained when taking the second derivative data over the 1900–1960 nm region after removal of 4 outliers. The model had a root mean square error of prediction (RMSEP) of 10.20 μg H_2O/vial with 2 PLS factors. The MC correlation plot of the data set for the best EMSC model is given in Fig. 2A. For this plot, a slope of 1 and an offset of

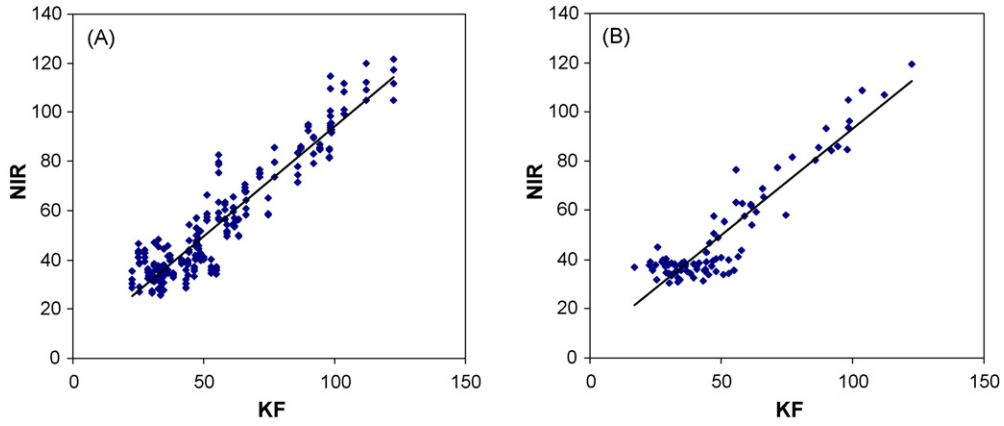

Fig. 2. The MC correlation plot between NIR prediction and KF results for multi-products models: (A) model based on original data set (RMSEP 10.20 μg H_2O/vial and 2 PLS factors): slope is 0.82; offset is 8.95; correlation is 0.90 and RMSED is 9.94 μg H_2O/vial; and (B) model based on reduced data set (RMSEP 8.99 μg H_2O/vial and 2 PLS factors): slope is 0.86; offset is 6.92; correlation is 0.93 and RMSED is 8.64 μg H_2O/vial.

0 should be ideal. In this case, a slope of 0.82 and an offset of 8.95 were achieved. The correlation was 0.90 and the root mean square error of deviation (RMSED) was 9.94 μg/vial.

Four NIR spectra were recorded for each sample at different angles. These four spectra can be regarded as four replicates, and therefore should be averaged to one spectrum for each sample. A total of 95 spectra for 95 samples were obtained after data reduction. Calibration models were built with this reduced data set using different pre-treatments and assessed by full cross validation. The performance of each model is not shown here. The best calibration model was obtained with EMSC data over the 1900–1960 nm range after removing 5 outliers. The model had a RMSEP of 8.99 μg H_2O/vial corresponding to 0.9% (w/w) with 2 PLS factors. The MC correlation plot between the NIR predication and KF results is shown in Fig. 2B. The coefficients of the plot were slightly better than the one based on the entire

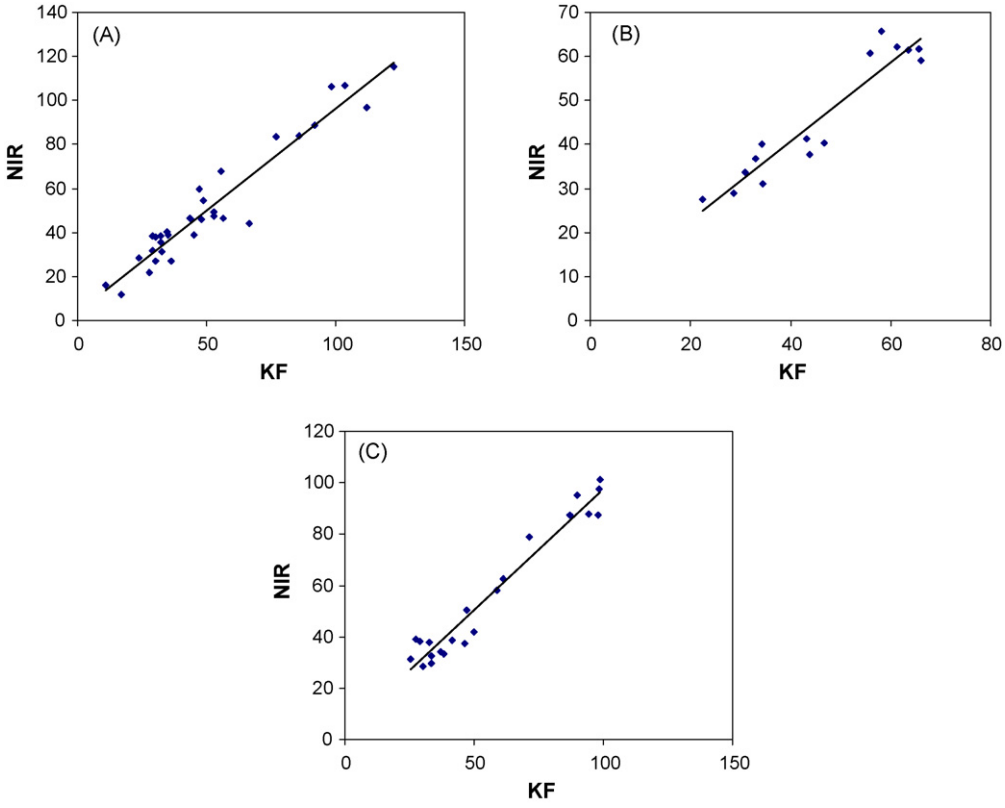

Fig. 3. The MC correlation plot between NIR prediction and KF results for product-specific models: (A) cat hair model (RMSEP 8.86 μg H_2O/vial and 2 PLS factors): slope is 0.93; offset is 3.85; correlation is 0.96 and RMSED is 7.70 μg H_2O/vial; (B) Grass pollen model (SMSEP 5.54 μg H_2O/vial and 1 PLS factor): slope is 0.89; offset 4.85; correlation is 0.95; and RMSED is 4.68 μg H_2O/vial; (C) Wasp venom model (SMSEP 6.66 μg H_2O/vial and 2 PLS factor): slope is 0.95; offset is 2.80; correlation is 0.97 and RMSED is 5.85 μg H_2O/vial.

data set as seen from the higher correlation (0.93) and lower RMSED (8.65).

The model based on the reduced data set did not change significantly compared to the model with the original data set. The results may indicate that the angle effect is not important. However, a spinner rotator may benefit the measurement, since it rotates the sample during detection which results in a truly average spectrum.

3.3. Product-specific calibration model for MC

The multi-calibration model set included five different products; however, different protein antigens may interact with water in different ways affecting the final calibration model. To investigate the product-specific effect and to improve the predictions, calibration models were built for each product except the house dust mite vaccine, because the sample size ($n = 7$) was too small. Calibration models were constructed on the reduced data set (the average spectrum of the four spectra resulting from four different angles for each sample) and assessed by full cross validation using different data pre-treatment methods in two wavelength windows: whole wavelength region and 1900–1960 nm. The detailed data of the models are not shown here.

The results indicated that both EMSC and second derivative pre-treatment improved the performance of the calibration models for each product. Wavelength selection, i.e. 1900–1960 nm, also benefited the models for cat hair and wasp venom vaccines. The best calibration model for the cat hair vaccine was obtained using the 2nd derivative of the data over the significant *X* variables in region of 1900–1960 nm. This model had a RMSEP of 8.86 μg H_2O/vial with 2 PLS factors. For the grass pollen vaccines, the best calibration model was achieved using the EMSC data over the significant *X* variables in the whole NIR region. The model possessed a RMSEP of 5.54 μg H_2O/vial with one PLS factor. The best model for the wasp venom vaccine was obtained using the 2nd derivative of the data over the significant *X* variables in the region of 1900–1960 nm. The model has a RMSEP of 6.66 μg H_2O/vial with 2 PLS factors. The MC correlation plots between NIR prediction and KF results for the best models were shown in Fig. 3.

With regard to the bee venom product, it was difficult to build an acceptable calibration model for the determination of MC. This may be caused by the small sample size ($n = 16$) and large batch variance (3 batches).

In conclusion, the product-specific models give good results compared to the multi-product models. However, the sample sizes in the current product-specific models are limited. Normally a training set with 100 samples is ideal [18]. Thus, further investigations are required for the product-specific models. Nevertheless, a multi-product calibration model is more useful for routine analysis because it is more user-friendly. The MC correlation plot demonstrate that NIR models consistently produce higher moisture levels compared to KF results, see Figs. 2 and 3. It may be due to the headspace moisture in NIR measurement in light of the refrigeration and warming.

4. Conclusion

This preliminary study demonstrates that NIR spectroscopy has the possibility to be used as a non-invasive, non-destructive method to rapidly discriminate whether moisture content fit into the specifications of lyophilized allergen vaccines. The moisture content results obtained from NIR spectroscopy show acceptable agreement with results obtained by KF titration. The models based on product-specific data are slightly better than models based on multi-product data; however, a general multi-product model is more useful in routine quality analysis, since one model can be applied for all product samples.

Acknowledgements

We gratefully acknowledge the financial support from the Ministry of Science, Technology and Innovation under its Industrial PhD initiative.

References

[1] A.M. Klibanov, J.A. Schefiliti, Biotechnol. Lett. 26 (2004) 1103–1106.
[2] H.R. Costantino, K. Griebenow, P. Mishra, R. Langer, A.M. Klibanov, Biochim. Biophys. Acta 1253 (1995) 69–74.
[3] S.P. Schwendeman, H.R. Costantino, R.K. Gupta, G.R. Siber, A.M. Klibanov, R. Langer, Appl. Biol. Sci. 92 (1995) 11234–11238.
[4] Food and Drug Administration. Guideline for the determination of residual moisture in dried biological products (Docket No. 89D-0140).
[5] W.A.P. Luck, Structure of Water and Aqueous Solution, Verlag Chemie-Physik, Weinheim, Germany, 1974.
[6] H. Büning-Pfaue, Food Chem. 82 (2003) 107–115.
[7] G. Reich, Adv. Drug Del. Rev. 57 (2005) 1109–1143.
[8] K.A. Bakeev, Spectroscopy 19 (2004) 39–42.
[9] T.P. Lin, C.C. Hsu, PDA J. Pharm. Sci. Technol. 56 (2001) 196–205.
[10] X. Dziki, J.F. Bauer, J.J. Szpylman, J.E. Quick, B.C. Nichols, J. Pharm. Biomed. Anal. 22 (2000) 829–848.
[11] M. Blanco, J. Coello, H. Itirriaga, S. Maspoch, C. de la Pezuela, Analyst 123 (1998) 135R–150R.
[12] X. Zhou, P. Hines, M.E. Borer, J. Pharm. Biomed. Anal. 17 (1998) 219–225.
[13] M.S. Kamat, R.A. Lodder, P.P. DeLuca, Pharm. Res. 6 (1989) 961–965.
[14] E.W. Ciurczak, R.P. Torlini, M.P. Demkowicz, Spectroscopy 1 (1986) 36–39.
[15] L. Sukowski, NIR-based process analytical technology: In-line residual moisture determination for a complete batch inspection of lyophilized end-products, PhD. Thesis, University of Basel, Deutschland, 2003.
[16] J.A. Jones, I.R. Last, B.F. MacDonald, J. Pharm. Biomed. Anal. 11 (1993) 1227–1231.
[17] H. Martens, J.P. Nielsen, S.B. Engelsen, Anal. Chem. 75 (2003) 394–404.
[18] P.C. Williams, K. Norris, in: P. Williams, K. Norris (Eds.), Near-Infrared Technology in the Agricultural and Food Industries, American Association of Cereal Chemists, Inc., St. Paul, Minnesota, USA, 2001, pp. 171–185.