ELSEVIER

Journal of Pharmaceutical and Biomedical Analysis
26 (2001) 73–78

JOURNAL OF
PHARMACEUTICAL
AND BIOMEDICAL
ANALYSIS

www.elsevier.com/locate/jpba

Near-infrared spectroscopy as a convenient analytical method for alkyl polyglycosides

Jong-Yun Kim [a], Young-Ah Woo [b], Hyo-Jin Kim [b], Jong-Duk Kim [a,*]

[a] *Department of Chemical Engineering, Korea Advanced Institute of Science and Technology, Taejon 305-701, South Korea*
[b] *College of Pharmacy, Dongduk Women's University, Seoul 136-714, South Korea*

Received 03 January 2001; received in revised form 10 January 2001; accepted 15 January 2001

Abstract

Near infrared (NIR) spectroscopy is used for the rapid determination of decyl glucopyranoside (10G1) and alkyl polyglycoside (AG) mixtures having different alkyl chain length and the number of glucose unit in aqueous solution. NIR spectroscopy is a much simpler spectroscopic analysis method compared to three analytical methods for mixture of AGs such as thin layer chromatography (TLC), high performance liquid chromatography (HPLC), and photometry method. NIR spectra of AGs between 0.030 and 0.540 mg/ml in aqueous solutions were utilized to develop a calibration model. Both raw spectra and the second derivatives of AGs were tested for the best fit. The best calibration was built with second derivative spectra by using multiple linear regression (MLR). The standard error of calibration (SEC) and the standard error of prediction (SEP) were used for the evaluation of the model. The best calibration provides an SEP of 0.052 and 0.061 mg/ml for the prediction set of 10G1 and AG mixture, respectively.

Keywords: Near-infrared spectroscopy; Alkyl polyglycosides; Decyl glucopyranoside; Multiple linear regression method; Non-destructive analytical method

1. Introduction

Although alkyl polyglycoside (AG) was first synthesized by Fischer 100 years ago [1], it draws interest recently due to its biodegradability, good dermatological properties and synergistic effects with other surfactants [2]. Commercial AGs are usually a mixture of AGs having a different number of head group and alkyl chain length. AGs cannot be determined by the conventional methods for surfactants such as Epton titration or determination of bismuth-active substances. Four analytical methods for AG by GC (gas chromatography), TLC (thin layer chromatography), HPLC (high performance liquid chromatography), and photometric method (UV-Vis spectroscopy) for the determination of the total concentration in aqueous solution has been reported [3–8]. But these techniques also require sample preparations and several steps for quanti-

* Corresponding author. Tel.: +82-42-8698969/8693921; fax: +82-42-8693910.
E-mail address: kjd@cais.kaist.ac.kr (J.-D. Kim).

0731-7085/01/$ - see front matter
PII: S0731-7085(01)00355-7

tative analysis. Therefore, a simple, rapid, and non-destructive analytical technique to determine AG concentration in aqueous solution is in great demand.

The near infrared (NIR) region of the spectrum encompasses radiation with wavelength ranging from 800 to 2500 nm. Absorption in the NIR region always occurs due to the vibration caused by combination and overtone of fundamental vibrations in the infrared (IR) region [9]. In particular, absorption in the NIR mainly occurs in the functional groups having a hydrogen atom such as C–H, N–H, and O–H [10–12]. Although absorption in the NIR region is relatively weak, broad, and overlapped, NIR spectroscopy combined with multivariate calibration technique has been known as a convenient tool for component analysis. Quantitative analytical methods of glucose [13], sugar [14], and polymer [15] using NIR spectroscopy have been reported. Therefore NIR spectroscopy can be proposed as a simple and convenient method for quantitative determination of AG mixture and pure decyl glucopyranoside (10G1) in aqueous solution in this study.

2. Experimental

2.1. Materials and method

10G1 was obtained from Sigma Co. Glucopon 220UP, where the mixture of AGs with C8 and C10 alkyl chains and average degree of polymerization of 1.4, is obtained from Henkel Co. The general chemical structure of AG is shown in Fig. 1. NIR transmittance spectra were collected using NIRSystems model 6500 spectrometer (Foss NIRSystems Inc., MD) equipped with a quartz halogen lamp and PbS detector. The spectra were collected using Near Infrared Spectral Analysis Software (NSAS, Foss NIRSystems Inc., MD). To reduce noise, 32 scans per sample were performed and averaged. A background scan was obtained for air and used as the reference. The spectra were acquired with a cuvette cell with a quartz window. Multiple Linear Regression (MLR) and second derivative algorithm were performed using NSAS software packages (Foss NIRSystems Inc., MD). For both 10G1 and AG mixture, nine spectra were used for the calibration set and the prediction set as listed in Table 1. The concentrations in the calibration set and the prediction set ranged from 0.060 to 0.540 and 0.030 to 0.510 mg/ml, respectively.

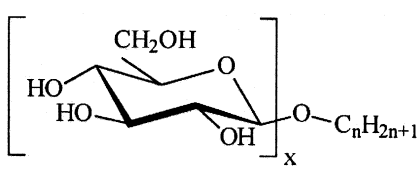

Fig. 1. Chemical structure of AG. x is degree of polymerization in head group and n alkyl chain length.

Table 1
Data set preparation for 10G1 and AG mixture

	Calibration set	Prediction set
Number of samples	9	9
Concentration range (mg/ml)	0.060–0.540	0.030–0.510
Average (mg/ml)	0.300	0.270
S.D. (mg/ml)	0.164	0.164

2.2. Data analysis using MLR

The MLR method is utilized to reduce NIR spectral data to concise and useful information for concentration. For the development of the MLR calibrations it was necessary to select the single wavelength with the highest correlation to the concentration of the analyte. In order to find the better calibration model, second wavelength was selected so that the SEC was minimum. Selection of further wavelengths continued until the improvements to the model were considered minimal. Only a few wavelengths were chosen, typically less than 5, for an MLR calibration.

The performance of the calibration model with respect to fitting the calibration spectra was assessed by the standard error of calibration (SEC), given in Eq. (1). It is calculated when building a model for each model with the selected wavelength(s).

$$\mathrm{SEC} = \left(\frac{\sum_{i=1}^{n} (x - \hat{x})^2}{n - m - 1} \right)^{0.5} . \quad (1)$$

In this equation, x is the actual AG concentration for the sample, $\hat{x}$ represents the predicted concentration for the sample, and m is the number of wavelengths used in the model and n is the number of samples in the data set. The minimum SEC shows the number of wavelengths which optimally fit the data. The standard error of prediction (SEP) is also calculated in Eq. (2) to avoid overfitting the data and confirm the robustness of the calibration model.

$$\mathrm{SEP} = \left(\frac{\sum_{i=1}^{n} (x - \hat{x})^2}{n} \right)^{0.5} . \quad (2)$$

The SEP is calculated for the samples that were not included in calibration set for SEC when developing the model. Because the actual concentration values are expressed in terms of mg/ml, all parameters in Eqs. (1) and (2) will have the units of mg/ml.

3. Results and discussion

Representative NIR spectrum of 10G1 is presented in Fig. 2. The dominant absorption bands around 1450 and 1950 nm in the spectrum are assigned to the first overtone of OH stretching and the combination of OH stretching and bending vibration, respectively. The spectral bands to be correlated with the concentration could not be specified in the raw spectra of 10G1. In order to enhance spectral features, second derivative spectra were calculated as shown in Fig. 3. In the analysis of NIR spectra, derivative spectroscopic techniques are typically utilized since baseline offsets are largely eliminated without compromising the signal-to-noise ratio [16]. Absorbance maxima in the second derivative spectrum are inverted to minima that are surrounded by a positive sidelobe on each side. The spectral region around 1392 nm is highly correlated to the concentration of 10G1 in second derivative spectra. The best-fit

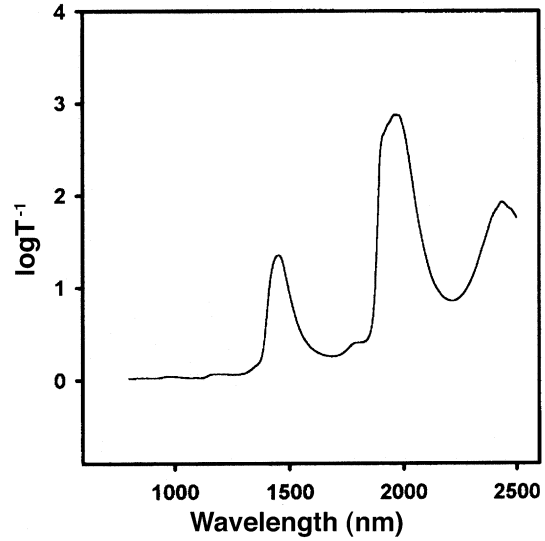

Fig. 2. The representative NIR spectrum of 10G1 samples in the calibration set. T: transmittance.

model was found by using MLR. The linear summation of wavelengths, which are most highly correlated to concentration, was used for MLR model. In this study, only four discrete wavelengths were investigated to develop the model to

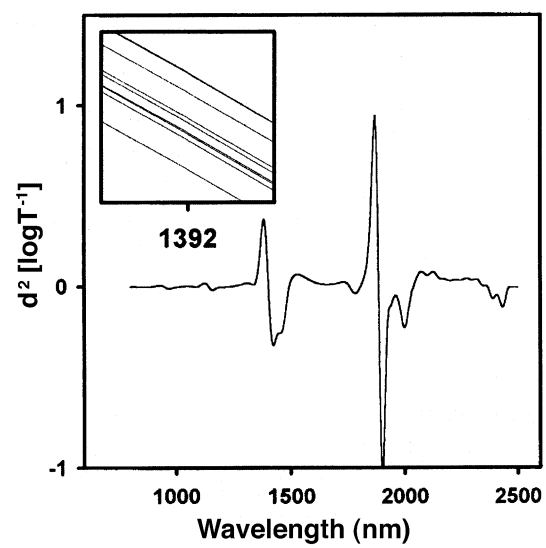

Fig. 3. The second derivative spectra of 10G1. Inset: magnification for clear comparison of samples with different concentration around 1392 nm. T: transmittance.

avoid overfitting. Multiple correlation coefficient (R), standard error of calibration (SEC), standard error of prediction (SEP) were calculated to evaluate the four wavelength model and find the best model. R represents the multiple correlation coefficient which is a measure of the agreement between the NIR data and actual concentration of AGs for the calibration set. The SEC and SEP were calculated for calibration set and prediction set, respectively. For each model, SEC and SEP were calculated and the results are shown in Table 2. For 10G1, the MLR model using the linear summation of four wavelength of 1392, 1532, 2320, and 2232 nm had best results, providing an SEC of 0.008 mg/ml and an SEP of 0.052 mg/ml. It is supposed that the absorbance bands around 1392 are due to combination of the second overtone of C–H stretching and C–H deformation. The bands around 1532 nm could be assigned to the first overtone of C–H stretching [9]. It is known that a region, 2083–2381 nm, contains glucose bands of high absorptivity [17]. The absorption bands around 2320 and 2232 nm in second derivative spectrum of 10G1 were assigned to glucosyl group. Using the linear summation of four wavelengths, 1392, 1532, 2320, and 2232 nm, the scatter plot shows good correlation between 10G1 concentration and NIR data in Fig. 4. Filled squares and open squares represent calibration and prediction data, respectively. The calibration and prediction data has good correlation with actual concentration and many points fall on or close to the unity line.

The spectral bands correlated to the concentration could not be specified in the raw spectra of AG mixture similar to the case for 10G1. Only

Table 2
Results of calibration and prediction for determination 10G1 in aqueous solution using MLR

λ (nm)	R	SEC (mg/ml)	SEP (mg/ml)
1392	0.9782	0.044	0.074
1392, 1532	0.9944	0.024	0.067
1392, 1532, 2320	0.9987	0.012	0.062
1392, 1532, 2320, 2232	0.9996	0.008	0.052

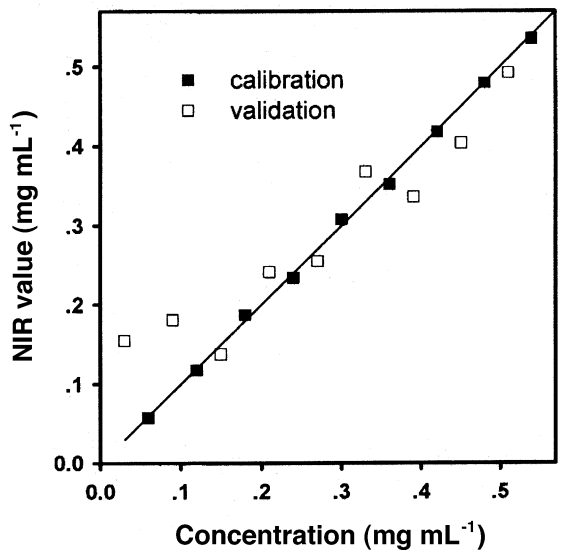

Fig. 4. Scatter plot showing correlation between actual concentration and NIR value using the four discrete wavelengths of 1392, 1532, 2320, and 2232 nm for 10G1.

large absorption bands around 1450 and 1950 nm in the spectrum are assigned to the OH stretching of water. The second derivative spectra were acquired to remove the baseline shift as shown in

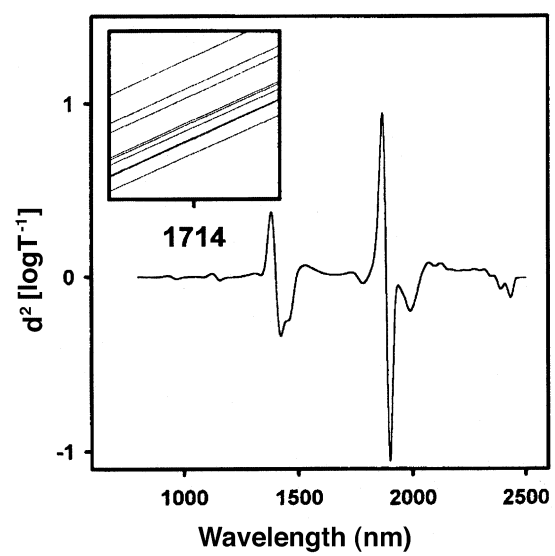

Fig. 5. The second derivative spectra of AG mixture. Inset: magnification for clear comparison of samples with different concentration around 1714 nm. T: transtmittance.

Table 3
Results of calibration and prediction for determination AG mixture in aqueous solution using MLR

λ (nm)	R	SEC (mg/ml)	SEP (mg/ml)
1714	−0.9198	0.074	0.079
1714, 878	0.9907	0.027	0.068
1714, 878, 2368	0.9961	0.019	0.066
1714, 878, 2368, 2408	0.9993	0.009	0.061

Fig. 5. As previously described, up to four wavelengths were considered to find the best model. It was found that absorption around 1714 nm is related to the concentration of AG mixture in second derivative spectra in the inset of Fig. 5. The absorbance bands around 1720 are due to the first overtone of C–H stretching. However, the model using only 1714 nm is not adequate for estimating the total concentration of AG mixture, since it provides low correlation coefficient of −0.9198 as listed in Table 3. Accordingly, second and third wavelengths were found to compensate the model. The SEC and SEP values are lower compared to one wavelength-model. However, the linear combination of the four discrete wavelengths of 1714, 878, 2368, and 2408 nm shows the best result with an SEC of 0.009 mg/ml and an SEP of 0.061 mg/ml as listed in Table 3. For the AG mixture, it is difficult to clearly assign the absorption at the four wavelengths, since the sample is a mixture of alkyl polyglycosides with C8 and C10 alkyl chains and the bands overlap each other in the NIR spectra. Fig. 6 shows the scatter plot showing correlation between AG concentration and NIR data using the linear summation of the four wavelengths, 1714, 878, 2368, and 2408 nm. The plot is highly linear throughout the concentration range from 0.030 to 0.540 mg/ml.

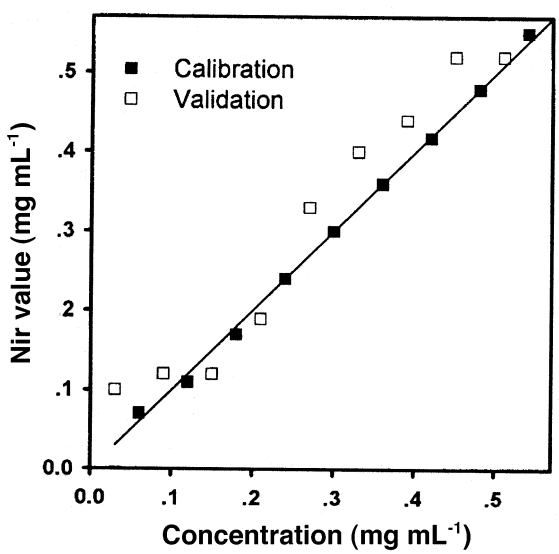

Fig. 6. Scatter plot showing correlation between actual concentration and NIR value using the four discrete wavelengths of 1714, 878, 2368, and 2408 nm for AG mixture.

NIR spectroscopy using MLR shows the possibility to determine total concentration of AG mixture with C8 and C10 alkyl chains in aqueous solution and pure 10G1 in aqueous solution. Furthermore, NIR spectroscopy can be applied to analyze other pure alkyl glycosides and total concentration of other AG mixture without pretreatment for measurement.

4. Conclusion

NIR spectroscopy as a rapid, accurate, and non-destructive determination of 10G1 and AG mixtures in aqueous solution was investigated. Enhanced spectral features were obtained after second derivatization of the raw spectra. It was found that the absorption around 1392 and 1714 nm in each representative spectrum was correlated to the concentration of 10G1 and AG mixtures, respectively. With the use of MLR in the NIR region from 800 to 2500 nm, the best calibration models were built, showing good correlation with the actual concentration of 10G1 and AG mixtures. These results indicate NIR spectroscopy can be used as a simple and convenient method for the determination of AG mixtures as well as pure 10G1 in aqueous solution.

References

[1] E. Fischer, Chem. Berichte 26 (1893) 2400.
[2] P. Busch, H. Hensen, H. Tesmann, Tenside Surf. Det. 30 (2) (1993) 116–121.

[3] N. Buschmann, S. Wodarczak, Tenside Surf. Det. 32 (4) (1995) 336–339.
[4] N. Buschmann, A. Kruse, S. Wodarczak, Agro-Food-Industry Hi-Tech 12 (1996) 6–8.
[5] H.S. Klaffke, T. Neubert, L.W. Kroh, Tenside Surf. Det. 35 (2) (1998) 108–111.
[6] H.S. Klaffke, T. Neubert, L.W. Kroh, Tenside Surf. Det. 36 (3) (1999) 178–184.
[7] H. Waldhoff, J. Scherler, M. Schmitt, J.R. Varvil, in: K. Hill, W. von Rybinski, G. Stoll (Eds.), Alkyl Polyglycosides: Technology Properties and Applications, VCH Verlagsgesellschaft mbH, Weinheim, 1997, pp. 23–38.
[8] M.A. Jermyn, Anal. Biochem. 68 (1975) 332–335.
[9] B.G. Osborne, T. Fearn, Near Infrared Spectroscopy In Food Analysis, Longman Scientific & Technical, New York, 1986.
[10] E. Stark, K. Luchter, M. Margoshes, Appl. Spectrosc. Rev. 22 (4) (1986) 335–399.
[11] L.G. Weyer, Appl. Spectrosc. Rev. 21 (1) (1985) 1–43.
[12] D.L. Wetzel, Anal. Chem. 55 (1983) 1165–1166A.
[13] H. Chung, M.A. Arnold, M. Rhiel, D.W. Murhammer, Appl. Spectrosc. 50 (1996) 270–276.
[14] J.B. Reeves, Appl. Spectrosc. 50 (1996) 154–160.
[15] S. Cossin, M. Conell, B. Cross, R. Winter, Appl. Spectrosc. 50 (1996) 900–995.
[16] T.C. O'Haver, T. Begley, Anal. Chem. 53 (1876–1878) 1981.
[17] M.A. Arnold, G.W. Small, Anal. Chem. 62 (1990) 1457–1464.