Determination of total concentration of alky polyglycosides in aqueous solution using near infrared transmittance spectroscopy

Jong-Yun Kim,^a Young-Ah Woo,^b Hyo-Jin Kim^b and Jong-Duk Kim^{a*}

^aDepartment of Chemical Engineering, Korea Advanced Institute of Science and Technology, Taejon 305-701, Korea.

^bCollege of Pharmacy, Dongduk Women's University, Seoul, 136-714, Korea.

Introduction

Commercial alkyl polyglycosides (AGs) are usually mixtures of different numbers of head groups and alkyl chain lengths. AGs cannot be determined by the conventional methods for surfactants such as Epton titration or determination of Bismuth-active substances. Three analytical methods for AG by 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.¹ But those techniques also require sample preparations and include several steps. Therefore a simple, rapid and non-destructive analytical technique to determine AG mixture in aqueous solution is in great demand. Quantitative analytical methods of glucose,² sugar³ and polymers⁴ using near infrared (NIR) spectroscopy and multivariate calibration have been reported. In this study, NIR spectroscopy is suggested as a simple and convenient method for quantitative determination of AG mixture as well as pure AG-decyl glucopyranoside (10G1) in an aqueous solution.

Materials and methods

10G1 was obtained from Sigma Co. Glucopon 225UP, mixtures of alkyl polyglycosides with 8 and 10 alkyl chain and average DP (degree of polymerisation) of 1.4, were obtained from Henkel Co. Gen-

eral chemical structure of alkyl polyglycosides is shown in Figure 1. NIR transmittance spectra were collected using an NIRSystems model 6500 spectrometer (Foss NIRSystems Inc., MD, USA) equipped with a quartz halogen lamp and PbS detector. The spectra were collected with 2 nm data intervals using NIR Spectral Analysis Software (NSAS, Foss NIRSystems Inc.). To reduce noise, 32 scans per sample were obtained and averaged. A background scan, *To*, of air was obtained and used as the reference. The spectra obtained consisted of -Log (Ts/To) over the wavelength range of 800 to 2500 nm, where *Ts* is the light transmit-



Figure 1. Chemical structure of alkyl polyglycoside. x is degree of polymerisation in head group and n alkyl chain length.

	Calibration set	Prediction set
Number of samples	9	9
Constituent range (mgmL ⁻¹)	0.030-0.570	0.060-0.540
Average	0.300	0.270
Standard deviation	0.164	0.164

Table 1. Data set preparation for decyl glucopyranoside(10G1) and AG mixture.

ted through the sample. 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 package (Foss NIRSystems Inc., MD, USA).

For 10G1 and AG mixtures, the calibration set consisted of a collection of 10 spectra, while the prediction set was composed of 10 spectra. The concentration of the calibration set and the prediction set ranged from 0.030 to 0.570 mg mL⁻¹ and 0.060 to 0.540 mg mL⁻¹, respectively as listed in Table 1.

Results and discussion

The representative NIR spectra are presented in Figure 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 correlating the concentration could not be specified in the raw spectra. In order to enhance spectral features, the second derivative spectra were calculated as shown in Figure 3. In the analysis of NIR spectra, derivative spectroscopic techniques are typically utilised since baseline offsets are largely eliminated without compromising the signal-to-noise ratio.⁵ Absorbance maximum in the second-derivative spectrum is inverted to minimum which is surrounded by a positive side-lobe on each side. It is recognised that the spectral region around 1392 nm is correlated more to the concentration of 10G1 in second derivative





Figure 2. The representative NIR spectra of decyl glucopyranoside (10G1) samples in the calibration set.

Figure 3. The second derivative spectra of AG mixture. Inset: magnification for clear comparison of samples with different concentration around 1714 nm.

λ	R	SEC	SEP
(nm)		$(mg mL^{-1})$	$(mg mL^{-1})$
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

Table 2. Results of calibration and prediction for determination decyl glucopyranoside(10G1) in aqueous solution using multiple linear regression (MLR).

spectra, when the wavelength most highly correlated to concentration of 10G1 was first found for MLR. The best model was found by using linear summation of wavelengths which are highly correlated to concentration with multiple linear regression (MLR). As previously described, only four discrete wavelengths were investigated to develop the model to avoid overfitting. To evaluate a model, multiple correlation coefficient (*R*), standard error of calibration (*SEC*) and standard error of prediction (*SEP*) were calculated. *R* represents the multiple correlation coefficient which is a measure of the agreement between the NIR data and actual concentration of AG for the calibration set. The *SEC* and *SEP* were calculated for the calibration set and prediction set, respectively. For each model, *SEC* and *SEP* were calculated as shown in Table 2. For 10G1, the MLR model, using the linear summation of four wavelengths of 1392, 1532, 2320 and 2232 nm, has better results, providing a *SEC* of 0.008 mg mL⁻¹ and a *SEP* of 0.053 mg mL⁻¹. The absorbance bands around 1392 are due to combination of the second overtone of C–H stretching. It is known that a region, 2083–2381 nm, contains glucose bands of high absorptivity.⁶ The absorption bands around 2320 and 2232 nm were assigned to the

glucosyl group of 10G1. 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 Figure 4. Open circles and filled circles represent calibration and prediction data, respectively. The calibration and prediction data have good correlation with actual concentration and many points fall on or close to the unity line.

The spectral bands correlating the concentration could not be specified in the raw spectra of AG. Only absorption bands around 1450 and 1950 nm in the spectrum are assigned to the OH stretching of water. In order to remove the baseline shift, the second derivative spectra were acquired as shown in Figure 3. As previously described, four wavelengths were investigated to find the best model. It was found that absorption around 1714 nm is more related to the concentration of AG mixture in the second derivative spectra in the inset of Figure 3. The absorbance bands



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

λ (nm)	R	$\frac{SEC}{(\text{mg mL}^{-1})}$	SEP (mgmL ⁻¹⁾
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

Table 3. Results of calibration and prediction for determination AG mixtures in aqueous solution using multiple linear regression (MLR).

around 1720 are due to the first overtone of C-H stretching. However, this single-wavelength model is not adequate for estimating the total concentration of AG mixture since the model provides a low correlation coefficient of -0.9198 as listed in Table 3. Accordingly, second and third wavelengths were found to compensate the model. It is shown that the SEC and SEP values are lower compared with a one wavelength model. However, the linear combination of the four discrete wavelengths 1714, 878, 2368 and 2408 nm show better result swith an SEC of 0.009 and SEP of 0.061, as shown in Table 3. For AG mixtures, it is difficult to clearly assign the absorption at the four wavelengths, since the samples are mixtures of alkyl polyglycosides with 8 and 10 alkyl chain and the absorptions overlapped with each other in the NIR spectra. Figure 5 shows the scatter plot showing correlation between AG concentration and NIR data using the linear summation of 1714, 878, 2368 and



Figure 5. 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 samples.

2408 nm. The plot is highly linear throughout the $0.030-0.570 \text{ mg mL}^{-1}$.

Conclusion

A rapid, accurate and nondestructive quantitative determination of 10G1 and total concentration of AG mixture in aqueous solution was performed using NIR spectroscopy. The NIR region from 800 to 2500 nm was used and the models for quantitative analysis were developed by using MLR. For 10G1 and AG mixtures, the second derivative NIR data were used to build the model. For NIR analysis of 10G1 and AG mixtures, the NIR data with the concentration range from 0.030 to 0.570 mg mL⁻¹ were used in this study. The best calibration was built by using MLR, providing an *SEP* of 0.053 and 0.061 mg mL⁻¹ for 10G1 and AG mixtures, respectively. These results indicate NIR spectroscopy can be used as a simple and convenient method for the determination of AG mixtures as well as 10G1 in aqueous solution.

References

- 1. N. Buschmann, A. Kruse and S. Wodarczak, Agro-Food-Industry Hi-Tech Jan, 6 (1996).
- 2. H. Chung, M.A. Arnold, M. Rhiel and D.W. Murhammer, *Applied Biochemistry and Biotechnology* **50**, 109 (1995).
- 3. J.B. Reeves, Appl. Spectrosc. 50, 154 (1996).
- 4. S. Cossin, M. Conell, B. Cross and R. Winter, Appl. Spectrosc. 50, 900 (1996).
- 5. T.C.O'Haver and T. Begley, Anal. Chem. 53, 1876 (1981).
- 6. M.A. Arnold and G.W. Small, Anal. Chem. 62, 1467 (1990).