Simple and Rapid Quantitative Determination of Di(2-ethylhexyl)phosphoric Acid Concentration Using Near Infrared Spectroscopy

Jong-Yun Kim,[*] Young-Ah Woo,[†,*] Myung-Ho Lee, Hyo-Jin Kim,[†] Yeong-Jae Park, and Won-Ho Kim

Nuclear Chemistry Research Division, Korea Atomic Energy Research Institute, Daejeon 305-353, Korea
[*]E-mail: kjy@kaeri.re.kr
[†]Department of Pharmacy, Dongduk Women's University, Seoul 136-714, Korea. [*]E-mail: wooy@dongduk.ac.kr
Received September 11, 2006

Amphiphilic di(2-ethylhexyl)phosphoric Acid (DEHPA) is an excellent cationic extractant in chemical, hydrometallurgical, and nuclear processes, especially in trivalent lanthanides or actinides separation process.[1] Since DEHPA has two long alkyl chains and a small phosphate group, and consequently is very hydrophobic as shown in Figure 1, it is usually used in apolar media such as hexane, octane, etc. Sodium salt of DEHPA has been also widely used in the form of a reverse micelle in an apolar media, and therefore used for a variety of extraction applications such as liquid-liquid protein extraction[2] and aminoglycoside antibiotic extraction.[3] In solid-liquid separation, impregnation of DEHPA into a porous polymer support in an organic solvent is an important process, and therefore adsorbed DEHPA in organic solvents should be quantitatively determined. To our best knowledge, gravimetric measurement,[4] phosphorus elemental analysis,[5] potentiometric titration with NaOH[6] were the conventional method reported for the determination of DEHPA, while the spectrophotometric analysis of DEHPA in an organic solvent has not been taken into account. Total organic carbon method as a general quantitative determination tool for most organic materials cannot be used in organic media such as hexane. Despite the disadvantages of requiring the time-consuming calibration and validation with plenty of samples as well as a proper understanding of chemometric techniques in near infrared (NIR) technique for industrial applications, practically useful instrumentations such as attractive portable NIR system have been suggested recently.[7] In this study, common glass vial was used for the convenient NIR quantification of DEHPA in the presence of the interference from a solvent absorption by using NIR spectroscopic technique equipped with a user-friendly commercial software for a multivariate calibration. This will be technically useful for various field applications. Simple wavelength selection and effect of pathlength were also investigated.

Experimental

N-Hexane was glass-distilled HPLC-grade and purchased from J. T. Baker (Phillipsburg, NJ). Di(2-ethylhexyl)phosphoric acid (DEHPA) was analytical-grade from Merck and used without additional purification. FT-NIR spectra were recorded on a Bruker MPA FT-NIR spectrometer using a 1-mm-pathlength rectangular glass cell (Bruker), 8-mm-pathlength cylindrical glass vial (Bruker), and 24-mm-pathlength cylindrical scintillation vial (Fisher Scientific). Blackman-Harris 3-term apodization function was used for Fourier transformation, and a zerofilling factor of 2 was applied to yield an encoding interval of approximately one data point per wavenumber. All data acquisitions and Fourier transformations were performed using the data acquisition software, OPUS 4.2 package supplied by Bruker in the spectral range of 830-2,500 nm. All FT-NIR spectra of DEHPA in hexane are the average of 32 scans collected at 8-cm^{-1} resolution. Two sets of the same-sized standard solutions for calibration and prediction were prepared using hexane and by the successive dilution of a DEHPA stock solution. Data sets were classified into three groups, 0.005-0.045 wt%, 0.05-0.45 wt%, and 0.5-4.5 wt%, respectively, in order to investigate the quality of PLS (Partial Least-Squares) model with each concentration ranges. Three replicate measurements were carried out for each data point. In the present study, PLS method[8] was used to find out the quantitative relationship between concentration of DEHPA in hexane and FT-NIR frequencies by reducing the dimensionality of the variable space from all available FT-NIR spectroscopic data, and identifying new meaningful underlying variables. Data preprocessing (Savitsky-Golay second-derivative) and PLS regression were performed with Unscrambler (v. 7.5, Camo AS, Oslo, Norway). Full cross-validation (Leave-one-out cross-validation) was employed for the development of a multivariate calibration model as well as the determination of the optimum number of factors in the calibration model in order to avoid overfitting of the models.
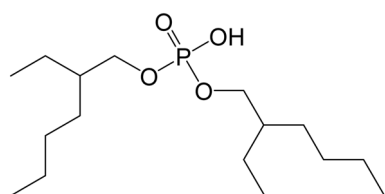

Figure 1. Chemical structure of DEHPA.

Results

The NIR spectra of DEHPA in hexane were shown in Figure 2(a). NIR absorbance increased remarkably with optical pathlength as well as DEHPA concentration. Interestingly, the absorbance difference depending on the DEHPA concentration in the region of 1,860-2,100 nm was much greater than in any other region, especially when the 24-mm cell was used. Figure 2(b) is a replot of Figure 2(a) in a wavelength window of 1,860-2,100 nm. In this spectral region, absorbance clearly increased with DEHPA concentration. But, it was difficult to use a classical univariate calibration method for DEHPA concentration in hexane due to the complex features of NIR spectrum such as the scattering effect on the spectra and the interferences from the solvent. PLS regression was, therefore, used for the development of a calibration model for DEHPA concentration in
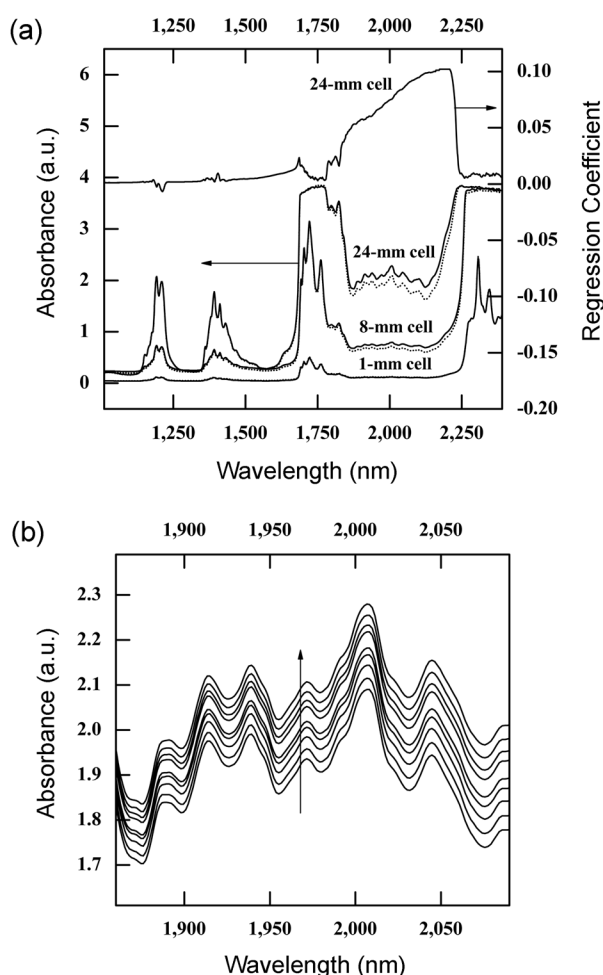

(a)

(b)

Figure 2. FT-NIR spectra of (a) 0.5 wt% DEHPA (dotted line), and 4.5 wt% DEHPA (solid line) using 1-mm-, 8-mm-, and 24-mm-pathlength glass cell, including the representative regression coefficient plot of 0.5-4.5 wt% DEHPA using 24-mm-pathlength glass cell for the comparison with FT-NIR spectrum, and (b) 0.5-4.5 wt% DEHPA using 24-mm-pathlength glass cell in the spectral range of 1,860-2,100 nm to show more clear differences with concentration. Arrow in (b) indicates the direction of DEHPA concentration increase from 0.5 to 4.5 wt%.

hexane to elucidate the relationship between NIR absorbance and concentration. PLS projects a data set from an original variable space to a new variable (latent variable or factor) space. The factors obtained are linear combination of the original variables. First few factors are selected by discarding uninformative factors, and results in a much smaller subspace. The number of factors in the calibration model with the best predictive performance was determined by full cross-validation until the root-mean-square error of cross-validation (RMSECV) had the lowest value. RMSECV gives a more realistic estimate of the prediction capability of the calibration than standard error of prediction (SEP) when the mean difference between measured value and reference value is not insignificant. Although RMSECV of the calibration model will generally decrease when an additional factor is introduced in the model, selection of too many factors introduces some noise in the model or spectral regions irrespective of the model prediction, and consequently deteriorates the quality of predicted model. Firstly, the effect of optical pathlength on the NIR spectrum and quality of the model were investigated for 0.5-4.5% DEHPHA. R^2 increased with optical pathlength. 1-mm-pathlength glass cell showed the lowest R^2, and 8-mm-pathlength glass vial showed the highest LV value. It is clear that the multivariate analysis using 24-mm-pathlength glass vial gives the most desirable model in all statistical respects (Table 1), *i.e.*, RMSECV, R^2, and LV, and therefore 24-mm-pathlength cell was used. PLS model over the entire concentration range (0.005-4.5 wt% DEHPA) was very satisfactory with only 1 LV, whereas the predictive ability of the model in a concentration window ranging from 0.05 wt% to 0.45 wt% DEHPA continues to improve until the regression model consists of 4 LVs, while R^2 in a concentration range of 0.005-0.045 wt% showed 0.518 and 0.399 for calibration set and validation set, respectively (Table 2). These R^2 values in a very low concentration range seem plausible because the spectral features from concentrated samples are clearer than those from diluted samples in NIR data as depicted in Figure 2. Here, the simple wavelength selection strategy was adopted to examine the improvement of the model quality in case of low-concentration samples, based on a manual selection of a single wavelength interval range by deleting uninformative variables with relatively small regression coefficient. There have been some reports on the advantages of the wavelength selection approach over the full-spectrum calibration.[9,10]

Table 1. Effect of optical pathlength on the statistical result of full cross-validation for 0.5-4.5 wt% DEHPA in hexane

Pathlength	LV[a]	Calibration set		Validation set	
		RMSEC (wt%)	R^2	RMSECV (wt%)	R^2
1 mm	2	0.1385	0.9940	0.1522	0.9927
8 mm	3	0.0970	0.9970	0.1303	0.9948
24 mm	1	0.0835	0.9978	0.0898	0.9975

[a]latent variables, which are optimum number of factors in PLS regression

Notes

Bull. Korean Chem. Soc. **2006**, Vol. 27, No. 12   2103

Table 2. Effect of spectral range on the statistical result of full cross-validation using 24 mm-glass cell

Range (wt%)	Full spectral range (830-2,500 nm)					Selected spectral range (1,860-2,100 nm)				
	LV[a]	Calibration set		Validation set		LV[a]	Calibration set		Validation set	
		RMSEC (wt%)	R^2	RMSECV (wt%)	R^2		RMSEC (wt%)	R^2	RMSECV (wt%)	R^2
0.005-4.5	1	0.1183	0.9960	0.1214	0.9957	1	0.1110	0.9964	0.1133	0.9963
0.005-0.045	1	0.0108	0.5175	0.0117	0.3992	3	0.0029	0.9734	0.0035	0.9626
0.05-0.45	4	0.0072	0.9984	0.0239	0.9823	2	0.0082	0.9979	0.0095	0.9972
0.5-4.5	1	0.0835	0.9978	0.0898	0.9975	1	0.0732	0.9983	0.0775	0.9981

[a]latent variables, which are optimum number of factors in PLS regression

There are a number of criteria for the wavelength selection from the individual spectral points,[11,12] and from spectral intervals,[13,14] including automatic wavelength selection. Statistical results of PLS regression using a wavelength window of 1,860-2,100 nm, selected based on the regression coefficient in Figure 1, showed the wavelength selection had a slight impact on the quality of a model prediction for 0.5-4.5 wt% DEHPA, while R^2 of 0.005-0.045 wt% DEHPA dramatically increased after a simple wavelength selection, from 0.399 to 0.947 for a validation data set, and from 0.518 to 0.975 for a calibration data set, respectively, although the number of LV also increased. RMSEC and RMSECV also became better than those obtained from the full-spectrum analysis. In case of intermediate concentration range (0.05-0.45% DEHPA), RMSEC and R^2 for a calibration data set and a validation data set slightly changed after a wavelength selection. It should be noted that RMSECV greatly decreased from 0.0239 to 0.0095, and the number of LV also decreased from 4 to 2, which indicates a much better model by removing the uninformative variables. The spectra pre-processing such as multiplicative scatter correction (MSC)[15] and the Savitsky-Golay derivative technique[16] were applied for further improvement of the model by the data pretreatment since the Savitsky-Golay second derivative for smoothing is usually useful when there are overlapping peaks in the original FT-NIR spectra to enhance the resolution, and for the baseline correction. But second-order differentiation is associated with the danger of losing some spectral information from these compounds. Therefore, a PLS regression is sometimes built from first-order-differentiated spectra. In our study, however, both first-order and second-order Savitsky-Golay treatments as well as MSC didn't improve

the quality of any model for low-concentration samples.

References

1. (a) Bond, A. H.; Dietz, M. L.; Chiarizia, R. *Ind. Eng. Chem. Res.* **2000**, *39*, 3442. (b) Cha, K. W.; Jeong, E. S. *Bull. Korean Chem. Soc.* **1994**, *15*, 9.
2. Hu, Z.; Gulari, E. *Biotechnol. Bioeng.* **1996**, *50*, 203.
3. Hu, Z.; Gulari, E. *J. Chem. Tech. Biotechnol.* **1996**, *65*, 45.
4. Meguro, Y.; Iso, S.; Sasaki, T.; Yoshida, Z. *Anal. Chem.* **1998**, *70*, 774.
5. Alexandratos, S. D.; Ripperger, K. P. *Ind. Eng. Chem. Res.* **1998**, *37*, 4756.
6. Gonzalez, M. P.; Saucedo, I.; Navarro, R.; Avila, M.; Guibal, E. *Ind. Eng. Chem. Res.* **2001**, *40*, 6004.
7. (a) Cho, S.; Chung, H.; Woo, Y.; Kim, H. *Bull. Korean Chem. Soc.* **2005**, *26*, 115. (b) Woo, Y.; Ahn, J.; Chun, I.; Kim, H. *Anal. Chem.* **2001**, *73*, 4964.
8. Gerlach, R. W.; Kowalski, B. R.; Wold, H. O. A. *Anal. Chim. Acta* **1979**, *112*, 417.
9. Xu, L.; Schechter, I. *Anal. Chem.* **1996**, *68*, 2392.
10. Rimbaud, D. J.; Walczak, B.; Massart, D. L.; Last, I. R.; Prebble, K. A. *Anal. Chim. Acta* **1995**, *304*, 185.
11. Westad, F.; Martens, H. *J. Near Infrared Spectrosc.* **2000**, *8*, 117.
12. Shaffer, R. E.; Arnold, M. A.; Small, G. W. *Anal. Chem.* **1996**, *68*, 2663.
13. Jiang, J. H.; Berry, R. J.; Siesler, H. W.; Ozaki, Y. *Anal. Chem.* **2002**, *74*, 3555.
14. Norgaard, L.; Saudland, A.; Wagner, J.; Nielsen, J. P.; Munck, L.; Engelsen, S. B. *Appl. Spectrosc.* **2000**, *54*, 413.
15. Isakson, T.; Naes, T. *Appl. Spectrosc.* **1988**, *42*, 1273.
16. Savitsky, M.; Golay, J. E. *Anal. Chem.* **1964**, *36*, 1627.