Organic Process Research & Development **2006**, *10,* 1101–1109

Full Papers

Measurement and Prediction of Solubility of Paracetamol in Water−Isopropanol Solution. Part 1. Measurement and Data Analysis

H. Hojjati and S. Rohani*

Department of Chemical and Biochemical Engineering, The University of Western Ontario, London, Ontario, Canada N6A 5B9

Abstract:

An attempt has been made to measure the concentration of paracetamol (98%, Aldrich Chemical Co. Inc., MO) in different solutions using an in situ ATR-FTIR device and chemometrics. A partial least-squares (PLS1) algorithm has been applied to construct two calibration models for paracetamol concentration, water mass percent, and temperature. The models and errors have been analyzed using validation data sets and diagnostic tools. The models are then used to evaluate the solubility of paracetamol (PA) in pure isopropanol, pure water, and isopropanol−water mixtures in the temperature range 5−40 °C. The solubility of paracetamol in isopropanol−water mixtures shows a maximum at almost 20 water mass percent. For some selected data points, the measured solubility by the FTIR is in good agreement with the corresponding solubility measured using the gravimetric method. Also the solubility in pure isopropanol and water is in reasonable agreement with the data reported in the literature.

Introduction

The production of a large number of chemicals, pharmaceuticals, and photographic materials is carried out in crystallizers. In crystallization processes, the supersaturation, the difference between the actual concentration and the saturation concentration (solubility), is the driving force for nucleation, growth, and agglomeration phenomena that influence the crystal size distribution, filterability, morphology, and polymorphic distribution of the product. Thus, an accurate measurement or estimation of supersaturation is needed to determine the crystallization yield and to design, operate, and control the process.

Several methods and analytical techniques have been proposed for the measurement of solute concentration in a crystallization process. The methods can generally be categorized as off-line, on-line, and in situ. In the on-line method, the samples are continuously withdrawn from the solution, passed through the measurement device and then recycled to the solution vessel. Delays between sampling and measuring, separation of liquid phase from crystal slurry, disturbing the solution, and changing the equilibrium state of the solution are the primary problems associated with the on-line method. On the other hand, in an in situ method, the concentration of the solute in the vessel is directly measured. There are several analytical devices that have been applied to measure the concentration. These include off-line gravimetric method,[1] on-line density meter,[2−6] and in situ conductivity meter.[7−10] The conductivity meter works for conducting slurries and is highly sensitive to temperature and may saturate at high solute concentrations. Other techniques include on-line calorimetry[11,12] on-line and in situ turbidity meters,[13] and in situ focused beam reflectance FBRM probe.[14] The latter has been used for the measurement of solubility and the metastable zone.

* Corresponding author. E-mail: rohani@eng.uwo.ca. Telephone: (519) 661-4116.

(1) Mohameed, H. A; Abu-Jdayil, B.; Khateeb, M. A. Effect of Cooling Rate on Unseeded Batch Crystallization of KCl. *Chem. Eng. Process.* **2002**, *41*, 297−302.
(2) Hojjati, H.; Rohani, S. Cooling and Seeding Effect on Supersaturation and Final Crystal Size Distribution (CSD) of Ammonium Sulphate in a Batch Crystallizer. *Chem. Eng. Process.* **2005**, *44*, 949−957.
(3) Tadayyon, A.; Rohani, S.; Bennett, M. Estimation of Nucleation and Growth Kinetics of Ammonium Sulphate from Transients of a Cooling Batch Seeded Crystallizer. *Ind. Eng. Chem. Res.* **2002**, *41*, 6181−6193.
(4) Qui, Y.; Rasmuson, A. C. Estimation of Crystallization Kinetic from Batch Cooling Experiments. *AIChE J.* **1994**, *40*, 799−812.
(5) Miller, S. M.; Rawlings, J. B. Model Identification and Control Strategies for Batch Cooling Crystallization. *AIChE J.* **1994**, *40*, 1312−1327.
(6) Gutwald, T; Mersmann, A. Batch cooling Crystallization at Constant Supersaturation: Technique and Experimental Results. *Chem. Eng. Technol.* **1990**, *13*, 229−237.
(7) Kubota, N.; Doki, N.; Yokota, M.; Sato, A. Seeding Policy in Cooling Crystallization. *Powder Technol.* **2001**, *121*, 31−38.
(8) Garcia, E.; Veesler, S.; Boistelle, R.; Hoff, C. Crystallization and Dissolution of Pharmaceutical Compounds: An Experiment Approach. *J. Cryst. Growth* **1999**, *198/199*, 1360−1364.
(9) Jagadesh, D.; Kubota, N.; Yokota, M.; Doki, N.; Sato, A. Seeding Effect on Batch Crystallization of Potassium Sulphate under Natural Cooling Mode and a Simple Design Method of Crystallizer. *J. Chem. Eng. Jpn.* **1999**, *32*, 514−420.
(10) Hložný, L.; Sato, A.; Kubota, N. On-Line Measurement of Supersaturation During Batch Cooling Crystallization of Ammonium Alum. *J. Chem. Eng. Jpn.* **1992**, *25*, 604−606.
(11) Monnier, O.; Févotte, G.; Hoff, C.; Klein, J. P. Model Identification of Batch Cooling Crystallization Through Calorimetry and Image Analysis. *Chem. Eng. Sci.* **1997**, *52*, 1125−1139.
(12) Févotte, G.; Klein, J. P. Calorimetric Methods for the Study of Batch Crystallization Processes. *Trans. IChemE* **1996**, *74*, Part A, 791−796.
(13) Moscosa-Santillán, M.; Bals, O.; Fauduet, H.; Porte, C.; Delacroix, A. Study of Batch Crystallization and Determination of an Alternative Temperature Profile by On-Line Turbidity Analysis: Application to Glycine Crystallization. *Chem. Eng. Sci.* **2000**, *55*, 3759−3770.
(14) Barrett, P.; Glennon, B. Characterizing the Metastable Zone Width and Solubility Curve Using LasenTec FBRM and PVM. *Trans. IChemE* **2002**, *80*, Part A, 799−805.

10.1021/op060073o CCC: $33.50
Published on Web 09/13/2006

Recently, the feasibility of in situ attenuated total reflection Fourier transform infrared (ATR-FTIR) spectroscopy with the help of some useful mathematical tools, chemometrics, has shown reliability and potential for the real-time measurement and control of batch crystallizers. For instance, the solubilities and supersaturation measurements of aqueous citric acid,[15] maleic acid in water,[16] isotactic polystyrene,[17] bifenox in methanol, isoproturon in ethanol, a pharmaceutical compound in ethanol,[18] potassium dihydrogen phosphate (KDP) in water,[19] and paracetamol in water[20] were investigated using ATR-FTIR spectroscopy. All the results showed that the method is sensitive and accurate enough to measure small changes in concentration and is applicable to multicomponent systems and monitoring in situ polymorphic transformation.

In this study, the solubility of paracetamol in pure isopropanol, pure water, and isopropanol−water mixtures is measured using an *in situ* AIR-FTIR spectroscopy and the chemometric technique (partial least-squares).

Materials and Methods

Besides its applications as a prescribed drug for pain relief or reduction of fever and as an intermediate component for photographic material production, paracetamol (acetaminophen or 4-acetamidophenol) has been used as a model compound for many pharmaceutical crystallization studies. Paracetamol (98%, Aldrich Chemical Co. Inc., MO), isopropanol (99.5%, EMD Chemical Inc., NJ), and deionized water supplied in our lab were used in this study. The relatively low purity of paracetamol used in this study imposes limitations on the experimental results. However, higher-purity paracetamol would have been very costly because of the large amounts needed to conduct the solubility and control experiments.

Experimental Setup. All experiments were carried out in a 100-mL jacketed flask (Bellco, NJ). A magnetic stirrer was employed for gentle mixing of the solution. The solution temperature was measured by a thermocouple (J type) with 0.1 °C resolution. The temperature was controlled by manipulating the set-point temperature of a cooling water circulator (F-32, Julabo LABORTECHNIK GMBH, Germany) using a two-level cascade controller. A LabView (National Instruments, TX) hardware/software module was employed for data acquisition and control. All spectra were monitored and analyzed using the ATR-FTIR (Hamilton Sundstrand, CA) probe (DMD-270 diamond ATR immersion probe) and its software (FX-90). The ATR-FTIR was calibrated using solutions with known concentrations of paracetamol at different temperatures. The solutions were prepared by adding an appropriate mass of paracetamol and solvent (pure isopropanol, pure water, or isopropanol−water mixtures) to the flask, covering the top, keeping the solution temperature constant, and mixing until the paracetamol was totally dissolved at a given temperature. The spectra and concentrations were recorded for all runs.

To measure the saturation concentration of paracetamol in each solvent individually, 80 g of the solvent and an excess amount of solids were added to the flask at a given temperature, and the solution was stirred until no change in the spectra was observed for 30 min (10 consecutive readings). This procedure was repeated for eight temperatures between 5 and 40 °C. In the case of measuring the solubility of paracetamol in isopropanol−water mixtures, some data points were also measured using a gravimetric technique. In the gravimetric technique, a preheated syringe equipped with 0.2-μm syringe filter was used to withdraw a clear solution from the slurry. After reaching room temperature, the weight of the sample before and after evaporation of the solvent was measured using a precision (0.1 mg) balance (AX205, Mettler Toledo, Switzerland). The measured values were used to calculate the solubility of paracetamol in the cosolvent.

Chemometrics. Chemometrics refers to the statistical processing of analytical chemistry data with various numerical techniques to handle, interpret, and predict the data. A combination of principal component factor analysis and multiple linear regression allows displaying multivariable regression and extracting useful information by reducing the dimensionality. Principle component regression (PCR) and partial least-squares (PLS) are the most important techniques in this respect; however, the PLS is often considered as an effective methodology for doing quantitative analysis on complex systems and as the major regression technique for multivariable data. Two major algorithms for PLS, PLS1 and PLS2, have been reported in the literature. The PLS1 algorithm calibrates each component individually, while PLS2 uses multiple components simultaneously. In general, when nonlinearity, rather than random noise, is the major problem, then it may be advantageous to use separate PLS1 modeling for each independent variable for the final calibration results.[21] In our case, the PLS1 algorithm proved to be more accurate than the PLS2 for all data sets examined and was chosen for the regression.

In the chemometric methods all data have to be organized in sets of matrices, namely, *training* and *validation* sets. After setting up a chemometric method and before running the calibration for unknown samples, the model and errors should be analyzed using some diagnostic tools. Both the validation set and diagnostic tools need to detect the outliers and to

(15) Dunuwila, D. D.; Carrol, L. B., II; Berglund, K. A. An Investigation of the Applicability of Attenuated Total Reflection Infrared Spectroscopy for Measurement of Solubility and Supersaturation of Aqueous Citric Acid Solution. *J. Cryst. Growth* **1994**, *137*, 561−568.

(16) Dunuwila, D. D.; Berglund, K. A. ATR-FTIR Spectroscopy for in situ Measurement of Supersaturation. *J. Cryst. Growth* **1997**, *179*, 185−193.

(17) Kimura, T.; Ezure, H.; Tanaka, S.; Ito, E. In-Situ FTIR Spectroscopic Study on Crystallization Process of Isotactic Polystyrene. *J. Polym. Sci., Part B: Polym. Phys.* **1998**, *36*, 1227−1233.

(18) Lewiner, F.; Kelin, J. P.; Puel, F.; Févotte, G. On-Line ATR-FTIR Measurement of Supersaturation during Solution Crystallization Processes: Calibration and Applications on Three Solute / Solvent Systems. *Chem. Eng. Sci.* **2001**, *56*, 2069−2084.

(19) Togkalidou, T.; Fujiwara, M.; Patel, S.; Braatz, R. D. Solute Concentration Prediction Using Chemometrics and ATR-FTIR Spectroscopy. *J. Cryst. Growth* **2001**, *231*, 534−543.

(20) Fujiwara, M.; Chow, P. S.; Ma, D. L.; Braatz, R. D. Paracetamol Crystallization Using Laser Backscattering and ATR-FTIR Spectroscopy: Metastability, Agglomeration, and Control. *Cryst. Growth Des.* **2002**, *2*, 363−370.

(21) Martens, H.; Næs, T. *Multivariate Calibration*; John Wiley & Sons: UK, 1989.

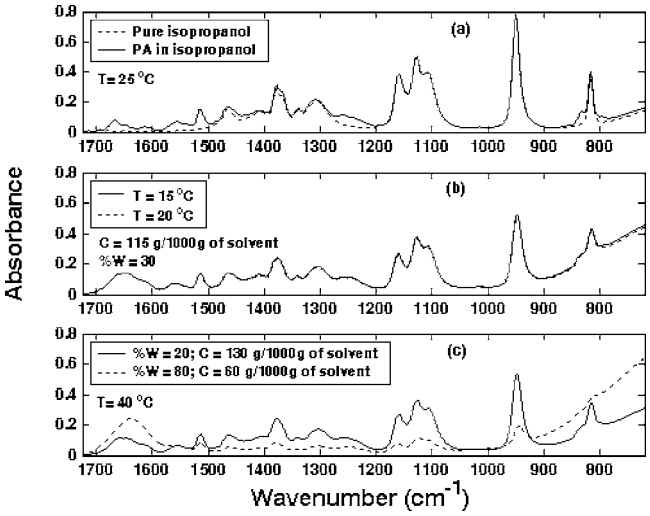

Figure 1. **ATR-FTIR absorption spectrums of paracetamol (PA) solution; (a) in isopropanol at 25 °C for 0 and 110.5 g PA/1000 g of solvent; (b) in isopropanol−water mixture (30 water mass percent on a solute-free basis) for constant concentration (115 g PA/1000 g of solvent) and different temperatures (15 and 20 °C); (c) in isopropanol−water mixture (80 and 20 water mass percent based on solute free) at 40 °C for different concentrations (60 and 130 g PA/1000 g of solvent).**

find the optimum model. There are various diagnostic tools; however, *predicted* vs *experimental*, *correlation coefficient squared* (R^2), and *residuals analysis* were used in this study.

Analyzing the spectra and selecting an appropriate range of wavenumbers are needed for the calibration model. Preprocessing is an integral part of the chemometrics to remove or minimize the unwanted features such as noise and background scatter. In this investigation, two different preprocessing methods were considered: *smoothing* using the Savitzky−Golay algorithm with 11-point averaging and *zero baseline correction*. Also, by removing the unwanted information and narrowing the range of wavenumber, the quality of the calibration models is enhanced, and computational efficiency is improved. Figure 1 shows the effects of temperature, paracetamol concentration, and water mass percent (solute free) on the FTIR absorbance spectra of paracetamol solutions for the entire wavenumber regions between 720 and 1725 cm^{-1}. The spectra are plotted after the preprocessing. To find the best region representing the paracetamol concentration, the spectra of pure isopropanol and of paracetamol (PA) in isopropanol (C = 110.5 g PA/1000 g of solvent) at 25 °C were plotted (Figure 1a). The major differences between two spectra are in the wavenumber range 1200−1725 cm^{-1}. Thus, the wavenumber ranges 1200−1285 cm^{-1} and 1490−1725 cm^{-1} were selected for the calibration model of PA in both pure isopropanol and pure water. At constant concentration, the temperature effect on PA spectra is shown in Figure 1b. Although the zero baseline correction compensates for the temperature effect in some regions, there are still some deviations between two spectra in other wavenumber regions that should be considered in the calibration model. In the case of paracetamol in isopropanol−water mixtures, Figure 1c shows the effects of paracetamol concentration and water mass percent (solute free) on the FTIR spectra at constant temperature. In the entire region between 720 and 1725 cm^{-1}, there are peaks representing paracetamol concentration and solvent composition changes. Thus, the entire range, 720 − 1725 cm^{-1}, was chosen for the calibration model of paracetamol in isopropanol−water mixtures.

In the case of the crystallization of paracetamol in isopropanol−water mixtures, several variables (e.g., paracetamol and water concentrations) must be predicted from the same input data (e.g., absorbance at various wavenumbers). Considering the temperature effects and using the PLS1 algorithm, eq 1 was applied for the calibration procedure.

$$C = a_0 + a_T T + \sum a_i A_i \tag{1}$$

where A_i, a_i, T, and C are the absorbance of IR spectra, adjustable parameters, temperature, and concentration, respectively. If the model were constructed to classify temperature as the third predicted variable, the eq 1 could be simplified as:

$$\bar{C} = a_0 + \sum a_i A_i \tag{2}$$

where $\bar{C}$ is the matrix of paracetamol concentration, water mass percent (solute free), and temperature.

For the *training set*, we used the rule of 10 that determines the minimum required sample size to create a calibration.[22] Thus, 10 times the number of components gives the minimum number of sample size of training sets. The *validation sample* numbers were at least half of *training set* numbers.

The desirable operating range is between the solubility and metastable curves. Therefore, for the calibration, the ranges 60−180, 5−25, and 10−350 g PA/1000 g of solvent were used for the concentration of paracetamol in isopropanol, water, and isopropanol−water mixtures, respectively. The temperature was varied between 3 and 45 °C. The solvent composition was over the range of 10−90% (water mass percent based on solute free).

Results and Discussion

Solubility of Paracetamol in Isopropanol−Water Mixture. The spectra of paracetamol in isopropanol−water mixture in the range 720−1720 cm^{-1} were used to construct a calibration model. The model relates the concentration and temperature to the IR spectra. Solubility was measured over the range of 5−40 °C. The solvent composition was changed between 10 and 90 water mass percent (solute free). On the basis of the "rule of 10" and three components involved in the model, the minimum data points needed for a solid calibration were 30 samples. Considering the validation set size and the outliers, 65 samples were measured for the calibration model, 45 samples for the *training set*, and 20 samples for the *validation set*. The data space was designed according to the procedure described above. The data points of *validation set* were selected randomly. The method was based on selection of data points within the entire range of

(22) Kramer, R. *Chemometric Techniques for Quantitative Analysis*; Marcel Dekker: New York, 1998.

solute concentration, but the intermediate solute concentrations were weighted.

Determination of Factor Numbers and Outliers. The first step in PLS1 analysis is to determine the factor numbers of each component individually. Predicted residual error sum of squares (*PRESS*) and statistical *F*-test are sufficient criteria to determine the optimum number of factors.

$$PRESS_{comp} = \sum_{n} (C_{pre} - C_{exp})^2 \quad (3)$$

where C_{pre} and C_{exp} represent the predicted and experimental concentration of each component, respectively, and n is the number of samples. Forty five samples were selected randomly for the *training set*. The maximum factor number 25 was chosen for each component. The *PRESS* values were calculated for all factor numbers of each component. As the difference between the minimum *PRESS* and the other *PRESS* values becomes smaller, the probability that each additional factor provides significant improvement to the model becomes smaller. Thus, the ratio of all *PRESS* values to the smallest *PRESS* value (P_{ratio}) and $F_{\alpha,n,n}$ values (F distribution with n and n degrees of freedom) at 95% ((1 − α) × 100) confidence limit were used in this study for a good choice of the statistical cut-off point of factor numbers.

$$P_{ratio} = \frac{PRESS_i}{PRESS_{min}} \quad (4)$$

The smallest factor number was chosen in such a way that the P_{ratio} of that particular *PRESS* becomes less than $F_{\alpha,n,n}$.[23]

The possible outliers were detected using the residuals of data points of each component and residual standard deviation of the corresponding component ($\sigma_{r,comp}$). If the absolute value of residuals is greater than $2\sigma_{r,comp}$, the corresponding data point was considered as an outlier.[21] Then the outliers were eliminated from the data set to provide a new training set. The factor number determination procedure was repeated for the new training set until all possible outliers were removed.

Two different models were investigated using this procedure. The first model was constructed based on eq 1 considering paracetamol concentration and water mass percent as predicted variables. Graphs a and b of Figure 2 show the semilogarithmic plots of *PRESS* value versus factor numbers of PA concentration and water mass percent for the final training set, respectively. Applying the P_{ratio} statistic, the factor numbers for each component were determined. The PLS1 regression was then conducted for the selected factor numbers, and the regression model was used to predict the solubility of paracetamol. Note that the *PRESS* values of both components were relatively too high (order of magnitude 10^4) for low factor numbers. This was due to the fact that the concentration and water mass units applied for the regression were in g PA/1000 g of solvent and percent, respectively. However when the factor numbers were increased, the *PRESS* values decreased rapidly to 10^2 and 10^0 for PA concentration and water mass percent, respectively. In the second model, temperature was considered as the third predicted variable, and the regression model was built using eq 2. The same procedure was applied to construct the second calibration model. Although the outlier data points of this model were different from those of the previous model, the factor number analysis of PA concentration and of water mass percent showed almost the same results for both models. For the final training set, the semilogarithmic plot of *PRESS* for the temperature is also shown in Figure 2c. This model was also used for the solubility prediction.

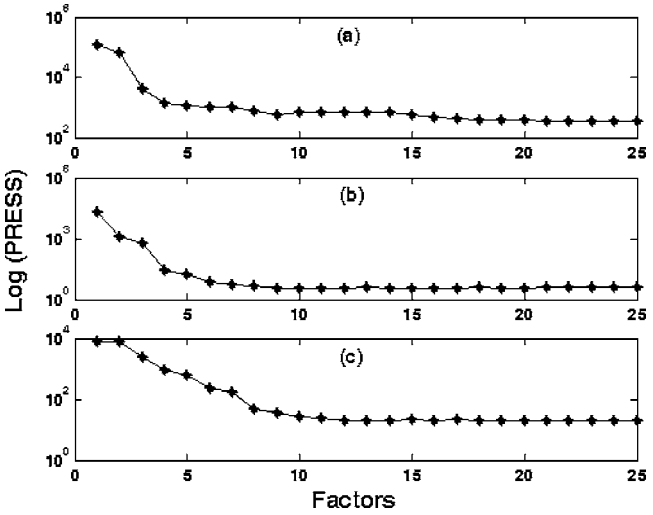

Figure 2. **Semilogarithmic plot of *PRESS* versus factor numbers for the calibration model of paracetamol in isopropanol−water mixtures using the PLS1 algorithm: (a) for paracetamol concentration, (b) for water mass percent, (c) for temperature.**

Predicted versus Experimental Solubility Data. The calibration constructed on the first model was used to predict the paracetamol concentration and water mass percent of the training set. This diagnosis gives a complete description of the calibration model. It also helps predict how well the calibration set will analyze the component concentration of unknown samples. Graphs a and b of Figure 3 show the predicted versus experimental PA concentration and water mass percents of the *training set*, respectively. Similar results were obtained when temperature was considered as the third predicted variable. In Figure 3c the predicted temperature is plotted versus experimental temperature to examine the predictive capability of the second model. In both models and for three components, all data points of the *training set* were closely distributed to the diagonal line. The percentage errors calculated for all components were less than 1%. It can also be inferred that the outliers were well recognized and eliminated from the data set. Note that for any outlier the difference between the predicted and experimental value is relatively high, and the data points appear in an area that is far from the diagonal line. The correct match is an indication of an accurate calibration model; however, a good fit to the *training set* does not guarantee that the model has a good predictive capability for the unknown samples.

(23) Haaland, D. M.; Thomas, E. V. Partial Least-Squares Methods for Spectral Analyses. 1. Relation to Other Quantitative Calibration Methods and the Extraction of Qualitative Information. *Anal. Chem.* **1988**, *60*, 1193−1202.

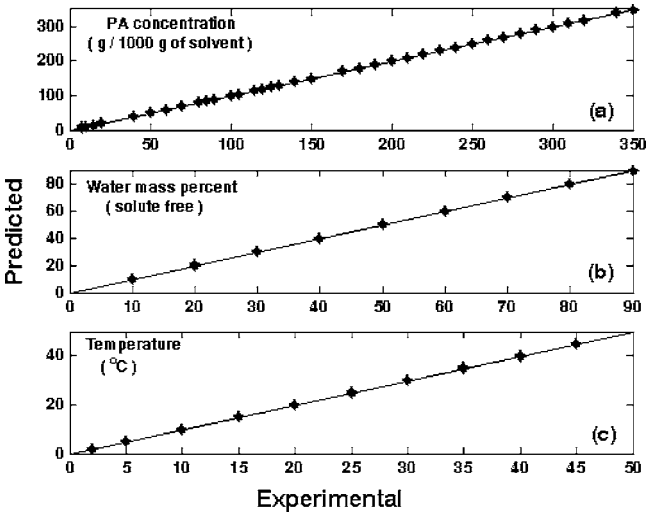

Figure 3. **Predicted versus experimental value of each component for the training set.**

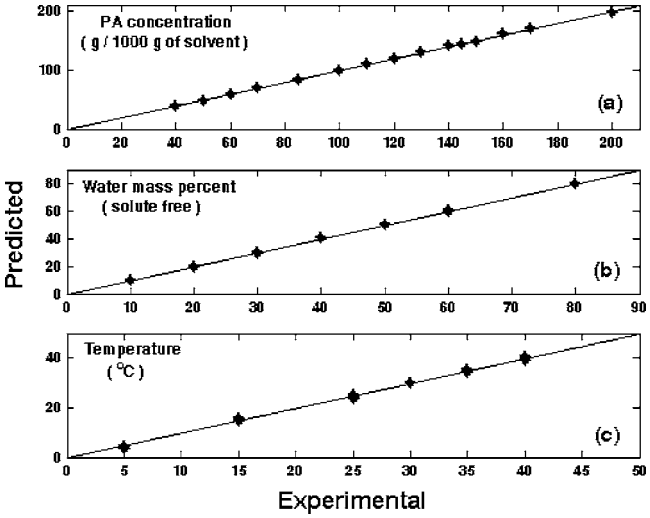

Figure 4. **Predicted versus experimental value of each component for the validation set.**

Table 1. **Square of correlation coefficient (R^2) for each component and for both models**

	PA concentration	water mass %	temperature
first model	0.9999	0.9999	–
second model	0.9996	0.9998	0.9994

Correlation Coefficient (R^2). Instead of covariance, the correlation coefficient is widely used between the predicted and experimental values to determine how closely they match. For both models, the squares of correlation coefficients of all components were calculated using eq 5 and are listed in Table 1.

$$R^2 = \left[\frac{\sum (C_{\text{pre}} - \bar{C})(C_{\text{exp}} - \bar{C})}{\sqrt{\sum (C_{\text{pre}} - \bar{C})^2 \sum (C_{\text{exp}} - \bar{C})^2}} \right]^2 \qquad (5)$$

where $\bar{C}$ is mean value of each individual component. Although the R^2 values for the second model were relatively lower than those calculated for the first model, all values of both models were close to 1, showing the high predictive capability of models.

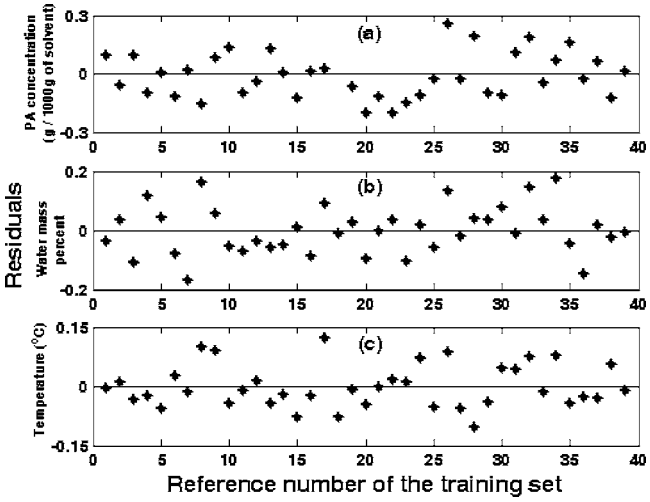

Figure 5. **Component residuals of the training set (a), (b) for the first model, and (c) for the second model.**

Validation of Models. Twenty data points, selected randomly, were used to validate both models. Graphs a−c of Figure 4 show the predicted versus experimental values for PA concentrations and water mass percents of the first model, and temperature of the second model, respectively.

It is apparent that the component values were predicted very well. All data points were located on or close to the diagonal line, indicating no considerable amount of scatter between the predicted and experimental values. The percentage absolute differences between the predicted and experimental values calculated for all data points were less than 1% (the maximum error percentages were 0.99, 0.85, and 0.74% for PA concentration, water mass percent, and temperature, respectively).

Residuals. An alternative and useful way of looking at predicted results is to plot the *residuals* for each component. The plots can help analyze the reliability of the calibration models and show how well the outliers are identified.

The *residuals* for each component were calculated and plotted versus the reference number of the *training set* in Figure 5. The *residuals* for a reference data set were very low in comparison to other points, indicating that there was no significant outlier among the samples.

Solubility Measurement. The temperature was changed from 5 to 40 °C every 5 °C. At each constant temperature, the water mass percent (solute free) was varied from 10% to 90% with a step of 10%. Experimental data and the first calibration model were used to estimate the saturation concentration of the solutions. The results of solubility measurements are given in Table 2. Using the first model, the predicted water mass percents are also listed in Table 2 to compare with the experimental values. The predicted temperatures using the second model are also listed in Table 2. The maximum and average deviations from the experimental values were 2.0% and 0.5% for the water mass percent, and 8.0% and 2.2% for the temperature, respectively.

Table 2. **Solubility data for paracetamol in pure isopropanol, pure water, and isopropanol−water mixtures**

water mass %	solubility (g/1000 g of solvent) 40 (°C)	35 (°C)	30 (°C)	25 (°C)	20 (°C)	15 (°C)	10 (°C)	5 (°C)
0	**169.69**	**151.57**	**135.74**	**122.34**	**110.70**	**97.80**	**88.32**	**79.04**
10	**280.03**	**250.71**	**223.11**	**198.95**	**178.55**	**158.07**	**139.47**	**126.77**
	(10.0)[a]	(10.1)	(9.9)	(9.8)	(10.1)	(10.0)	(10.2)	(9.8)
	(39.4)[b]	(35.1)	(29.8)	(25.3)	(19.9)	(15.1)	(10.2)	(4.9)
20	**331.72**	**300.63**	**274.01**	**247.5**	**224.95**	**203.52**	**183.45**	**168.50**
	(19.9)	(20.1)	(20.0)	(19.9)	(19.9)	(20.2)	(19.9)	(20.1)
	(41.0)	(35.0)	(29.2)	(24.9)	(20.1)	(14.8)	(10.1)	(5.3)
30	**348.46**	**312.54**	**282.61**	**254.31**	**226.68**	**205.10**	**181.67**	**163.94**
	(29.9)	(29.7)	(30.4)	(29.8)	(30.1)	(29.8)	(30.0)	(30.2)
	(39.8)	(34.9)	(30.4)	(24.4)	(19.6)	(15.5)	(10.6)	(4.9)
40	**341.83**	**303.33**	**268.81**	**239.80**	**212.27**	**187.28**	**166.44**	**145.57**
	(40.4)	(39.7)	(39.8)	(40.2)	(39.8)	(39.7)	(40.1)	(39.7)
	(40.3)	(35.4)	(29.6)	(25.6)	(19.3)	(14.8)	(9.5)	(5.1)
50	**302.26**	**263.19**	**230.02**	**199.05**	**172.94**	**150.31**	**131.67**	**114.16**
	(50.3)	(49.7)	(49.8)	(49.8)	(50.1)	(49.6)	(49.7)	(50.1)
	(39.5)	(34.9)	(29.6)	(25.0)	(20.4)	(14.8)	(9.5)	(5.2)
60	**247.44**	**213.08**	**182.58**	**154.74**	**133.15**	**114.03**	**95.96**	**82.55**
	(59.8)	(60.3)	(59.6)	(60.2)	(59.7)	(60.0)	(59.9)	(59.8)
	(39.4)	(35.2)	(30.4)	(25.3)	(19.5)	(15.1)	(9.9)	(4.7)
70	**183.25**	**152.03**	**125.34**	**105.53**	**87.01**	**72.28**	**58.82**	**48.94**
	(70.3)	(69.5)	(69.7)	(70.0)	(70.1)	(69.7)	(69.6)	(69.8)
	(40.4)	(33.9)	(29.9)	(25.6)	(20.7)	(14.1)	(10.6)	(4.8)
80	**107.67**	**86.50**	**70.45**	**56.53**	**45.69**	**36.95**	**29.24**	**23.58**
	(79.8)	(80.4)	(80.0)	(79.9)	(79.7)	(79.8)	(80.1)	(80.2)
	(39.4)	(36.0)	(30.1)	(24.8)	(19.9)	(15.4)	(10.4)	(5.4)
90	**61.14**	**47.43**	**37.66**	**30.01**	**23.25**	**18.32**	**14.20**	**11.44**
	(89.8)	(89.9)	(90.4)	(90.3)	(89.4)	(89.7)	(90.0)	(89.6)
	(39.0)	(35.1)	(29.1)	(25.2)	(20.8)	(14.8)	(10.6)	(5.2)
100	**24.75**	**20.80**	**17.36**	**14.98**	**12.22**	**10.71**	**9.12**	**8.09**

[a] Predicted water mass percent (solute free) using the first model. [b] Predicted temperature (°C) using the second model.

Table 3. **Comparison of measured solubility using gravimetric technique and predicted solubility using ATR-FTIR technique for paracetamol in isopropanol−water mixtures**

water mass %	*T* (°C)	measured solubility (g/1000 g of solvent)	predicted solubility (g/1000 g of solvent)	% error
20	10	188.16	183.45	−2.5
20	15	199.63	203.52	+1.9
20	20	226.71	224.95	−0.8
20	25	252.80	247.50	−2.1
40	10	162.96	166.44	+2.1
40	15	190.81	187.28	−1.8
40	20	210.76	212.27	+0.7
60	20	129.98	133.15	+2.4
60	25	155.92	154.74	−0.8
60	30	180.55	182.58	+1.1
80	25	57.99	56.53	−2.5
80	30	70.17	70.45	+0.4
80	35	87.10	86.50	−0.7

Comparing the experimental and predicted values suggests good prediction by both models.

To evaluate the predicted values, the solubility of paracetamol in the cosolvent was measured gravimetrically for some selected data points. The results are given in Table 3. The corresponding value predicted by the ATR-FTIR method along with the error percentage is also presented in Table 3. The results show good agreement between both techniques.

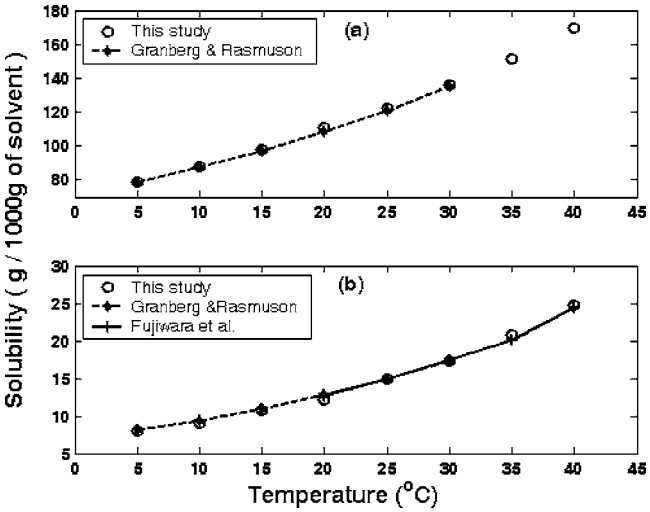

Figure 6. **Solubility of paracetamol in (a) pure isopropanol and (b) pure water.**

Solubility of Paracetamol in Pure Isopropanol and Pure Water. The solubility of paracetamol in pure isopropanol and pure water was also estimated. The wavenumber ranges 1200−1285 cm^{-1} and 1490−1725 cm^{-1} were selected for both solvents. Using eq 1 as the calibration model, the same procedure was applied to measure the solubility of PA in the pure solvents. All steps including factor numbers determination, outliers detecting, and diagnosis were conducted. Only the results of solubility of PA in both pure solvents are listed in Table 2. To evaluate the predictive capability of the models, the results of this study are compared in Figure 6 with the literature. Using an off-line gravimetric method, Granberg and Rasmuson[24] reported the solubility of paracetamol in 26 different pure solvents including isopropanol and water over the temperature range −5 to +30 °C. Fujiwara, Chow, Ma, and Braatz[20] measured the solubility of PA in water in the range from 20 to 50 °C using an in situ ATR-FTIR probe and chemometric techniques. The solubility measured in this work is in good agreement with those reported in the literature.

Analyzing the Solubility Data. Paracetamol has a simple molecular structure showing polar property. Barra, Lescure, Doelker, and Bustamate[25] studied the hydrogen bonding ability of paracetamol as a donor and an acceptor. High solubility of paracetamol in polar protic solvents such as water and isopropanol is expected. Isopropanol and water molecules have high ability to form hydrogen bonds with paracetamol molecules. The reason for the greater solubility of paracetamol in alcohols than in water lies in the balance of nonpolar and polar portions of the molecule.

According to *regular-solution* theory, the solubility curve of a solid in a cosolvent shows a maximum at a specific solvent composition (the effective solubility parameter of a cosolvent is a function of the solvent composition), if the

(24) Granberg, R. A.; Rasmuson, A. C. Solubility of Paracetamol in Pure Solvents. *J. Chem. Eng. Data* **1999**, *44*, 1391−1395.
(25) Barra, J.; Lescure, F.; Doelker, E.; Bustamate, P. The Expanded Hansen Approach to Solubility Parameter. Paracetamol and Citric Acid in Individual Solvents. *J. Pharm. Pharmacol.* **1997**, *49*, 644−651.

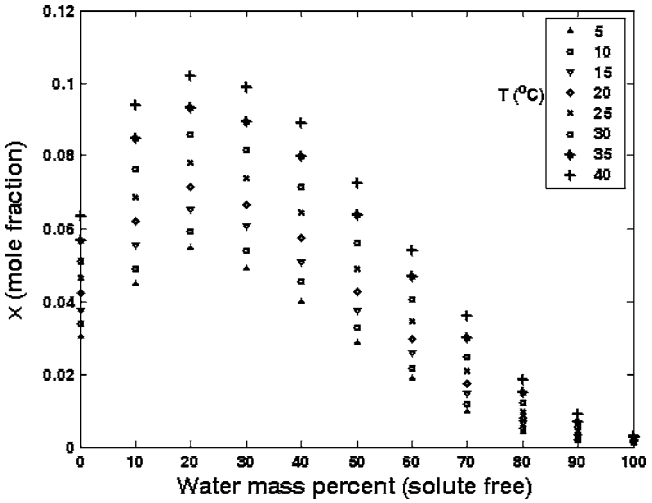

Figure 7. **Solubility of paracetamol in isopropanol−water mixtures at temperature range 5−40 °C.**

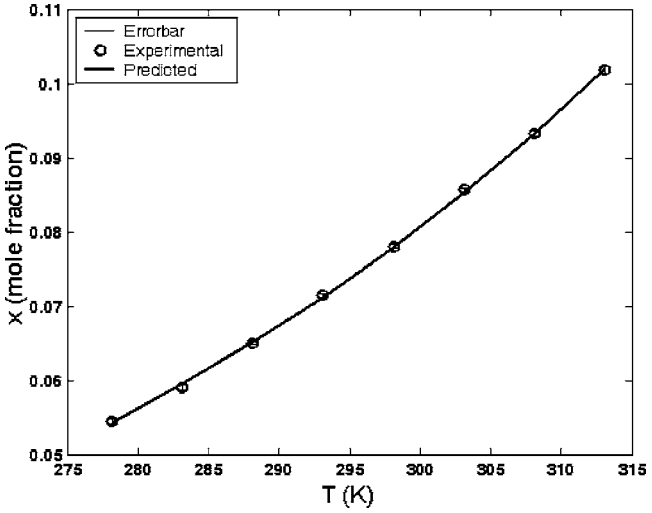

Figure 8. **Fitted curve of paracetamol solubility at 20 water mass percent (solute-free basis) using eq 6.**

solubility parameter of solute is between the solubility parameters of solvents. Using partial solubility parameters, Barra, Lescure, Doelker, and Bustamate[25] reported 27.80 (J $cm^{-3})^{1/2}$ for the solubility parameter of paracetamol (with the confidence level <0.001), whereas the solubility parameters of isopropanol and water are 23.575 and 47.813 (J $cm^{-3})^{1/2}$, respectively. Thus, it can be inferred that at any temperature the solubility of paracetamol in isopropanol−water mixture has a maximum at a specific solvent composition. Using the data in Table 2, the solubility of paracetamol at different temperatures was calculated in mole fractions and plotted versus solvent composition (solute-free basis) in Figure 7. At any constant temperature, there is a maximum solubility at 20 water mass percent, suggesting that the solubility parameter of paracetamol is between the solubility parameters of isopropanol and of water, much closer to the solubility parameter of isopropanol.

Mullin[26] proposed several correlations to relate the solubility, in mole fractions, to absolute temperature. Among them, the following equation provides the best fit for the temperature dependence of solubility of paracetamol on the solvents.

Table 4. **Adjustable parameters of eq 6 for the solubility of paracetamol at different constant solvent compositions along with confidence intervals (CI) for each parameter (at 95% confidence limit) and the corresponding squares of correlation coefficient**

water mass %	β_1	$\pm CI_1$	$\beta_2 \times 10^2$	$\pm CI_2 \times 10^4$	R^2
0	9.6097×10^{-5}	1.0700×10^{-5}	2.0719	3.710	0.9997
10	1.1896×10^{-4}	1.5480×10^{-5}	2.1312	4.321	0.9996
20	3.6380×10^{-4}	3.1936×10^{-5}	1.8001	2.922	0.9998
30	1.9383×10^{-4}	$2,7233 \times 10^{-5}$	1.9913	4.670	0.9995
40	7.6503×10^{-5}	1.0673×10^{-5}	2.2558	4.628	0.9996
50	1.8398×10^{-5}	1.6816×10^{-5}	2.6445	3.024	0.9999
60	4.2163×10^{-6}	7.2623×10^{-7}	3.0218	5.684	0.9997
70	3.7478×10^{-7}	6.8948×10^{-8}	3.6626	6.047	0.9998
80	3.0066×10^{-8}	4.9491×10^{-9}	4.2567	5.393	0.9999
90	2.7825×10^{-9}	7.1217×10^{-10}	4.7814	8.364	0.9998
100	8.958×10^{-8}	4.1944×10^{-8}	3.3184	15.00	0.9982

$$\ln x = \ln \beta_1 + \beta_2 T \quad \text{or} \quad x = \beta_1 e^{\beta_2 T} \tag{6}$$

where x is in mole fractions, T is temperature in K, and β_1 and β_2 are adjustable parameters that can be estimated from the experimental data. Using a nonlinear regression algorithm, the solubility data of paracetamol at constant solvent composition were fitted by eq 6. Figure 8 shows good correlation between the experimental data and the fitted solubility curve. The error bars calculated at the 95% confidence limit are also plotted.

All adjustable parameters for various solvent compositions are listed in Table 4. The confidence interval calculated at the 95% level for each parameter and the corresponding square of correlation coefficients is also given in Table 4. Except for the results of the paracetamol solubility in pure water, the other results show good agreement between the experimental and the estimated solubility using eq 6.

For the solubility of the polar compounds in water, the following equation has been proposed for the temperature dependence of the solubility in mole fractions.[27]

$$\ln x = -\beta_1 \frac{1}{T} + \beta_2 \ln T + \beta_3 \tag{7}$$

where

$$\beta_1 = \frac{a}{R}$$

$$\beta_2 = \frac{b}{R}$$

R is the universal gas constant in J mol^{-1} K^{-1}, and the coefficients of a and b can be used to calculate the apparent enthalpy (ΔH_s^*) of solution and the apparent heat

(26) Mullin, J. W. *Crystallization*, 3rd ed.; Butterworth-Heinemann: UK, 1993.
(27) Grant, D. J. W.; Higuchi, T. *Solubility Behavior of Organic Compounds*; John Wiley & Sons: New York, 1990.

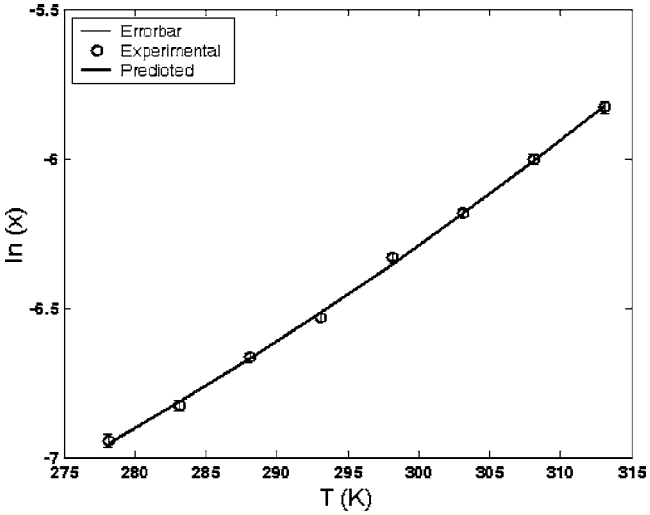

Figure 9. **Fitted curve of paracetamol solubility in pure water using eq 7.**

capacity of paracetamol (C_p^*).

$$\Delta H_s^* = a + bT \tag{8}$$

$$c_p^* = b \tag{9}$$

Using a nonlinear regression, the adjustable parameters were estimated for the solubility data of paracetamol in water.

$$\ln x = \frac{10495.900}{T} + (45.11344 \ln T) - 298.59288 \tag{10}$$

Comparison between the measured and calculated solubility is presented in Figure 9. The error bars are plotted at the 95% confidence limit. The square of correlation coefficient is 0.9991 showing a better model for paracetamol solubility in water in comparison with eq 6. The confidence intervals for the adjustable parameters are not reported in this section because they show very wide ranges. This statistical value is related inversely to degree of freedom and directly to *t* distribution if the numbers of data are large enough. Also the degree of freedom is defined as the difference between the number of data (n) and the number of parameters (p). In this case, the n and p values are 8 and 3, respectively, and thus neither the n nor the degree of freedom is large enough to rely on the confidence interval of adjustable parameters as a statistical criterion.

Using eq 7 and the estimated adjustable parameters, the apparent enthalpy of solution and the apparent heat capacity of paracetamol can be estimated.

$$\Delta H_s^* = 24.30 \text{ kJ mol}^{-1} \qquad c_p^* = 374.2 \text{ kJ mol}^{-1} \text{ K}^{-1}$$

The apparent enthalpy of solution is in good agreement with those reported by Fujiwara, Chow, Ma, and Braatz,[20] 24.05 kJ mol^{-1}, and Grant, Mehdizadeh, Chow, and Fairbrother,[28] 21.7 kJ mol^{-1}.

Conclusions

An approach to measure paracetamol concentration and solubility in pure water, isopropanol and water−isopropanol mixture using an in situ ATR-FTIR spectroscopy and chemometric technique is investigated. On the basis of two calibration models constructed for the prediction of two or three variables, validated by an independent data set, and analyzed with the diagnostic tools, the solution concentration of paracetamol, the solvent composition, and the solution temperature are predicted with enough accuracy. The solubility of paracetamol in different water−isopropanol compositions is in good agreement with the measured solubility using a gravimetric technique. The solubility in a mixture of solvents shows a maximum at a specific solvent composition. The predicted solubility data in pure isopropanol and pure water are also in reasonable agreement with those reported in the literature. In general, the results indicate that the method can provide a powerful tool for real-time concentration measurement in multicomponent systems.

Acknowledgment

This work was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada and Canada Foundation for Innovation (CFI).

NOTATION

A_i	absorbance of spectra at particular wavenumber of i
a_i, b_i	adjustable parameters
C	concentration (g/1000 g of solvent)
$\bar{C}$	mean value for each component
$\bar{\bar{C}}$	matrix data defined in eq 2
C_{exp}	experimental value for each component
C_{pre}	predicted value for each component
C_p^*	apparent heat capacity of solute (kJ mol^{-1} K^{-1})
$F_{\alpha,n,n}$	F distribution
ΔH_s^*	apparent enthalpy of solution (kJ mol^{-1})
n	number of samples
p	number of parameters
P_{ratio}	ratio of *PRESS* value to the smallest *PRESS* value
R	universal gas constant (J mol^{-1} K^{-1})
R^2	correlation coefficient squared
T	temperature (K or °C)
x	solubility in mole fraction
Greek letters	
α	probability value
β_i	adjustable parameter
$\sigma_{r,comp}$	residual standard deviation for each component

(28) Grant D. J. W.; Mehdizadeh, M.; Chow, A. H. L.; Fairbrother, J. E. Non-Linear van't-Hoff Solubility-Temperature Plots and Their Pharmaceutical Interpretation. *Int. J. Pharm.* **1984**, *18*, 25−33.

Abbreviations

ATR-FTIR	attenuated total reflection-Fourier transform infrared
CI	confidence interval
CL	confidence limit
IR	infrared
MLR	multiple linear regression
PA	paracetamol
PCR	principal component regression
PLS	partial least squares
PRESS	predicted residual error sum of squares

Received for review April 4, 2006.

OP060073O