Online Estimation and Monitoring of Diastereomeric Resolution Using FBRM, ATR-FTIR, and Raman Spectroscopy

Sze-Wing Wong,[†,‡] Christos Georgakis,*[,†] Gregory D. Botsaris,[†] Kostas Saranteas,[§] and Roger Bakale[§]

System Research Institute for Chemical and Biological Processes and Department of Chemical and Biological Engineering, Tufts University, Medford, Massachusetts; and Chemical Process Research and Development, Sepracor Inc., Marlborough, Massachusetts

This paper first demonstrates that the estimation of fractional solid composition of two diastereomers is possible by incorporating Raman spectra, slurry density, and temperature into a partial least squares (PLS) model. Since slurry density measurement is not readily available, the online estimation is obtained through another PLS model that utilizes online Raman spectroscopy, focus beam reflectance measurement (FBRM), attenuated total reflection Fourier transform infrared (ATR-FTIR), and/or additional process information such as temperature and agitation rate to infer slurry density online. It is argued that the model, which infers slurry density from IR spectra, gives more accurate estimation of the fractional composition of two diastereomers. Furthermore, the paper also successfully demonstrates through a third and a more general PLS model the ability of real time monitoring several crystallization process variables by the use of the aforementioned online measurements. Besides monitoring the changing diastereomeric composition and slurry density, one can also monitor the solute concentration of each diastereomer in a complex crystallization system.

1. Introduction

The active pharmaceutical ingredient (API) is often produced either in the form of racemates (i.e., compounds with 50:50 proportion of enantiomers) or as pure enantiomers, which are chiral molecules that are mirror images of each other. Although the chemical and physical properties of both enantiomers (R = right handed and S = left handed) are essentially the same in an achiral environment, they are readily distinguished by biological systems. Therefore, the enantiomers often have different pharmacokinetic properties such as absorption and distribution rate and may also have different toxicological effects.[1] As a result, the Food and Drug Administration (FDA) requires the manufacturer to provide pharmacokinetic properties and toxicological effects of each enantiomer and of the quantitative isomeric composition of the final product.[2] Since the physical properties of both enantiomers are the same, a traditional separation method such as crystallization by seeding is feasible but it becomes very sensitive to the experimental conditions.[3,4] There are generally two ways to obtain one of the enantiomers, either by achiral synthesis of the enantiopure product or by the reaction of enantiomers with another chiral acid or base to produce different diastereomers. Diastereomers are molecules that have two or more chiral centers and are not mirror images of each other (Figure 1). The resulting diastereomers have different physical properties such as solubility and often crystallize into different crystal structures (Figure 2). With these differences, crystallization can effectively separate the desired product with very high purity.[5,6]

In order to comply with strict government regulations, the purity of the final product is of crucial importance. Thus, the capability of online monitoring the optical purity and solute concentration of the diastereomers will help to develop a robust crystallization procedure. Several online analytical measurements such as the Raman spectroscopy, focus beam reflectance measurement (FBRM), and attenuated total reflection Fourier transform infrared (ATR-FTIR) are suitable for such application. Several studies showed that Raman spectroscopy was capable in monitoring a solvent-mediated polymorphic transformation[7–13] online and along with multivariate analysis it is able to predict the percent composition of the polymorphs and/or slurry density.[12–15] In addition, ATR-FTIR along with chemometric tools are proven to be a reliable analytical tool to monitor the changing solute concentration throughout crystallization in real time.[17–22] With the recent advances in applying online analytical measurements in crystallization systems, several authors have combined FBRM and ATR-FTIR to better understand the thermodynamic and kinetic characteristics of different crystallization systems.[23–25]

Recent research has shown that Raman spectroscopy is capable of differentiating and quantifying diastereomers in a crystallization slurry, and the accuracy of the estimation model increases provided the changing process parameters of temperature and slurry density are included in the model.[26,27] However,

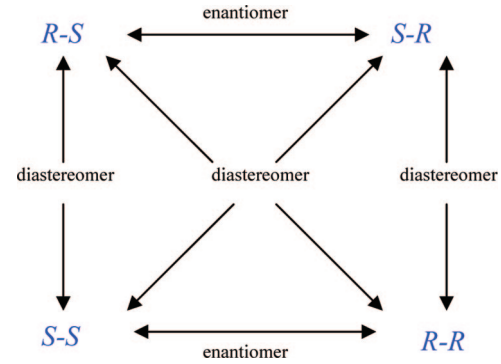

Figure 1. Relationships between enantiomers and diastereomers. The compounds used here consist of two diastereomers each with one different chiral center with one being either right-handed (R) or left-handed (S).

* To whom correspondence should be addressed. Tel.: 617-627-2573. Fax: 617-627-3991. E-mail: Christos.Georgakis@tufts.edu.
[†] Tufts University.
[‡] Currently at Chemical Product R&D, Eli Lilly & Company, Indianapolis, IN 46285.
[§] Sepracor Inc.

10.1021/ie071179q CCC: $40.75
Published on Web 07/01/2008

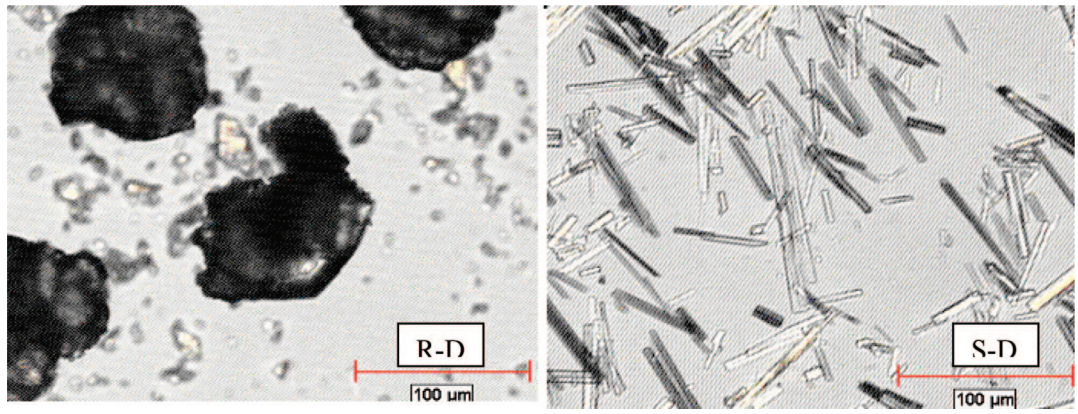

Figure 2. (Left) Polarized-light microscopy of R-D agglomerate. (Right) Polarized-light microscopy of S-D crystals.

slurry density is not easily measured online without a sampling loop. It is essential to find an alternative online measurement from which to infer slurry density. The choice for the alternative online measurement is ATR-FTIR and FBRM. ATR-FTIR is often used to monitor how solute concentration changes in solution phase over time during crystallization. It serves as an indirect measure of the solid phase dynamics. On the other hand, FBRM is essentially a particle counter and it gives a more direct measure of slurry density in measuring the total number of crystals scanned per second. In the present work, a complex crystallization system is examined to determine whether FBRM or ATR-FTIR measurements are sufficient to replace slurry density measurement along with Raman spectra in providing an accurate estimation of the solid composition of one of the two diastereomers. Furthermore, a methodology is developed to systemically estimate other process variables from the vast amount of data obtained from all three online measurements. Under this methodology, four partial least squares (PLS) estimation models were developed and used to convert the spectroscopic data into (1) slurry density, (2) diastereomeric composition of the crystals, (3) solute concentration of compound A (Sepracor Compound), and (4) the percent composition of diastereomer in solution. The converted information can then be used to model the crystallization dynamics to optimize and control the process.

2. Experimental Section

2.1. Materials. HPLC grade solvents were used as received from commercial suppliers without further purification. The starting materials (racemic free base of compound A and a chiral acid, denoted here as D) that met the specifications defined by Sepracor Inc. were used as received from qualified suppliers without further purification.

2.2. Raman Spectroscopy. A RamanRxn1 analyzer (Kaiser Optical System, Inc.) coupled with an immersion fiber optic probe was used for the in situ measurements. Raman spectra were recorded using near-IR excitation radiation at 785 nm and the spectroscopy incorporates the TE-cooled CCD detector technology. The Raman detector was set to have an exposure time of 5 s with five accumulation spectra. Each recorded Raman spectrum was the sum of five accumulation spectra acquired in a 1-min interval.

2.3. ATR-FTIR. The ReactIRiC10 coupled with a flexible FiberConduit and a 6 mm DiComp probe (Mettler Toledo AutoChem, Inc.) was used for the in situ measurements. The spectrometer was purged with nitrogen gas 24 h before the experiments and during measurement to reduce water and carbon dioxide absorption. Air was used as the background measurement of the IR spectrometer, and liquid nitrogen was used to cool the spectrometer. All the IR spectra were collected at 258 scans per spectrum and recorded every minute.

2.4. FBRM. The Lasentec FBRM D600L was used to monitor the particle chord length distribution and particle counts (Mettler Toledo AutoChem, Inc.) with the acquisition software version 6.0 build 14. A measurement period of 20 s was chosen for all crystallization experiments, and the data point was recorded every minute.

2.5. Experimental Setup. A 500 mL jacketed curved-bottom reactor was used for all the crystallization experiments. The reactor has an inner diameter of 100 mm and a bottom drain valve. A four-blade glass stirrer was connected to a digital overhead agitator to ensure adequate mixing. Both the temperature and agitation rate were automated and controlled by the Advantage 4100 system (Biotage AB) and the data was collected at 1-min intervals.

All the online measurements such as the Raman, FBRM, ATR-FTIR, and the thermocouple were inserted top-down into the reactor at a slight angle through a Teflon reactor head (Figure 3a). The probe windows were positioned slightly above the glass stirrer to ensure a representative sample of the slurry and to minimize coating of the probe windows (Figure 3b). A condenser is also attached to the Teflon reactor head to reduce solvent evaporation during the experiments. The equipment setup remained the same throughout all the experiments.

2.6. Calibration Experiments. Ten seeded crystallization experiments with different operating conditions were conducted to build and test the four estimation models. The seeded crystallization experiment procedure and analytical method were detailed in previous publications and will not be detailed further.[26,27] These estimation models were developed utilizing the analytical result of the samples, which were taken sequentially in the experiments at different temperatures. In order to ensure the estimation models covered a wide enough model space which will be representative of future crystallization experiments, the design of experiments was based on a fractional factorial design with varying process variables such as the amount of seeding, agitation rate, and cooling rate (Table 1). It should be noted that seven of the 10 experiments were used for model development, while three of the remaining 10 experiments (experiment no. 4, 6, and 9) were set aside and used to test the model performances.

2.7. Data Pretreatment. The Raman and IR spectral data was analyzed using either the PLS_Toolbox 3.5 by Eigenvector Research, Inc. (Manson, WA) or the Unscrambler Chemometrics Software from Camo Inc. (Trondheim, Norway). Data pretreatment using first derivative was performed on both the IR and

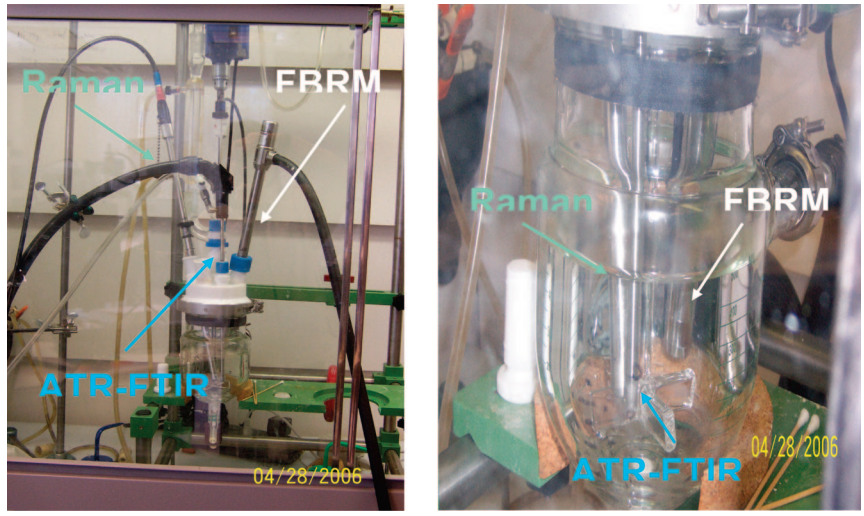

Figure 3. (a, left) Photograph of the reactor setup with all the online measurements inserted top-down into the reactor. (b, right) A close-up view of the probe windows positioned slightly above the glass stirrer.

Table 1. Experimental Conditions

expt no.	seed (wt %)	cooling rate (°C/min)	agitation (rpm)
1	2	0.1	550
2	4	0.1	350
3	2	1	350
4	4	1	550
5	3	0.5	450
6	4	0.5	300
7	2	nonlinear cooling profile 1	350
8	2	nonlinear cooling profile 2	250
9	6	nonlinear cooling profile 2	350
10	6	nonlinear cooling profile 1	250

Raman spectrum to correct any possible baseline drift. It can be observed that there is slight baseline drifting from the raw spectra (Figures 4 and 5) and it is removed by taking the first derivative of the spectra. The first derivative of all the spectra was computed with the Savitzky−Golay method using second-degree polynomial fit and 11-points window.[28] It was found that the 11-points window was sufficient to achieve a high level of smoothing without attenuating the extrema in the data. Higher order derivative did not show any significant improvement in model accuracy and hence the first derivative of the spectral measurement was chosen as the only data pretreatment step.

Three measurement data matrices are used to develop the four PLS estimation models, which are composed of different combinations of online measurements such as the entire Raman spectrum (spectrum range 75−3300 cm^{-1}), FBRM statistical data (no. of total particles scanned per second), FTIR spectrum (940−1880 cm^{-1}), temperature, agitation rate (rpm), and/or slurry density as shown in eqs 1 –3. The solid composition of the diastereomers, solute concentration of compound A, percent composition of the diastereomers in solution phase, and slurry density were the *y*-variables of the four different PLS models.

$$D_1 = [\text{Temp (T)} \quad \text{Slurry Density (D)} \quad \text{Raman (Ra)}] \tag{1}$$

$$D_2 = [\text{Temp (T)} \quad \text{FTIR (IR)}] \tag{2}$$

$$D_3 = [\text{Temp (T)} \quad \text{RPM} \quad \text{FBRM (P)} \quad \text{Raman (Ra)}] \tag{3}$$

In order to compare the contribution of each factor equally, temperature (T), slurry density (D), FBRM statistical data (P), agitation rate (RPM), FTIR spectra (IR), and Raman spectra (Ra) were first scaled to have zero mean and variance of one as shown in eqs 4–10.

$$\hat{T}_i = \frac{T_i - T_{\text{avg}}}{\sigma_T} \tag{4}$$

$$\hat{D}_i = \frac{D_i - D_{\text{avg}}}{\sigma_D} \tag{5}$$

$$\hat{P}_i = \frac{P_i - P_{\text{avg}}}{\sigma_P} \tag{6}$$

$$\widehat{\text{RPM}}_i = \frac{\text{RPM}_i - \text{RPM}_{\text{avg}}}{\sigma_{\text{RPM}}} \tag{7}$$

$$F_i = \int f_i(w)\, dw \tag{8}$$

$$F_{i,\text{avg}} = \frac{\sum_{i=1}^{n} F_i}{n} \tag{9}$$

$$\tilde{S}_i(w) = \frac{f_i(w) - F_{i,\text{avg}}}{\sigma_{F_i}} \tag{10}$$

Here $\hat{T}$, $\hat{D}$, $\hat{P}$, $\hat{\text{RPM}}$, and $\hat{S}_i(w)$ represent the scaled values, obtained by subtracting the average values (T_{avg}, D_{avg}, P_{avg}, RPM_{avg}, and $F_{i,\text{avg}}$) and dividing by the corresponding standard deviation (σ). However, the spectral data (both FTIR and Raman), distributed over a frequency spectrum, was scaled slightly different, using $F_{i,\text{avg}}$ (eq 9), the average of the under curve areas of all spectra data (eq 8), and σ_{F_i} the corresponding standard deviation. It should be noted that $f_i(w)$ is the spectrum function with respect to w, denoting the wavenumber that corresponds to each frequency.

3. Estimation Model

3.1. Partial Least-Squares Regression Model (PLS). The four estimation models were calculated by regressing the measurement data matrix (eqs 1–3) of the training group with different *y* variables using PLS. The training group (85 samples from the seven crystallization experiments) was divided into 10 subgroups with 8/9 data points per subgroup for cross-validation. The relative percent error (% E), the root mean square of % E (E_1), and the root mean square error (E_2) would be used to compare the different PLS models (eqs 11–13) with the prediction results from the testing groups (34 samples from three

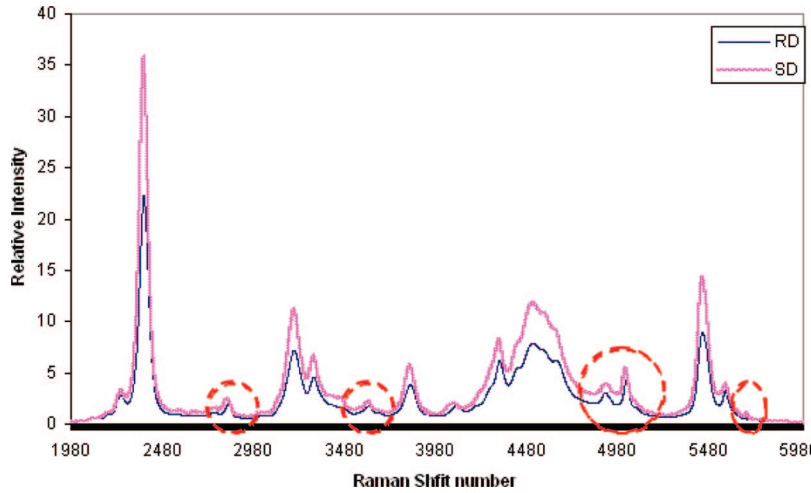

Figure 4. Raman spectra of pure R-D and S-D at the same temperature and slurry density to illustrate the baseline offset. The circles indicate different regions of slight peak shifts between the diastereomers.

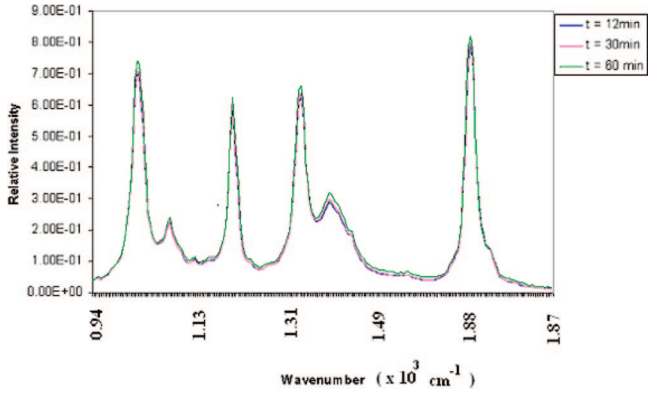

Figure 5. IR spectra collected at different time to illustrate the baseline drift in crystallization experiment no. 3.

experiments). All the PLS estimation models are developed using the Unscrambler Chemometrics Software.[33]

$$\%E = \left| \frac{(y_{i,\text{estimate}} - y_{i,\text{experiment}})}{y_{i,\text{experiment}}} \right| \times 100\% \quad (11)$$

$$E_1 = \sqrt{\frac{\sum_{i=1}^{n} (\%E)^2}{n}} \quad (12)$$

$$E_2 = \sqrt{\frac{\sum_{i=1}^{n} (y_{i,\text{estimate}} - y_{i,\text{experiment}})^2}{n}} \quad (13)$$

where n is the number of the total data points.

3.2. Estimation Model of Solid Composition of Diastereomer (PLS Model I). It was shown previously that model accuracy is greatly improved by incorporating process variables such as temperature and slurry density into the model along with Raman spectral data (eq 3).[26,27] However, real time slurry density measurement cannot be easily obtained without a sampling loop connected to a densitometer. The disadvantages of using a sampling loop are potential clogging and it requires precise temperature control. If the temperature is not well controlled within the loop, nucleation and dissolution may occur inside the sampling loop and hence affect the slurry density reading. The first PLS model addresses the issue of whether an alternative online measurement may be used to replace densitometer or sampling. Three estimation models with the same measurement data matrix (eq 1) were developed in which slurry density measurement was either obtained through sampling or calculated with other online measurements. The objective was to compare all three models and check whether the alternative online measurements have at least the same or better prediction accuracy of the solid diastereomeric composition.

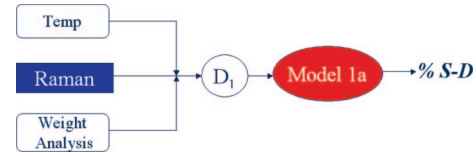

Figure 6. Model structure of PLS model Ia.

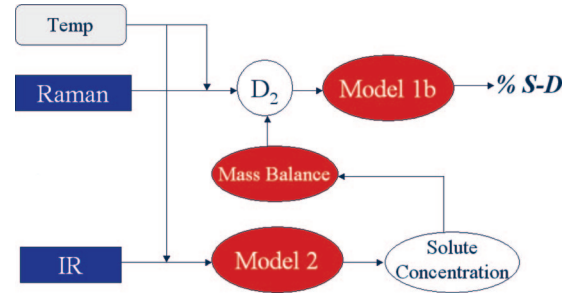

Figure 7. Model structure of PLS model Ib.

The first estimation model (model Ia) acts as the control group in which the values of the slurry density are obtained from weight analysis of the samples (Figure 6). The slurry density was calculated by weighing the dry solid sample and dividing the weight by the sampling volume. The second estimation model (model Ib) estimated slurry density with the alternative measurement of ATR-FTIR. The IR spectral data along with the temperature measurement (eq 2) were used to develop a second estimation model to predict solute concentration of compound A. Since the initial solute concentration of compound A was known, a simple mass balance is used to calculate the slurry density measurement at any time (Figure 7). Finally, the third estimation model (model Ic) used a different combination of several online measurements (eq 3) to directly estimate slurry density through a third PLS estimation model (Figure 8). It should be noted that the three estimation models (models Ia, Ib, and Ic) only differ in how slurry density is obtained. The models have the same measurement data matrix (eq 1) and the

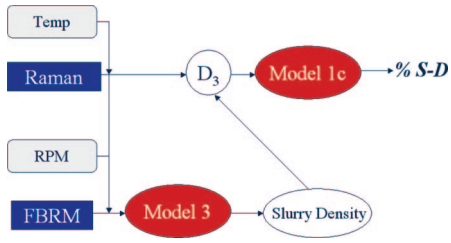

Figure 8. Model structure of PLS model Ic.

data matrix was used to regress with the percent composition of the solid diastereomer.

3.3. Estimation of Crystallization Process Variables. As discussed previously, besides monitoring the relative percent composition of the diastereomer and the slurry density, one should be able to estimate other important crystallization process variables such that the kinetics of the complex system can be modeled. Three different PLS models are developed in which each model involved different combination of several online measurements to estimate the slurry density, the solute concentration of compound A, and the relative composition of diastereomer in solution phase.

3.3.1. Estimation Model of Solute Concentration of Compound A (PLS Model II). A second estimation model is constructed with the IR spectral data and temperature (eq 2) to predict the solute concentration of compound A. Since the IR spectra of compound A has multiple absorbances at different wavelengths as a function of concentration and solution temperature, PLS is used to convert a range of wavelength ($940-1880\ cm^{-1}$) along with temperature to convert absorbance data into solute concentration.[17,21]

3.3.2. Estimation Model of Slurry Density (PLS Model III). The number of particles scanned by the FBRM per second is directly proportional to slurry density, which is defined as the total mass of crystals per unit volume. As a result, a simple linear model ($y = mx + b$) should be sufficient to relate the FBRM measurement with slurry density. However, the FBRM measurement is extremely sensitive to the agitation rate. As the agitation rate increases, the number of air bubbles generated along with the different mixing environment affects the FBRM measurement such that a simple linear model will have larger uncertainty.

To address the above problem, a third estimation model is constructed with the combination of measurements such as temperature, agitation rate (RPM), FBRM statistical data (no. of particles scanned per second), and Raman spectra (eq 3) to directly predict slurry density. The reason for including Raman spectra in the measurement data matrix is because the number of crystals inside the Raman detection zone will influence the intensity of the returned Raman signal. Furthermore, the Raman spectrum is affected by the relative amount of solvent and solids in its path. Consequently, slurry density should impact the returned Raman signal intensity. A recent study showed that Raman spectra along with multivariate analysis were capable of predicting both the slurry density and percent composition of an anhydrous/monohydrate system.[12] However, the authors did not investigate whether temperature has any effects, as all the experiments were run isothermally. In the present work, the estimation model of slurry density takes into account both the temperature and agitator rate by incorporating the process variables into the model structure.

3.3.3. Estimation Model of Percent Composition of Diastereomer in Solution Phase (PLS Model IV). In order to monitor and understand the complex nature of a diastereomer resolution, information regarding the solute concentration of each diastereomer will be needed to calculate the driving force (supersaturation) of nucleation and growth. In general, when both diastereomers (S-D and R-D) are dissolved in solution, their chemical structures are only different in one of the chiral centers. Since both IR and Raman cannot differentiate chiral molecules, the solute concentration of each diastereomer will have to be calculated through mass balance. However, in this particular system, one of the diastereomers is a solvate such that IR spectroscopy is capable in differentiating the diastereomers in solution phase (Figure 9). As a result, a fourth estimation model is developed to predict the percent composition of diastereomer in solution phase with the combined measurements of the IR spectral data and temperature (eq 2).

4. Result and Discussion

In the previous section, we detailed the four families of PLS models (models I, II, III, and IV), which were developed to convert three online measurements into four important variables of the crystallization process. These models can be used to monitor the complex crystallization system. In the first model family (I), three estimation models (models Ia, Ib, and Ic) were constructed with the same modeling objective (eq 1) to predict the solid diastereomeric composition. In the other three model families (II to IV), the modeling objective was to explore how to systematically convert the online measurements into useful crystallization process variables that enable real time process monitoring.

4.1. PLS Model Development. The selection of the optimal number of latent variables (LV) was based on the cross-validation results (Figure 10). Detailed information of all the models is summarized in Tables 2 and 3. Each of the four PLS models was used to predict the different prediction variables, y, with the same testing group (34 samples) to check the model performance.

4.1.1. Alternative Online Measurements As Slurry Density Measurement (Model Ia vs Ib vs Ic). The first objective of the paper is to test the hypothesis whether there exists alternative online measurements that can be used as a replacement of slurry density measurement in estimating the diastereomer solid composition. In order to investigate whether the above hypothesis is true, three PLS estimation models were developed in which slurry density was either obtained through sampling or inferred from the alternative measurements. The first step in comparing the performance of the three models is to plot the measured y variable (% S-D) against the predicted y variable (Figure 11). Although the data points from model Ib are slightly different from the data points in model Ia and model Ic, it is observed that the three models in general have very good agreement and the data points from each model are almost overlapping each other. As a result, it is expected that the three models should have similar model performance. In addition, the R^2 value of 0.9 indicates that there is good correlation between the measurement data matrix and the relative percent composition of the solid diastereomer.

The three estimation models were applied into the test group of data, and the accuracy of each model is summarized in Table 4. It is not a surprise that the three models have similar model performance. The model in which slurry density was calculated with the IR spectral data (model Ib) has a slightly better prediction accuracy. The possible cause for model Ib having better prediction accuracy is due to the smaller experimental error associated with the HPLC analysis of the mother liquor. It should be noted that as the sample was taken from the

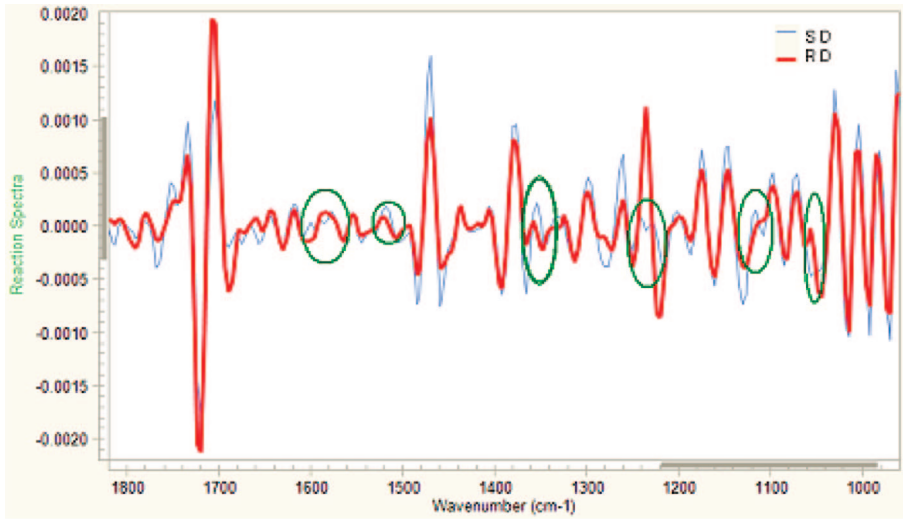

Figure 9. First derivative of the IR spectrum of both pure diastereomers dissolved in solution. The circled regions indicate the differences between the two diastereomers.

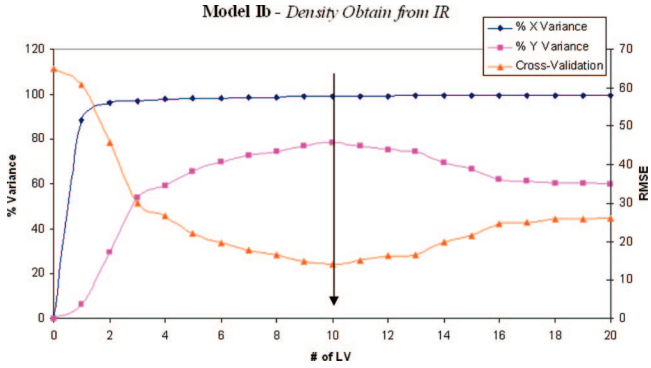

Figure 10. An illustration of how the optimal number of latent variables is selected.

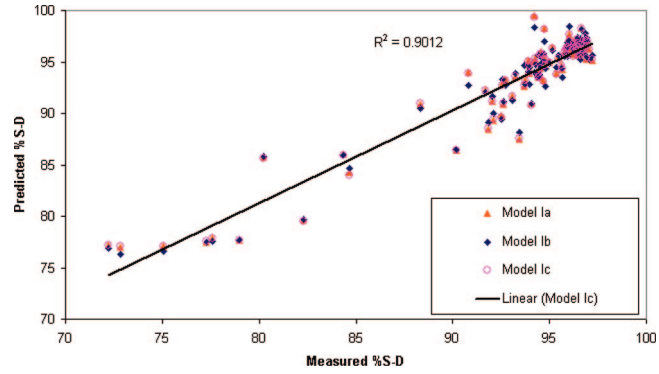

Figure 11. Comparison of models Ia, Ib, and Ic with the predicted *y* variable vs the measured *y* variable.

Table 2. Detailed Information of Model I (% solid composition of diastereomer)

model	no. of LV	% *X* variance	% *Y* variance
Ia (density from weight analysis)	9	98.99	75.2
Ib (density obtained from IR)	10	99.05	78.35
Ic (density obtained from FBRM and Raman)	9	98.97	76.55

Table 3. Detailed Information of Modes l II−IV

model	no. of LV	% *X* variance	% *Y* variance	prediction variable
II	5	98	86	solute concentration of A
III	6	98	98	slurry density
IV	5	99	88	% diastereomer in solution

experiment, it was quickly filtered to separate the solid and solution phase samples. While the solution was directly submitted for HPLC analysis which has an experimental error of around 2% assuming the solution temperature remained constant during the filtration step, a limited amount of the solid phase sample has to go through drying which may incur additional experimental error in the calculation of slurry density. Nevertheless, the estimation models successfully demonstrated the ability to replacing slurry density measurement with the alternative measurements such as FBRM and ATR-FTIR. Since both online measurements are noninvasive and have fast data acquisition times, the need of a sampling loop is eliminated and the system now has the capability to monitor the crystallization experiments in real time.

Table 4. Prediction Error of Model I with Testing Group

model	max % E	E_1	E_2
Ia (density from weight analysis)	11.2	5.34	4.84
Ib (density obtained from IR)	*10.3*	*5.26*	*4.79*
Ic (density obtained from FBRM and Raman)	11.7	5.29	4.80

4.1.2. PLS Model Performance (Models II−IV). In addition to developing the estimation model of the solid composition of the diastereomer, three other models with different combinations of online measurements were constructed to predict three key process parameters needed to monitor and model a complex crystallization system. This is particularly important in the chemical process development area; as more online analytical measurements become available, there is a need to systematically convert the large amount of raw data into useful information. Furthermore, the estimated process variables obtained from the online measurements provide important insights of the processes.

It is observed through the plot of measured *y* variable vs predicted *y* variable (Figures 12–14) that there is no apparent bias in the estimation models and the regression models have satisfactory correlation between the data matrix and the *y* variables. Finally, all three models were applied on the testing group data and the prediction accuracy is summarized in Table 5. In the two model families in which the IR spectral data were applied along with the temperature measurements (models II and IV), the prediction error (E_2) of each model is slighter higher. The slightly higher prediction error associated with each model might be due to the increase of IR measurement noise

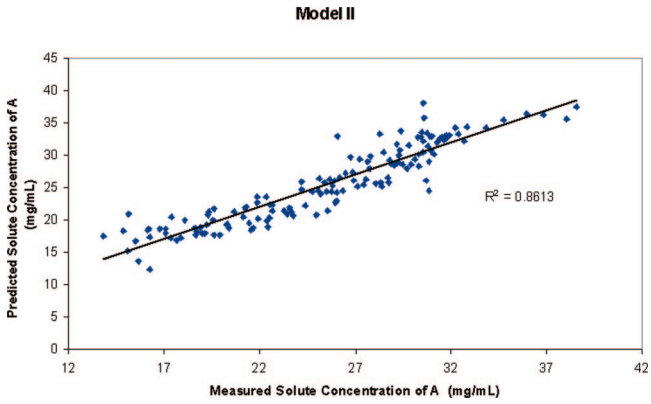

Figure 12. Model II: predicted vs measured solute concentration of compound A.

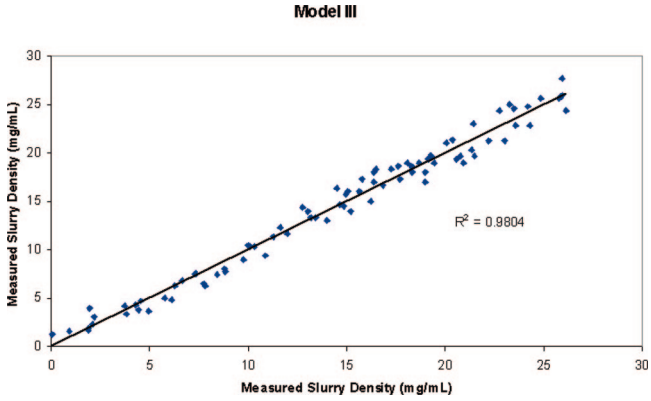

Figure 13. Model III: predicted vs measured slurry density.

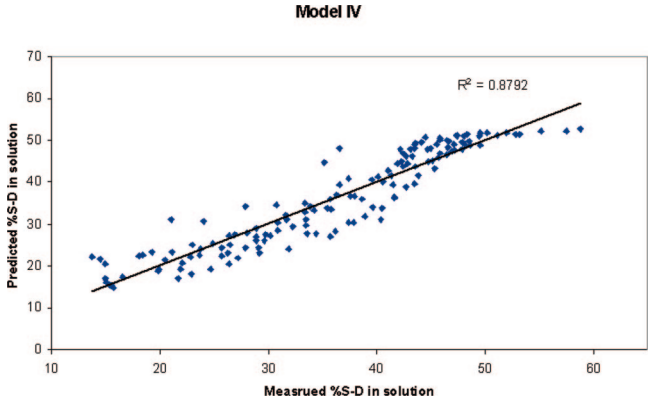

Figure 14. Model IV: predicted vs measured relative percent composition of diastereomer in solution phase.

Table 5. Prediction Error of Models II−IV with Testing Group

model	prediction variable	E_2
II	solute concentration of A	4.76
III	slurry density	2.97
IV	% diastereomer in solution	4.40

during nucleation which introduced interference from the solid phase. Although the ATR-FTIR instrument is designed to have a penetration depth of a few micrometers and should not detect the solid phase, it was observed during the crystallization experiments that there were fouling at the IR probe window from the newly formed crystals. As a result, the IR probe was taken out of the reactor periodically and flushed with solvent to clean the probe window. Furthermore, it should be noted that model III showed excellent correlation between slurry density,

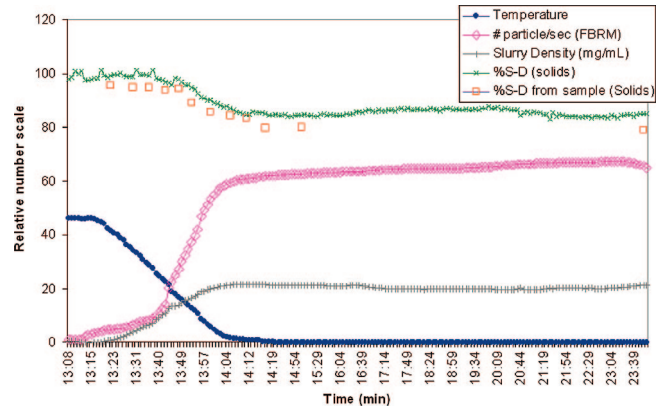

Figure 15. Crystallization example 1. Information regarding the solid phase such as the slurry density, FBRM measurement, and the relative percent composition of diastereomer in solid phase is monitored. Note that the open square symbol (□) represents the measured slurry density from the samples.

FBRM measurement, and Raman spectral data. The average prediction error of model III is also the lowest among all the estimation models.

4.2. Online Monitoring of Crystallizations. With the use of three online analytical measurements such as the Raman, ATR-FTIR, and FBRM, acquiring data throughout a crystallization experiment, one should be able to monitor several important process variables and thus gain process understanding. To illustrate how process insight can be obtained by using the developed models to monitor the crystallization process, four estimation models (models I−IV) are applied to two crystallization examples with different operating conditions. The first crystallization example employed a fast linear cooling rate (1 °C/min) in which the nucleation of the undesired diastereomer (R-D) was observed. In the second crystallization example, a nonlinear cooling rate was employed in which the initial cooling rate was slow to suppress nucleation and maximize growth of the desired diastereomer (S-D). The information obtained from the models through real-time measurements is essential to improving the operation of the crystallization process and ensuing the production of products with consistent quality.

4.3. Crystallization Example 1. In the first crystallization experiment, 4 wt % of the S-D diastereomer seeds was added into the solution with the agitation rate set at 550 rpm. The solution was then cooled at 1 °C/min from 46 to 0 °C and the temperature was held constant at 0 °C for a prolonged period of time. It may be observed from the FBRM data (no. of particles/s) and from the slurry density estimation (model III) when nucleation occurred, but it was unclear whether the desired S-D diastereomer or the undesired one (R-D) began to nucleate (Figure 15). However, when the information of the relative composition of the diastereomers in both the solid and the solution phase are examined, it becomes apparent that at the onset of nucleation, only the S-D diastereomer crystallized out of the solution. The nucleation point of the S-D diastereomer can be clearly identified at the moment when the FBRM measurement recorded a jump in particle count while the solid percent composition of the diastereomer remained unchanged (Figure 15). Furthermore, the solution phase information (models II and IV) obtained from the IR data also supports the above finding. As the R-D solute concentration dropped, the solid percent composition of diastereomer dropped, which signals the nucleation point of the undesired diastereomer (Figure 16). It should be noted in the circled region in Figure 16 that the increase of solute concentration was due to fouling of the IR

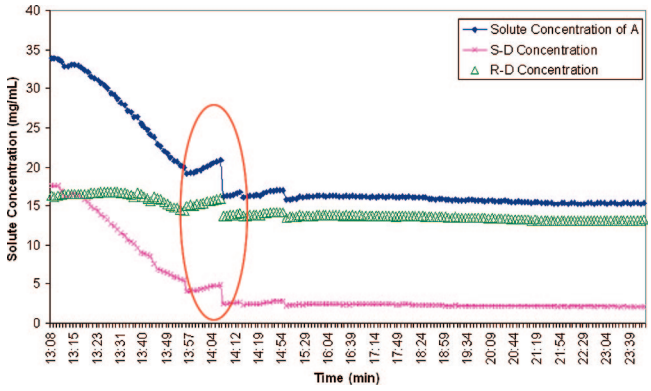

Figure 16. Crystallization example 1. Information regarding the solution phase such as the overall solute concentration of compound A and the solute concentration of each diastereomer is monitored.

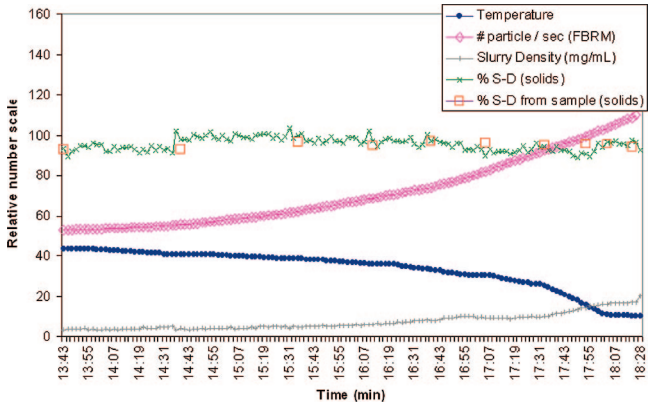

Figure 17. Crystallization example 2. Information regarding the solid phase such as the slurry density, FBRM measurement, and the relative percent composition of diastereomer in solid phase is monitored. Note that the open square symbol (□) represents the measured slurry density from the samples.

probe window by the newly formed crystals. As the IR probe was taken out of the crystallizer and flushed with solvent, the predicted solute concentration dropped to normal level.

4.4. Crystallization Example 2. In the second crystallization experiment, 6 wt % of the S-D diastereomer seeds were added into the solution with the agitation rate set at 350 rpm The solution was cooled with a nonlinear cooling rate designed to suppress nucleation at high supersaturation region and maximize growth of the desire diastereomer. It may be observed that nucleation was minimized in the beginning of the experiment (FBRM data) and the estimated slurry density remained constant until the cooling rate increased more rapidly at the end of the experiment (Figures 17 and 18). Furthermore, the nucleation of the undesired diastereomer was successfully suppressed with the use of a slower cooling rate. Both the solid composition of diastereomers (% S-D) and the R-D concentration suggested that the R-D diastereomer remained in solution throughout the experiment.

5. Conclusion

Three PLS models (models Ia, Ib, and Ic) were constructed with the same model structure to demonstrate that the alternative measurements of ATR-FTIR and FBRM were indeed able to estimate slurry density. The direct measurement of slurry density through the use of sampling loop was avoided. It is argued that the estimation model, which infers slurry density from IR spectral data, provides the best prediction of the solid composition of the diastereomer. Furthermore, a more general PLS model was developed to systematically convert online Raman, ATR-FTIR, and FBRM measurements into key process variables needed to model crystallization kinetics. The key process variables estimated from the four PLS models (models I−IV) are the solid composition of diastereomer, slurry density, the overall solute concentration of compound A, and the relative composition of diastereomer in solution phase. The four estimation models were applied to two crystallization experiments in which the ability to use the models to monitoring a complex crystallization system was successfully demonstrated. The paper showed that valuable insight of the process may be obtained through real time monitoring. Most importantly, the methodology may be used to obtain both solid and solution phase information needed to model the complex nature of a crystallization process.

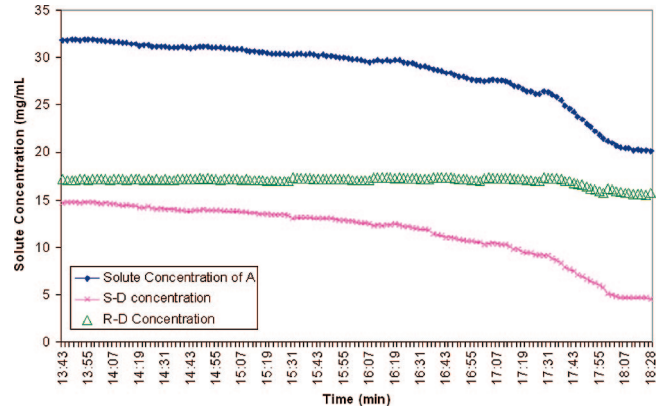

Figure 18. Crystallization example 2. Information regarding the solution phase such as the overall solute concentration of compound A and the solute concentration of each diastereomer is monitored.

Acknowledgment

The financial support from Sepracor Inc. in the form of an internship to Sze-Wing Wong and the availability of their experimental facilities is greatly appreciated.

Note Added after ASAP Publication: The author affiliations in the version of this paper that was published on the Web July 1, 2008 were incorrect. The correct version of this paper was reposted to the Web July 30, 2008.

Literature Cited

(1) Reddy, I. K.; Mehvar, R. *Chirality in Dry Design and Development*; Marcel Dekker Inc.: New York, 2004.

(2) FDA's Policy Statement for the Development of New Stereomisomeric Drugs; http://www.fda.gov/cder/guidance/stereo.tm.

(3) Qian, R. Y.; Botsaris, G. B. A New Mechanism for Nuclei Formation in Suspension Crystallizers: The Role of Interparticle Forces. *Chem. Eng. Sci.* **1997**, *52*, 3429–3440.

(4) Elsner, M. P.; Menendez, D. F.; Muslera, E. A.; Seidel-Morgenstern, A. Experimental Study and Simplified Mathematical Description of Preferential Crystallization. *Chirality* **2005**, *17*, S183–S195.

(5) Jacques, J.; Collet, A.; Wilen, S. H. *Enantiomers, Racemates, and Resolution*; Wiley: New York, 1981.

(6) Wilen, S. H.; Collet, A.; Jacques, J. Strategies in Optical Resolutions. *Tetrahedron* **1977**, *33*, 2725–2736.

(7) Wang, F.; Wachter, J. A.; Antosz, F. J.; Berglund, K. A. An Investigation of Solvent Mediated Polymorphic Transformation of Progesterone Using in Situ Raman Spectroscopy. *Org. Process Res. Dev.* **2000**, *4*, 391–395.

(8) O'Sullivan, B.; Barrett, P.; Hsiao, G.; Carr, A.; Glennon, B. In Situ Monitoring of Polymorphic Transitions. *Org. Process Res. Dev.* **2003**, *7*, 977–982.

(9) Ono, T.; Horst, H. T.; Jansens, P. J. Quantitative Measurement of the Polymorphic Transformation of L-Glutamic Acid Using In-Situ Raman Spectroscopy. *Cryst. Growth Des.* **2004**, *4*, 465–469.

(10) Hu, Y.; Liang, J. K.; Myerson, A. S.; Taylor, L. S. Crystallization Monitoring by Raman Spectroscopy: Simultaneous Measurement of Desupersaturation Profile and Polymorphic Form in Flufenamic Acid System. *Ind. Eng. Chem. Res.* **2005**, *44*, 1233–1240.

(11) Schöll, J.; Bonalumi, D.; Vicum, L.; Mazzotti, M.; Muller, M. In Situ Monitoring and Modeling of the Solvent-Mediated Polymorphic Transformation of L-Glutamic Acid. *Cryst. Growth Des.* **2006**, *6*, 881–889.

(12) Caillet, A.; Fevotte, G. In-line Monitoring of Partial and Overall Solid Concentration during Solvent-Mediated Phase Transition Using Raman Spectroscopy. *Int. J. Pharm.* **2006**, *307*, 201–208.

(13) Starbuck, C.; et al. Process Optimization of a Complex Pharmaceutical Polymorphic System via in Situ Raman Spectroscopy. *Cryst. Growth Des.* **2002**, *2*, 515–522.

(14) Falcon, J. A.; Berglund, K. A. In Situ Monitoring of Antisolvent Addition Crystallization with Principle Components Analysis of Raman Spectra. *Cryst. Growth Des.* **2004**, *4*, 457–463.

(15) Rades, T.; Pratiwi, D.; Fawcett, J. P.; Gordon, K. C. Quantitative Analysis of Polymorphic Mixtures of Ranitidine Hydrochloride by Raman Spectroscopy and Principal Components Analysis. *Eur. J. Pharm. Biopharm.* **2002**, *54*, 337–341.

(16) Zhou, G.; Wang, J.; Ge, Z.; Sun, Y. Ensuring Robust Polymorph Isolation Using In-Situ Raman Spectroscopy. *Am. Pharm. Rev.* **2002**, 2002.

(17) 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.

(18) Lewiner, F.; Klein, J. P.; Puel, F.; Fevotte, G. On-line ATR-FTIR measurement of supersaturation during solution crystallization process. Calibration and application of three solute/solvent systems. *Chem. Eng. Sci.* **2001**, *56*, 2069–2084.

(19) Fevott, G. New Perspectives for the on-line monitoring of pharmaceutical crystallization processes using in situ infrared spectroscopy. *Int. J. Pharm.* **2002**, *241*, 263–278.

(20) Yu, Z. Q.; Chow, P. S.; Tan, R. Application of Attenuated Total Reflectance - Fourier Transform Infrared (ATR-FTIR) Technique in the Monitoring and Control of Anti-solvent Crystallization. *Ind. Eng. Chem. Res.* **2006**, *45*, 438–444.

(21) Liotta, V.; Sabesan, V. Monitoring and Feedback Control of Supersaturation Using ATR-FTIR to Produce an Active Pharmaceutical Ingredient of a Desired Crystal Size. *Org. Process Res. Dev.* **2004**, *8*, 488–494.

(22) Feng, L.; Berglund, K. A. ATR-FTIR for Determining Optimal Cooling Curves for Batch Crystallization of Succinic Acid. *Cryst. Growth Des.* **2002**, *2*, 449–452.

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

(24) Barrett, P.; Glennon, B. Characterizing the Metastable Zone Width and Solubility Curve Using Lasentec FBRM and PVM. *Trans. IChemE* **2002**, *80*, 799–805.

(25) Togkalidou, T.; Tung, H. H.; Sun, Y.; Andrews, A. T.; Braatz, R. D. Parameter Estimation and Optimization of a Loosely Bound Aggregating Pharmaceutical Crystallization Using in Situ Infrared and Laser Backscattering Measurements. *Ind. Eng. Chem. Res.* **2004**, *43*, 6168–6181.

(26) Wong, S. W.; Georgakis, C.; Botsaris, G. D.; Bakale, R.; Saranteas, K. Factor Affecting On-Line Estimation of Diastereomer Composition Using Raman Spectroscopy. *J. Control Eng. Practice* **2007**, *15*, 1257–1267.

(27) Wong, S. W.; Georgakis, C.; Botsaris, G. D.; Bakale, R.; Saranteas, K. On-Line Estimation of Diastereomer Using Raman: Differentiation in High and Low Slurry Density PLS Models. Cryst. Growth Des., in press.

(28) Madden, H. H. Comments on the Savitzky-Golay Convolution Method for Least-Squares Fit Smoothing and Differentiation of Digital Data. *Anal. Chem.* **1978**, *50*, 1383–1386.

(29) Beebe, K. R.; Pell, R. J.; Seasholtz, M. B. In *Chemometrics: A Practical Guide*; John Wiley & Sons, Inc.: New York, 1998.

(30) Esbensen, K. H. In *Multivariate Data Analysis*, 5th ed.; CAMO Process AS: Trondheim, Norway, 2002.

(31) Wise, B. M.; Gallagher, N. B.; Bro, R.; Shaver, J. M.; Windig, W.; Koch, R. S. *PLS_Toolbox Version 3.5 for user with MATLAB*; Eigenvector Research Inc.: Manson, MA, 2005.

(32) Caillet, A.; Puel, F.; Fevotte, G. In-line Monitoring of Partial and Overall Solid Concentration during Solvent-mediated Phase Transition using Raman Spectroscopy. *Int. J. Pharm.* **2006**, *307*, 201–208.

(33) The Unscrambler 9.0; Camo Inc.: Trondheim, Norway

Received for review August 30, 2007
Revised manuscript received April 1, 2008
Accepted April 30, 2008

IE071179Q