

Comprehensive kinetic model for the dissolution, reaction, and crystallization processes involved in the synthesis of aspirin

David E. Joiner^a, Julien Billeter^a, Mary Ellen P. McNally^b, Ron M. Hoffman^b and Paul J. Gemperline^{a*}

Kinetic modeling of batch reactions monitored by *in situ* spectroscopy has been shown to be a helpful method for developing a complete understanding of reaction systems. Much work has been carried out to demonstrate the ability to model dissolution, reaction, and crystallization processes separately; however, little has been performed in terms of combining all of these into one comprehensive kinetic model. This paper demonstrates the integration of models of dissolution, temperature-dependent solubility, and unseeded crystallization driven by cooling into a comprehensive kinetic model describing the evolution of a slurry reaction monitored by *in situ* attenuated total reflectance ultraviolet–visible spectroscopy. The model estimates changes in the volume of the dissolved fraction of the slurry by use of the partial molar volume of the dissolved species that change during the course of reagent addition, dissolution, reaction, and crystallization. The comprehensive model accurately estimates concentration profiles of dissolved and undissolved components of the slurry and, thereby, the degree of undersaturation and supersaturation necessary for estimation of the rates of dissolution and crystallization. Results were validated across two subsequent batches via offline high-performance liquid chromatography measurements. Copyright © 2014 John Wiley & Sons, Ltd.

Keywords: batch process monitoring; batch process modeling; slurries; kinetic modeling; crystallization and dissolution

1. INTRODUCTION

The development of kinetic models for use in modeling and monitoring of batch reaction processes has been an active area of research in the last decade [1–4]. Modeling of reaction systems is helpful for acquiring a complete understanding of a production process and has demonstrated its usefulness for monitoring large-scale batch reactions [5,6]. Monitoring of reactions in this manner has both safety and financial benefits. For instance, in the event that a fault is detected because of the inability to fit the data within acceptable limits, corrections may be made in order to save the batch and prevent wasting materials and time.

Much work has been carried out in order to model reactions and reaction processes. Modeling of reactions, dissolutions, and crystallization processes has been a frequent area of study [7–10], and each has become a well-understood process. However, the number of real applications has been rather limited because of the little work that has been performed so far to combine these methods into one cohesive model. Although models for each of these processes separately provide valuable information, a combined model that incorporates all these processes simultaneously would be more desirable for the complete understanding of a slurry-based reaction system because it could be readily applied to the corresponding industrial manufacturing process.

Recently, we demonstrated the ability to apply a kinetic model to a system that includes both dissolution and crystallization processes [11]. This work extends this idea by modeling simultaneously the dissolution, reaction, and crystallization processes and by incorporating the changes in volume of the dissolved

fraction of the solution calculated from the partial molar volume of each individual component of the solution. Also, a temperature-dependent solubility curve is included in the comprehensive model to account for the effect of temperature on the crystallization rates. These additions render this combined model robust and make it possible to explain with a high level of accuracy the temporal changes of all the component concentrations throughout the course of the reaction.

2. THEORY

Under certain assumptions described hereafter, kinetic models have the ability to accurately predict absorbance and concentration profiles for all components of a reaction system. In order to do so, a mathematical model is fitted to a set of measured spectra and is used to estimate pure component spectra without the use of any calibration. The interested reader is referred to Puxty *et al.* [12] and Billeter *et al.* [13] for a comprehensive under-

* Correspondence to: Paul J. Gemperline, Department of Chemistry, East Carolina University, Greenville, NC 27858, USA.
E-mail: gemperlinep@gmail.com

^a D. E. Joiner, J. Billeter, P. J. Gemperline
Department of Chemistry, East Carolina University, Greenville, NC, 27858, USA

^b M. E. P. McNally, R. M. Hoffman
E. I. DuPont de Nemours and Co., Inc., Crop Protection Products and Engineering Technologies, Stine Haskell Research Center, Newark, DE, 19711, USA

standing of this technique applicable to both full-rank and rank-deficient spectroscopic data.

2.1. Beer's law

Calculation of the predicted time and wavelength-resolved absorbance profiles relies on Beer's law and relates the concentrations to the molar spectra of the pure components as follows:

$$\mathbf{Y} = \mathbf{CA} + \mathbf{R} \quad (1)$$

where \mathbf{Y} is a $nt \times nw$ matrix containing the measured spectra at nt times and nw wavelengths, \mathbf{C} is a $nt \times ns$ matrix containing the concentration profiles of all the ns absorbing (or spectroscopically active) species (or pure components), \mathbf{A} is a $ns \times nw$ matrix containing the pure component spectra (molar absorptivities times the pathlength), and \mathbf{R} is a $nt \times nw$ matrix of residuals that contains the lack of model fit and the random measurement errors.

The product of the matrices \mathbf{C} and \mathbf{A} , which provides an estimate of the absorbance spectra, is used to identify a kinetic model and adjust model parameters. Hence, the quality of the model fit can be assessed via the sum of the squares of the elements of the residual matrix \mathbf{R} .

2.2. Concentration profiles

The construction of accurate concentration profiles is the foundation for an accurate kinetic model. The concentration profiles are produced by numerical integration of a differential model that can be described as a system of ordinary differential equations (ODEs), which explains at any time the changes in concentration for each species present in the system. For most multistep reaction systems, it is often impossible to integrate analytically the corresponding set of ODEs, and thus, a numerical integration must be performed. The classical Runge–Kutta method [14] of numerical integration was used for this purpose. The differential equation of each component is written in terms of rate-law equations corresponding to each postulated reaction step taking place in the system. A detailed description of this construction procedure can be found in Puxty *et al.* [12] and Gemperline *et al.* [15]. The remaining information to provide to the differential model is the inlet and dilution terms of each dosed species, the initial concentrations of each species, and the reaction rate constants of each reaction. The result of the numerical integration is a set of computed concentration profiles for each of the components of the reaction mixture, which depend on the postulated rate laws (model structure), on the values of rate constants (parameter estimation), and on the initial and dosing conditions of the experiment. In this work, the initial and dosing conditions were presumed to be known accurately, and the rate constants were adjusted. In order to estimate the values of the rate constants, a nonlinear optimization (fitting) problem had to be solved.

2.3. Newton–Gauss–Levenberg–Marquardt algorithm

The Newton–Gauss–Levenberg–Marquardt (NGLM) algorithm [16,17] is a gradient-based optimization method that is used to adjust the values of unknown nonlinear parameters, here the rate constants. The values of these constants vary according to certain intensive properties of the system, such as the temperature and the pressure. An optimization method based on the NGLM algorithm is able to find accurate estimations of unknown parameters, so as to best describe the measured data.

In order to initiate the calculation of the concentration profiles, initial guesses for all adjustable parameters are to be specified. Using successive estimates of the concentration profiles $\hat{\mathbf{C}}$, the corresponding estimated pure component spectra $\hat{\mathbf{A}}$ are calculated as follows:

$$\hat{\mathbf{A}} = \hat{\mathbf{C}}^+ \mathbf{Y} \quad (2)$$

The difference between the measured and modeled absorbance spectra corresponds to the residual matrix \mathbf{R} that is used to calculate the sum of squares ssq .

$$\mathbf{R} = \mathbf{Y} - \hat{\mathbf{C}}\hat{\mathbf{A}} = \mathbf{Y} - \hat{\mathbf{C}}\hat{\mathbf{C}}^+ \mathbf{Y} \quad (3)$$

$$ssq = \text{vec}(\mathbf{R})^T \text{vec}(\mathbf{R}) \quad (4)$$

where $\text{vec}(\cdot)$ denotes the operation of vectorization that converts a matrix into a vector.

The optimization algorithm proceeds iteratively, calculating shifts in the adjustable parameters that result in a decrease in the sum of squared residuals. The first term in a truncated Taylor series expansion [12,18] is used to calculate a shift toward the minimum of the residuals for each of the nonlinear parameters, coinciding with a set of modeled absorbance spectra that better represent the original measurements. This recursive procedure of minimizing the residuals by computing a shift vector continues until ssq converges to a minimum value, defined here as the relative difference between two successive sums of squares. For the purpose of this work, the relative convergence was set to 10^{-4} .

3. EXPERIMENTS

The chemical system chosen in this work was the synthesis of aspirin (acetylsalicylic acid), which involves the dissolution of the solid salicylic acid (SA), the reaction of SA in solution with acetic anhydride (AA) under acid catalysis to form acetylsalicylic acid (ASA), and the subsequent crystallization of ASA. This system was chosen because it has well-understood dissolution, reaction, and crystallization processes that take place in a relatively short time period.

3.1. Equipment

All experiments were run in a custom 50 mL computer-controlled reactor made in-house at East Carolina University including a heated oil jacket. This reactor works on both the principles of heat flow and power compensation [19]. Thermostated heated silicone oil is pumped through the jacket (heat flow principle), and a 20 W hastelloy immersed heating coil with a Proportional-Integral-Derivative (PID) controller was used to maintain the set temperature inside the reactor vessel (power compensation). All of these were controlled by an H.E.L. Inc. automate system operating under the H.E.L. WinISO software package (Lawrenceville, NJ). The custom reactor lid was designed in-house and contained openings for an ultraviolet–visible (UV/Vis) submersible attenuated total reflectance (ATR) probe, an immersed heating coil, a thermocouple made of hastelloy, and an inlet tubing connected to an automatic syringe pump, controlled by the Winiso software. This reactor setup was described in more details in [11].

All UV/Vis measurements were taken using a tec5USA multichannel spectrophotometer (Plainview, NY) with a resolution of 3 nm in the 190–1100 nm region connected to a fiber-optic ATR probe (Hellma 661.821 model, 205 mm, 125 mm outside diameter (o.d.) sapphire crystal Plainview, NY). UV/Vis absorbance spectra were measured at a rate of 1 spectrum/s and averaged at 30 s intervals using MultiSpec Pro Process software (tec5USA). Dark current spectra were acquired before the start of each experiment and were used for correcting the measured spectra. A mixture containing AA and phosphoric acid as a catalyst was used to acquire a reference spectrum before each run.

Validation of the spectroscopic results was performed using an offline high-performance liquid chromatography (HPLC) system. The chromatograms were obtained with an Agilent (Santa Clara, CA) Zorbax (Norcross, GA) Eclipse (Waltham, MA) C-18 (25 cm × 4.6 mm, 5 μm) column at 40 °C, with an injection volume of 15 μL and a flow rate of 1 mL/min. The mobile phase was a 60:40 water (adjusted with 1% v/v acetic acid) and methanol mixture. A ConstaMetric 4100 gradient pump (Milton-Roy, Ivyland, PA), an Alcott (Norcross, GA) 728 autosampler, an Eppendorf column heater (Hamburg, Germany), a Spectraflow 757 UV/Vis detector (Applied Biosystems, Foster City, CA) set at 254 nm, and PC/Chrom data acquisition software (H&A Scientific Inc, Greenville NC) were used for all HPLC measurements.

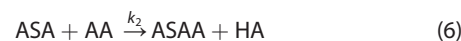
3.2. Method

A 20 mL aliquot of AA (Fisher Scientific, Waltham, MA) and a 0.2 mL aliquot of phosphoric acid catalyst (Fisher Scientific, Waltham, MA) were added to the reactor and were then heated to 55 °C. The oil jacket was maintained at 55 °C by a Julabo F25-HD heater/chiller (Allentown, PA). A 9.5 g of solid SA (Fisher Scientific, Waltham, MA) was then added to the solution and allowed to dissolve. During this period, both dissolution and reaction took place simultaneously, and the reaction mixture was continuously stirred at a rate sufficiently high to keep the solid SA suspended in the reaction mixture. Because of the large excess of AA present in solution, the desired product, ASA, underwent a second reaction and formed a side product, the acetylsalicylic anhydride (ASAA) [20–22]. After 60 min, an aliquot of 4 mL of deionized water was delivered by the syringe pump at a rate of 1 mL/min to eliminate the anhydrous side product and the remaining AA. The side product reacted with water to reform the desired product ASA along with acetic acid HA. At this point, the solution was supersaturated with ASA, and unseeded nucleation and crystallization occurred spontaneously after a short stochastic delay. The slurry of precipitated ASA and its reaction mixture was allowed to reach equilibrium for a period of 10 min, after which the reactor was cooled at a rate of 1.5 °C/min until the reactor reached 5 °C to complete the precipitation. The computer-controlled system maintained a constant temperature difference of 1.5 °C between the jacket and the reaction mixture during the cooling ramp. This portion of the experiment was added to demonstrate the robustness of the comprehensive kinetic model across varying temperatures and its ability to predict a temperature-induced crystallization. Representative examples of measured spectra are provided in Figures 1 and 2.

3.3. Reaction rate laws

Building a kinetic model, as previously mentioned, requires postulating a set of ODEs expressing the change of concentrations

for each of the species involved in the reaction. Such a construction assumes that the reaction mechanism is known or can be postulated and consequently that a correct set of rate-law equations can be established. Four reaction steps were considered in this work. In the first part of the experiment, SA was added in large excess of AA and formed the main product ASA as well as the side product ASAA.



After the addition of water, the system underwent another set of reactions that converted the side product ASAA to the main product ASA and consumed the excess of AA.



Chemical equations (5–8) can be translated into series of rate-law equations (9–12), which can then be used to construct the set of ODEs describing the evolution of the concentration profiles.

$$r_1 = k_1 c_{SA}(t) c_{AA}(t) \quad (9)$$

$$r_2 = k_2 c_{ASA}(t) c_{AA}(t) \quad (10)$$

$$r_3 = k_3 c_{ASAA}(t) c_{H_2O}(t) \quad (11)$$

$$r_4 = k_4 c_{AA}(t) c_{H_2O}(t) \quad (12)$$

In Equations (9–12), $c_s(t)$ represents the concentration of species s at time t , namely an element of the concentration matrix \mathbf{C} where s is an absorbing species. In order to construct a comprehensive model, rate-law equations for the dissolution and crystallization processes must also be included.

3.4. Dissolution and crystallization rate laws

Modeling physical transformations such as dissolution and crystallization (crystal growth) processes is different than modeling chemical transformations. Detailed models have been proposed in the literature and provide accurate descriptions of the dissolution and crystallization rates.

3.4.1. Crystal growth

The rate of crystal growth [23,24] for a solute s , denoted $r_{g,s}$, can be described as follows:

$$r_{g,s} = \frac{\Phi_s M_s k'_{c,s}}{3d_s \Phi_v} \eta_r (c_s(t) - c_s^{sat}(T))^g \quad (13)$$

where Φ_s and Φ_v are the surface and volumetric shape factors, respectively; M_s is the molecular weight; d_s is the density, η_r is the effectiveness factor [25]; and $k'_{c,s}$ is the crystallization rate constant. The term $c_s^{sat}(T)$ represents the saturation concentration of the solute s at the temperature T . In this work, the surface and volumetric shape factors were assumed to be constant

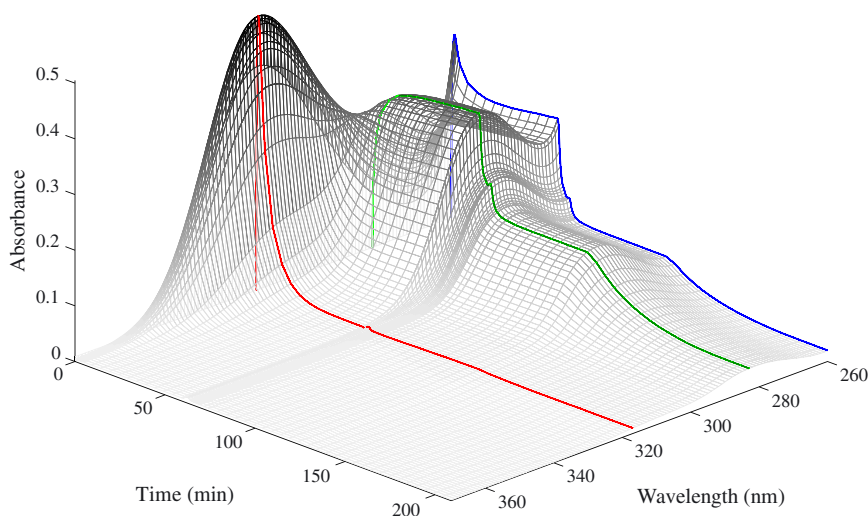


Figure 1. Attenuated total reflectance ultraviolet-visible absorbance as a function of time and wavelength with selected kinetic traces at 260 (blue), 283 (green), and 317 nm (red).

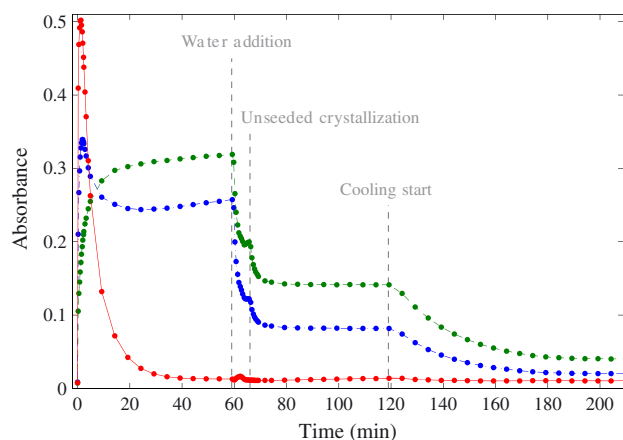


Figure 2. Kinetic traces at 260 (blue), 283 (green), and 317 nm (red), and main operating steps of the reactor (addition and reactive dissolution of SA, then addition of water, unseeded crystallization, and temperature cooling).

during the course of the experiments so that the rate law could be expressed in a simpler form, which depends on a macroscopic rate constant $k_{c,s}$ representing the corresponding accumulated terms. Another assumption was that there was no agglomeration of crystal particles or breakage of large crystals from shear forces during stirring. All these assumptions allow expressing the rate of crystal growth in a simpler form [26], as follows:

$$r_{g,s} = k_{c,s}(c_s(t) - c_s^{sat}(T))^c \quad (14)$$

where $r_{g,s}$ represents a macroscopic rate of crystal growth, whose driving force is determined by a rate constant k_c representing the accumulated terms of Equation (13) and by the degree of supersaturation $(c_s(t) - c_s^{sat}(T))$ raised to a power c .

Replacing the name of the solute s by the name of the only species undergoing crystallization during this synthesis, that is, aspirin ASA, one obtains the following rate-law expression:

$$r_g = k_c(c_{ASA}(t) - c_{ASA}^{sat}(T))^c \quad (15)$$

3.4.2. Dissolution

The rate of dissolution [26,27] for a solute s , defined as $r_{d,s}$, can be expressed in a similar way to the formulation of the rate of crystal growth presented in Equation (13):

$$r_{d,s} = \frac{2M_s k'_{d,s}}{d_s} (c_s^{sat}(T) - c_s(t)) \quad (16)$$

where $k'_{d,s}$ is the dissolution rate constant. Unlike the rate of crystal growth, however, the rate of dissolution depends on the degree of undersaturation $(c_s^{sat}(T) - c_s(t))$. Similarly to the simplifications made for rewriting Equation (13) in (14), Equation (16) can be rearranged into a macroscopic dissolution rate:

$$r_{d,s} = k_{d,s}(c_s^{sat}(T) - c_s(t))^d \quad (17)$$

where $k_{d,s}$ is a macroscopic dissolution rate constant representing the accumulated terms of Equation (16) and d is an exponent introduced to provide a higher degree of freedom to the dissolution rate expression, similarly to the exponent c defined in the crystallization rate expression (Equation (14)). When the name of the solute s is replaced by the name of the only species undergoing dissolution during the synthesis of aspirin, that is, the reactant SA, one obtains the following rate-law expression:

$$r_d = k_d(c_{SA}^{sat}(T) - c_{SA}(t))^d \quad (18)$$

3.5. Ordinary differential equations for the concentration profiles

When rate-law equations are known or postulated, one can formulate the system of ODEs describing the change of concentration and mass for each liquid and solid component of the reaction system.

$$\begin{aligned}
 \dot{m}_{SA} &= -M_{SA} V r_d \\
 \dot{c}_{SA} &= r_d - r_1 - \frac{\dot{V}}{V} c_{SA} \\
 \dot{c}_{AA} &= -r_1 - r_2 - r_4 - \frac{\dot{V}}{V} c_{AA} \\
 \dot{c}_{HA} &= r_1 + r_2 + r_3 + r_4 - \frac{\dot{V}}{V} c_{HA} \\
 \dot{m}_{ASA} &= M_{ASA} V r_c \\
 \dot{c}_{ASA} &= r_1 - r_2 - r_3 - r_c - \frac{\dot{V}}{V} c_{ASA} \\
 \dot{c}_{ASAA} &= r_2 - \frac{\dot{V}}{V} c_{ASAA} \\
 \dot{c}_{H_2O} &= -r_3 - r_4 + \frac{f}{V} c_{H_2O}^{in} - \frac{\dot{V}}{V} c_{H_2O}
 \end{aligned} \tag{19}$$

Each of the terms r represent the rates previously established in Equations (9–18). The sign and number (only ones for this reaction system) in front of each rate indicates the stoichiometric coefficients (denoted γ 's in the following) of each species in the corresponding kinetic step. The term f is the flow rate of water into the system, and V is the total volume of solution. Note that time derivatives are written using Newton's dot notation.

3.6. Ordinary differential equation for the total volume of solution

The term \dot{V} represents the change in the total volume of solution as a function of time. We assumed that volumes were additive and included terms to account for volume changes due to the dissolution of SA, change of composition during the reaction, addition of water, and crystallization of ASA. To account for these changes, the total volume of the solution can be modeled by an algebraic expression corresponding to the sum of the partial molar volumes of each component in the solution [28].

$$V = \sum_{i=1}^{ns} V_i(t) = \sum_{i=1}^{ns} n_i(t) v_i \tag{20}$$

where n_i is the number of moles of the i th species and v_i is the partial molar volume of that species.

After replacement in Equation (20) of the numbers of moles with the product of the corresponding concentrations as defined in Equation (19) and the total volume of solution, and subsequent derivation, one obtains the following expression for the change in the individual volume V_i of the i th species:

$$\dot{V}_i = v_i V \left(\sum_{j=1}^4 \gamma_{j,i} r_j + \gamma_{d,i} r_d + \gamma_{c,i} r_c + \varepsilon_i \frac{f}{V} c_{H_2O}^{in} \right) \tag{21}$$

where $\gamma_{j,i}$, $\gamma_{d,i}$ and $\gamma_{c,i}$ represent the stoichiometric coefficients of the i th species in the j th reaction in the dissolution process and in the crystallization process respectively, and $\varepsilon_i = 0$ for all species except for water for which this variable takes the value 1.

When the changes of all individual volumes are summed, one obtains an expression for the total volume of the solution:

$$\begin{aligned}
 \dot{V} &= \sum_{i=1}^{ns} \dot{V}_i \\
 &= V \sum_{i=1}^{ns} v_i \left(\sum_{j=1}^4 \gamma_{j,i} r_j + \gamma_{d,i} r_d + \gamma_{c,i} r_c + \varepsilon_i \frac{f}{V} c_{H_2O}^{in} \right)
 \end{aligned} \tag{22}$$

3.7. Effect of temperature on the experimental and model parameters

In order to properly model the crystallization of ASA a separate calibration experiment was performed and the temperature dependence of the solubility curve of ASA, which represents the main effect of temperature on the model, was determined. The solubility of the reactive slurry mixture was estimated for various temperature intervals as follows. At the end of the cooling step that led to precipitation of ASA at 5 °C, the mixture was allowed to equilibrate for 10 min. Then the temperature was raised in seven intervals of 5 °C with a 10 min equilibrium period per interval until a final temperature of 50 °C was reached. The establishment of equilibrium at each interval was verified by observing a constant absorbance at 283 nm, the wavelength corresponding to the maximum absorption of ASA. At each equilibrium interval, the concentration of ASA in solution was estimated using UV/Vis ATR measurements at 283 nm and the absorptivity coefficient of ASA at this same wavelength. The temperature dependence of the ATR probe was estimated using a mixture of ASA in HA measured over a temperature range of 5 to 50 °C, and was found to be less than 2%, which was deemed negligible for the purpose of this modeling study. The measured and modeled solubility $c_{ASA}^{sat}(T)$ as a function of the temperature is shown in Figure 3. The calibration curve that was used to model the solubility as a function of the temperature (in Kelvin) was the following:

$$c_{ASA}^{sat} = 4.04 \cdot 10^{-4} T^2 - 2.23 \cdot 10^{-1} T + 31.2 \tag{23}$$

The kinetic model describing the synthesis of aspirin thus consists of a system of eight differential equations for the concentration profiles (Equation (19)) and one differential

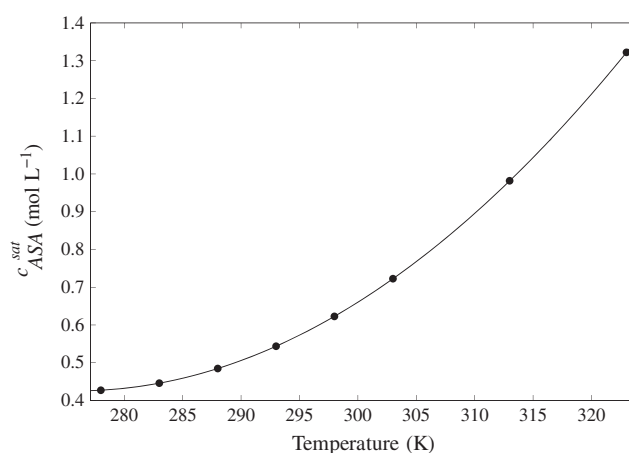


Figure 3. Measured (●) and modeled (–, Equation (23)) solubility curve of acetylsalicylic acid (ASA) as a function of temperature.

expression for the total volume of solution (Equation (22)), which depend on the rate laws (Equations (9–12), (15), and (18)) and on the solubility of ASA as a function of temperature (Equation (23)), with all equations to be solved simultaneously.

The system of ODEs formed by Equations (19) and (22) also formally depends on the temperature through the temperature dependence of the rate laws (Equations (9–12), (15), and (18)) and of the partial molar volumes (Equation (20)). However, as the first step of reactive dissolution was carried out at a constant temperature and because the excess AA, which could have potentially reacted afterward, was hydrolyzed prior to the cooling step, temperature-dependent parameters were not needed for these rate laws. Regarding the crystallization step enhanced by the cooling ramp, the assumption of quasi-stationary dynamic equilibrium was made to simplify the model description. Indeed, the rate of cooling was slower than the rate of crystallization, and hence, at any time during the cooling phase, the slurry mixture was in a situation of pseudo-equilibrium, which prevented observation of the temperature dependence of the crystallization rate constant. Additionally, the effect of the temperature drop on the partial molar volumes was neglected because the main source of volume change was the massive precipitation of ASA (see Section 4).

4. RESULTS

The NGLM algorithm adjusted the rate constants of dissolution, reaction, and crystallization to minimize the sum of squared residuals between the measured and modeled absorbance data, subject to the integration of the differential equations describing the kinetic model and subject to the estimation of the pure spectra for the absorbing species present in the slurry, that is, SA, ASA, and ASAA.

4.1. Estimated concentration profiles

The concentration profiles of the absorbing species after fitting are shown in Figure 4. Only the spectroscopically active species can be modeled directly from absorbance spectra using the method described in Equations (2–4). Attempting to directly

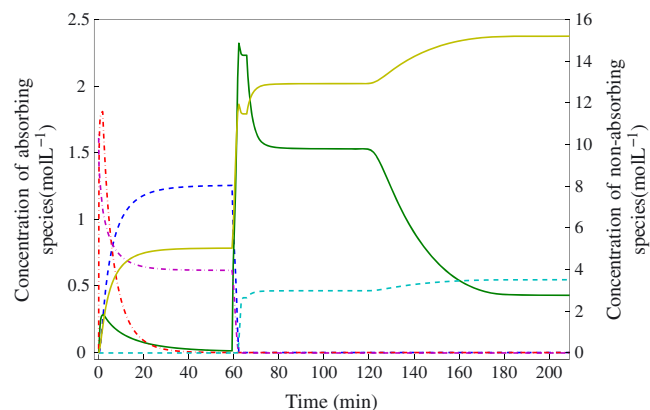


Figure 4. Concentration profiles of all species after the fitting of absorbance data by the Newton–Gauss–Levenberg–Marquardt algorithm. On the left axis, the modeled absorbing species (SA: red dash-dot line, ASA: green solid line, and ASAA: blue dashed line) are shown, and on the right axis, the reconstructed non-absorbing species (AA: violet dash-dot line, HA: yellow solid line, and H₂O: light blue dashed line).

model more than these species results in a rank-deficient matrix inverse problem during the model fitting process [29] (see Equation (2)). In this experiment, the rank of the concentration matrix **C** and of the pure component spectral matrix **A** (see Equation (1)) is three for the two following reasons: the mass balance for such a kinetic model only allows three spectroscopically active species to change in a linearly independent manner (rank condition on **C**); among these three possible spectroscopically active species, there are actually three absorbing species in this UV/Vis range (rank condition on **A**). In this particular case, the spectroscopically active species coincide with the absorbing species. From this information, the concentration profiles of the remaining non-absorbing species (AA, HA, and H₂O) can be indirectly reconstructed from the principles of mass balance, resulting in the concentration profiles shown in Figure 4.

The concentration profiles of each species follow a predictable trend (Figure 4). Solid SA, which is added at the onset of the experiment, simultaneously dissolves and reacts at 55 °C with AA and results in the main product ASA, the side product ASAA, and HA. At 60 min, water is added to eliminate ASAA and the excess AA, and regenerates ASA and HA *in situ*. This explains the small peak observed in the concentration profiles of HA and ASA at 62 min. However, to ensure a complete consumption of AA and ASAA, an excess of water is added by the syringe pump. This results in the dilution effect observed from 62 to 64 min, as the volume of the solution increases while the number of moles of each species remains constant. This decrease in concentration stops when the pump is turned off at 64 min.

4.2. Estimated absorbance profiles

Using the concentration profiles of the absorbing species at the optimum, an estimate of the absorbance spectra can be directly computed by use of the pseudo-inverse of the concentration matrix, $\hat{\mathbf{C}}^+$, in Equation (2). Computed spectra compared well with measured pure component spectra (not shown here) considering that the solvent composition changes significantly during the course of the reaction, from almost pure AA at the beginning to pure HA at the end. The comparison in Figure 5 of the measured and estimated UV/Vis kinetic traces at selected wavelengths shows the relevance of the fitting process.

The modeled absorbance profiles show good agreement with the measured ones (Figure 5), with an *ssq* value of 0.95 and a

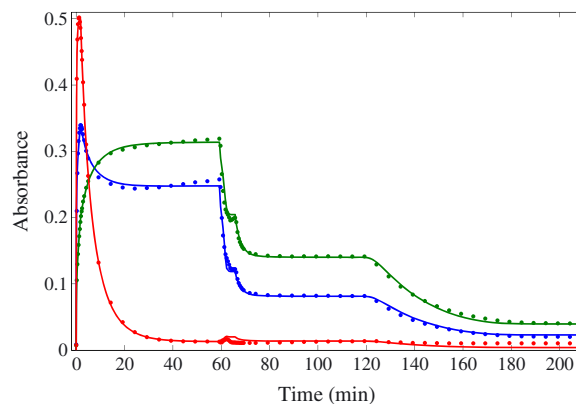


Figure 5. Measured (●) and modeled (–, Equations (1–4)) kinetic traces at 260 nm (blue), 283 nm (green), and 317 nm (red).

root mean square error of $4.29 \cdot 10^{-3}$. However a more careful look reveals some discrepancies. In particular, there is a small but noticeable increasing drift in absorbance between 20 and 60 min and again on the plateau between 65 and 67 min. We believe that this phenomenon is due to the weak adsorption of analyte on the surface of the ATR crystal, because it only takes place during periods of supersaturation. This type of phenomenon has regularly been observed in this work and has also been reported by others using UV/Vis ATR spectroscopic techniques similar to the one used here under supersaturated conditions [9].

Figure 5 shows in particular the kinetic trace isolated at 283 nm, the location of the peak maximum for the spectral band of the desired product ASA and near the maximum of the side product ASAA. At this wavelength, the estimated absorbance profile closely approximates the measured one, the main deviations being attributed to the aforementioned adsorption phenomena occurring under supersaturated conditions.

4.3. Estimated model parameters

The results of the nonlinear optimization are shown in Table I. In this work, the seven adjustable model parameters consisted primarily of reaction rate constants. In addition to the four reaction rate constants, k_1 , k_2 , k_3 and k_4 defined in Equations (9–12), and the two lumped rate constants, k_d and k_c , describing the dissolution and crystallization processes in Equations (15) and (18), the saturation concentration of SA, c_{SA}^{sat} , at 55 °C was also adjusted, because the solvent AA, in which this quantity should be estimated, also reacts with SA.

4.4. Comparison between two replicate batches

Two replicate batches were conducted for comparison purposes, as shown in Table I. Once a stable set of model parameters were found for each experiment, the nonlinear optimization process was repeated with small random perturbations of the initial values of k_2 to assess the stability of the model parameters and reproducibility of the fitting process. Optimization runs that converged to within 3% of the lowest *ssq* were included in the averages and standard deviations reported in Table I ($N=9$ and 5, respectively, for Experiments 1 and 2). The results in Table I demonstrate the relative consistency of the model fitting

process, although some noticeable differences can be observed in two of the model parameters estimated for these two datasets. In particular, the differences in the values of k_3 between the two batches are large. This can be explained by the fact that the reaction between water and ASAA (Equation (11)) is very fast and exothermic (the values of k_3 are 2 to 3 orders of magnitude larger than all the other kinetic parameters in the model); consequently, its effect can only be observed between 60 and 62 min (Figure 4). During this time, water is consumed almost as fast as it is added and stirred into solution. As a result, the parameter k_3 is poorly defined on the time scale of the measurements made in this work.

The value of k_d , which models the dissolution of SA, has poor reproducibility between the two batches because it is dependent on the rate at which powder SA is manually introduced into the reactor. Manual introduction of powdered SA was difficult to achieve in one instantaneous step, because a small amount would invariably adhere to the side of the reactor, which was subsequently rinsed into the reaction mixture with a plastic pipet and a small amount of the reaction mixture. Second-order global analysis [30] was not appropriate for these two experiments because of the conditions described in the previous two paragraphs.

In the modeling of dissolution and crystallization properties (Equations (15) and (18)), the exponents d and c were set to 1.90 and 1.34, respectively, for both datasets. It was manually determined that the values of the two exponents had little effect on the sum of squares of the residuals overall. A range of 0.20 to 4.60 was tested for the value of d , resulting in a minimum value of *ssq* for Experiment 1 at 1.90, but little or no change in the value of *ssq* in Experiment 2. The same test was performed to test the effect of changes in c . The values of *ssq* for Experiments 1 and 2 reached minimums at different values, but changes in the *ssq* value due to the increase in c became minimal at approximately 1.3, supporting tests that were performed in the beginning stages of this work to obtain the value of 1.34 that was used throughout.

Because of the simplifying assumptions made in this model, the estimated reaction rates do not hold much physical significance; however, it is important to note that the structure of the model produces accurate concentration profiles over a range of different experimental conditions. It is also important to note that the flexibility of the model to accommodate different experimental conditions is manifested as differences in fitted

Table I. Optimized values of model parameters with associated uncertainties (one standard deviation) found by nonlinear optimization for two replicate batches

	Units	Equation	Experiment 1	Experiment 2
k_1	L ² /mol ² /min	(9)	$3.40 (\pm 0.02) \cdot 10^{-2}$	$3.33 (\pm 0.05) \cdot 10^{-2}$
k_2	L ² /mol ² /min	(10)	0.49 (±0.12)	0.37 (±0.24)
k_3	L ² /mol ² /min	(11)	950 (±240)	2100 (±1200)
k_4	L ² /mol ² /min	(12)	84 (±32)	86 (±31)
k_d	L ^{0.90} /mol ^{0.90} /min	(18)	7.25 (±0.19)	1.56 (±0.30)
k_c	L ^{0.34} /mol ^{0.34} /min	(15)	1.20 (±0.15)	0.98 (±0.29)
c_{SA}^{sat} at 55 °C	mol/L	(18)	2.06 (±0.01)	2.32 (±0.01)
<i>ssq</i>		(4)	0.948 (±0.002)	1.27 (±0.02)

The values of the exponents c (Equation (15)) and d (Equation (18)) were set to 1.34 and 1.90, respectively. The *ssq* values that were used to establish the goodness of fit are also given. The rate constant, k_c , is obtained under the assumption of a quasi-stationary equilibrium with temperature.

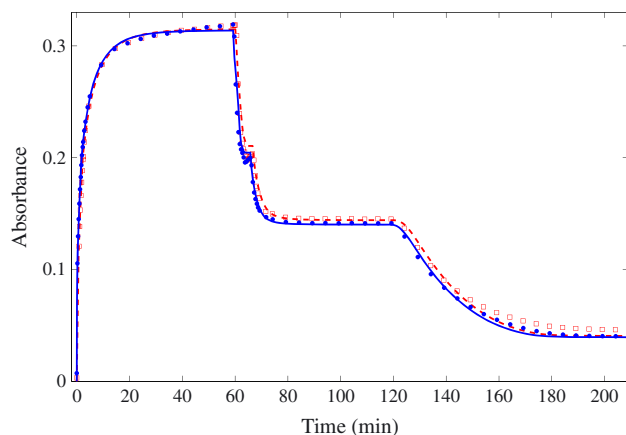


Figure 6. Measured (\bullet and \square) and modeled (— and - -) kinetic traces at 283 nm for two batch experiments (in blue and red).

parameters. Thus, the most significant benefit of this modeling approach is the calibration-free estimation of concentration profiles.

The similarities between the fitted absorbances at 283 nm for the two batches are shown in Figure 6, which displays an overlay of the results similar to those shown in Figure 5 but for the two datasets referred to in Table I. The overall shapes of the kinetic traces are very similar for the measured absorbance of the two batches as well as for the modeled and measured absorbance of each individual batch. This demonstrates the robustness both of the model and of the measurements. The differences between the two measured datasets can be attributed to slight differences in the experimental conditions of each reaction mixture.

4.5. Validation with offline high-performance liquid chromatography measurements

In order to confirm these findings, validation of the concentration profiles was performed via offline HPLC measurements. Adjustments were made to the initial experimental protocol mentioned in Section 3.2 in order to facilitate sampling for offline analysis. The cooling portion of the experimental protocol was altered as follows: the temperature was reduced by increments of 10 °C instead of being lowered at a constant rate. After equilibration at each temperature level, the stirrer was momentarily turned off, allowing the solid particles to settle to the bottom of the reactor. Because ATR spectroscopy only measures the dissolved components immediately surrounding the ATR crystal, only the solution phase was withdrawn from the reaction mixture and analyzed by HPLC. A series of three 50 μ L replicate samples were taken before crystallization at 55 °C and during crystallization at 45, 35, 15, and 5 °C. After dilution to 10 mL with mobile phase in a volumetric flask, a single injection of each sample was run through the HPLC system.

Three standard solutions of known concentrations were injected through the HPLC system in order to build a calibration curve and calculate the concentration of the unknown samples. One of the three standard solutions was injected in triplicate, thereby, the relative error in the method was estimated at approximately 1%. Figure 7 shows the comparison between the concentration of ASA in the samples determined by HPLC and the concentration profile of this species predicted by the kinetic model.

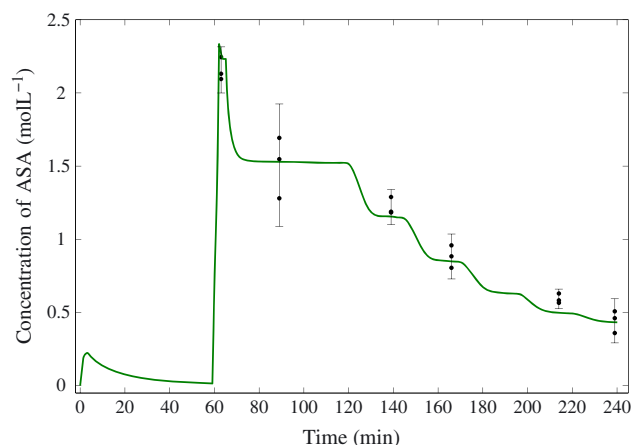


Figure 7. Measured (\bullet , offline high-performance liquid chromatography) and modeled (—) concentration profiles of ASA.

The concentration profile of ASA predicted by the model falls within the majority of the error bars, suggesting that the estimates from the kinetic model are reasonable for the experimental conditions described in Section 3. There is variability between the individual sample concentrations taken from the reaction mixture. Observations of the reaction mixture showed that despite the time allowed for the solid particles to settle to the bottom of the reactor, the solution was still slightly turbid. This indicated that some ASA solid particles may have remained suspended in the reaction mixture when the aliquot was taken, which explains the variability between the samples, especially at 90 min. It was not practical to filter the aliquots taken from the reaction mixture, because dissolved ASA would rapidly precipitate from the warm saturated solution on the colder filter.

Validation with offline HPLC measurements confirmed the proof of concept of this project, which consisted in the development of a comprehensive empirical model that accurately estimated the concentration profiles of the reactive species involved in the synthesis of aspirin, their corresponding pure component spectra, and the spectra of the reaction mixture as the process evolved over time.

5. CONCLUSION

This work established a comprehensive kinetic model that has the ability to accurately predict changes in concentration and measured absorbance of a reaction system that included dissolution, reaction, and crystallization steps. ATR UV/Vis spectroscopy was used to measure the absorbance of the solution phase of a complex reaction mixture, and a robust model was identified using an NGLM nonlinear optimization technique with numerical integration of a set of ODEs describing the concentration changes as a function of time for all species involved in the model. Incorporation of a calculation to estimate the change in reaction volume from the partial molar volumes for each species and incorporation of a temperature-dependent solubility curve to estimate the saturation concentration of the precipitated product further enhanced the precision of the proposed model. This modeling resulted in accurate predictions for the absorbance spectra, as demonstrated by the goodness of fit, as well as for the concentration profiles, as confirmed by offline HPLC measurements.

The modeling method presented here is very flexible and could be applied to other spectroscopic techniques where the relationship between absorbance and concentration is linear. The robustness of the method and its ease of implementation demonstrate the potential of this modeling technique for more complex applications encountered in chemical and pharmaceutical manufacturing.

Acknowledgements

This research was supported by the National Science Foundation (NSF) under grant number CHE-0750287, the Grant Opportunities for Academic Liaison with Industry (GOALI), and the E. I. DuPont de Nemours and Co., Inc., Crop Protection Products and Engineering Technologies. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

1. Puxty G, Maeder M, Rhinehart RR, Alam S, Moore S, Gemperline PJ. Modeling of batch reactions with in situ spectroscopic measurements and calorimetry. *J. Chemometr.* 2005; **19**(5-7): 329–340.
2. Puxty G, Neuhold YM, Jecklin M, Ehly M, Gemperline PJ, Nordon A, Littlejohn D, Basford JK, De Cecco M, Hungerbühler K. Multivariate kinetic hard-modelling of spectroscopic data: a comparison of the esterification of butanol by acetic anhydride on different scales and with different instruments. *Chem. Eng. Sci.* 2008; **63**(19): 4800–4809.
3. Nomikos P, MacGregor JF. Monitoring batch processes using multiway principal component analysis. *AIChE J.* 1994; **40**(8): 1361–1375.
4. Billeter J, Srinivasan S, Bonvin D. Extent-based kinetic identification using spectroscopic measurements and multivariate calibration. *Anal. Chim. Acta* 2013; **767**(1): 21–34.
5. Westerhuis JA, Gurden SP, Smilde AK. Spectroscopic monitoring of batch reactions for on-line fault detection and diagnosis. *Anal. Chem.* 2000; **72**(21): 5322–5330.
6. Billeter J, Neuhold YM, Puxty G, Hungerbühler K. Real-time kinetic hard-modelling for the optimisation of reaction conditions and the detection of process upset in semi-batch reactors. In *100th Annual Meeting of the American Institute of Chemical Engineering (AIChE)*. AIChE Conference Proceedings: Philadelphia (USA), 2008, available at <http://www3.aiche.org/Proceedings>.
7. Caillet A, Sheibat-Othman N, Fevotte G. Crystallization of monohydrate citric acid. 2. Modeling through population balance equations. *Cryst. Growth Des.* 2007; **7**(10): 2088–2095.
8. Cornel J, Mazzotti M. Estimating crystal growth rates using in situ ATR-FTIR and Raman spectroscopy in a calibration-free manner. *Ind. Eng. Chem. Res.* 2009; **48**(23): 10740–10745.
9. Simon LL, Nagy ZK, Hungerbühler K. Comparison of external bulk video imaging with focused beam reflectance measurement and ultra-violet visible spectroscopy for metastable zone identification in food and pharmaceutical crystallization processes. *Chem. Eng. Sci.* 2009; **64**(14): 3344–3351.
10. Yu ZQ, Chow PS, Tan RBH. 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**(1): 438–444.
11. Hsieh H, Billeter J, McNally MEP, Hoffman RM, Gemperline PJ. Kinetic modeling of dissolution and crystallization of slurries with attenuated total reflectance UV-visible absorbance and near-infrared reflectance measurements. *Anal. Chem.* 2013; **85**(11): 5367–5375.
12. Puxty G, Maeder M, Hungerbühler K. Tutorial on the fitting of kinetics models to multivariate spectroscopic measurements with non-linear least-squares regression. *Chemom. Intell. Lab. Syst.* 2006; **81**(2): 149–164.
13. Billeter J, Neuhold Y-M, Hungerbühler K. Systematic prediction of linear dependencies in the concentration profiles and implications on the kinetic hard-modelling of spectroscopic data. *Chemom. Intell. Lab. Syst.* 2009; **95**(2): 170–187.
14. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press: New York, NY, 1992; 1035.
15. Gemperline P. *Practical Guide to Chemometrics*. CRC/Taylor & Francis: Boca Raton, FL, 2006.
16. Levenberg K. A method for the solution of certain problems in least squares. *Q. Appl. Math.* 1944; **2**: 164–168.
17. Marquardt DW. An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* 1963; **11**(2): 431–441.
18. Kreyszig E. *Advanced Engineering Mathematics*. John Wiley & Sons: Jefferson City, MO, 2011.
19. Zogg A, Stoessel F, Fischer U, Hungerbühler K. Isothermal reaction calorimetry as a tool for kinetic analysis. *Thermochim. Acta* 2004; **419**(1-2): 1–17.
20. Domb AJ. Quantitative analysis of mixtures of symmetric and mixed anhydrides. *J. Chromatogr. A* 1994; **673**(1): 31–35.
21. March J. *Advanced Organic Chemistry: Reactions, Mechanisms, and Structure*. Wiley, Hoboken, NJ, 1992.
22. Florey, K. Ed., *Analytical Profiles of Drug Substances*, **3**, Academic Press: New York, NY 1974; 579.
23. Blandin AF, Mangin D, Nallet V, Klein JP, Bossoutout JM. Kinetics identification of salicylic acid precipitation through experiments in a batch stirred vessel and a T-mixer. *Chem. Eng. J.* 2001; **81**(1-3): 91–100.
24. Cornel J, Lindenberg C, Mazzotti M. Experimental characterization and population balance modeling of the polymorph transformation of L-glutamic acid. *Cryst. Growth Des.* 2009; **9**: 243–252.
25. Garside J. Industrial crystallization from solution. *Chem. Eng. Sci.* 1985; **40**(1): 3–26.
26. Higuchi WI, Nelson E, Wagner JG. Solubility and dissolution rates in reactive media. *J. Pharm. Sci.* 1964; **53**(3): 333–335.
27. Costa P, Sousa Lobo JM. Modeling and comparison of dissolution profiles. *Eur. J. Pharm. Sci.* 2001; **13**(2): 123–133.
28. Atkins PW, De Paula J. *Elements of Physical Chemistry*. Oxford University Press: Oxford, UK, 2005.
29. Billeter J, Neuhold YM, Hungerbühler K. Kinetic hard-modelling and spectral validation of rank-deficient spectroscopic data: a case study. *Chemom. Intell. Lab. Syst.* 2009; **98**(2): 213–226.
30. Dyson RM, Kaderli S, Lawrance GA, Maeder M. Second order global analysis: the evaluation of series of spectrophotometric titrations for improved determination of equilibrium constants. *Anal. Chim. Acta* 1997; **353**(2-3): 381–393.