International Journal of Pharmaceutics 405 (2011) 63–78

Contents lists available at ScienceDirect

International Journal of Pharmaceutics

journal homepage: www.elsevier.com/locate/ijpharm

Quality-by-Design (QbD): An integrated process analytical technology (PAT) approach for a dynamic pharmaceutical co-precipitation process characterization and process design space development☆

Huiquan Wu*, Maury White, Mansoor A. Khan

Division of Product Quality Research (DPQR, HFD-940), OTR/OPS/CDER/FDA, The FDA White Oak Campus, Life Science Building 64, 10903 New Hampshire Ave, Silver Spring, MD 20993, United States

ARTICLE INFO

Article history:
Received 29 September 2010
Received in revised form
15 November 2010
Accepted 24 November 2010
Available online 5 December 2010

Keywords:
Quality-by-Design (QbD)
Process analytical technology (PAT)
Co-precipitation
Real-time process monitoring
Critical process variable
Design space

ABSTRACT

The aim of this work was to develop an integrated process analytical technology (PAT) approach for a dynamic pharmaceutical co-precipitation process characterization and design space development. A dynamic co-precipitation process by gradually introducing water to the ternary system of naproxen–Eudragit L100–alcohol was monitored at real-time *in situ* via Lasentec FBRM and PVM. 3D map of count-time-chord length revealed three distinguishable process stages: incubation, transition, and steady-state. The effects of high risk process variables (slurry temperature, stirring rate, and water addition rate) on both derived co-precipitation process rates and final chord-length-distribution were evaluated systematically using a 3^3 full factorial design. Critical process variables were identified via ANOVA for both transition and steady state. General linear models (GLM) were then used for parameter estimation for each critical variable. Clear trends about effects of each critical variable during transition and steady state were found by GLM and were interpreted using fundamental process principles and Nyvlt's transfer model. Neural network models were able to link process variables with response variables at transition and steady state with R^2 of 0.88–0.98. PVM images evidenced nucleation and crystal growth. Contour plots illustrated design space via critical process variables' ranges. It demonstrated the utility of integrated PAT approach for QbD development.

Published by Elsevier B.V.

1. Introduction

1.1. Some key concepts and tools in relevant regulatory documents

According to the definition of ICH Guideline Q8(R2) (FDA/ICH, 2009), "Quality-by-Design (QbD) is a systematic approach to development that begins with predefined objectives and emphasizes product and process understanding and process control, based on sound science and quality risk management." During the past few years, the pharmaceutical quality (Woodcock, 2004) and QbD implementation strategies (Yu, 2008; Rathore and Winkle, 2009; Smith et al., 2009) for various pharmaceutical product areas have received a lot of attentions. This pharmaceutical quality paradigm shift as collectively highlighted by the FDA's PAT Guidance (FDA, 2004) and the ICH Q8, Q8(R2), Q9, and Q10 guidelines (FDA/ICH, 2006a,b, 2007, 2009) has laid a foundation for the science-based and risk-based regulatory processes.

In the QbD framework, understanding the impact of raw material attributes and process parameters on the critical quality attributes (CQAs) as well as identification and control of sources of variability are essential. However, pharmaceutical products and processes are complicated and multivariate in nature. Understanding of the relevant multi-factorial relationships among formulation parameters, process variables, and product quality attributes usually requires the use of multivariate approaches, such as statistical design of experiments (DOE) and multivariate data analysis (MVDA). Although both DOE and MVDA as standard techniques (Montgomery, 2000; Johnson and Wichern, 2007) have been widely used for many years in various areas, the integrated use of both of them are relatively few when applied to pharmaceutical development (Hwang et al., 1998; Naelapaa et al., 2008; Huang et al., 2009). Over the past few years, some internal efforts (Wu and Hussain, 2005b; Wu and Khan, 2009; Wu et al., 2009; Xie et al., 2008) have been made in the aspect of integration of DOE and MVDA for pharmaceutical product and process understanding, process monitoring and control. In principle, this integrated

☆ *Disclaimer*: The views and opinions expressed in this paper are only those of the authors, and do not necessarily reflect the views or policies of the FDA.

* Corresponding author at: DPQR, HFD-940, OTR/OPS/CDER/FDA, The FDA White Oak Campus Life Science Building 64, Room 1080, Silver Spring, MD, United States. Tel.: +1 301 796 0022; fax: +1 301 796 9816.
E-mail addresses: Huiquan.Wu@fda.hhs.gov, huiquan_wu@yahoo.com (H. Wu).

0378-5173/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.ijpharm.2010.11.045

approach in conjunction with appropriate real time process analytical technology (PAT) process analyzer should provide ample opportunities for enhanced process understanding, process monitoring, and process control of other pharmaceutical processes (such as crystallization, co-precipitation, etc.) under the QbD/PAT framework.

1.2. Measurement and control of particle size and particle size distribution in pharmaceutical particulate processes

In general the control of particle size of active pharmaceutical ingredients (APIs) is critical to the successful manufacture of most drug product formulations (Am Ende and Rose, 2006). It is a well known fact that many formulation and process variables could manifest together to impact particle size distribution (PSD), which is ultimately linked to the drug dissolution and bioavailability. Therefore, the ability to accurately measure and effectively control the crystal shape and crystal size distribution (CSD) is essential to ensure that desired properties can be achieved for crystalline APIs or drug components.

Co-precipitation as an important technique has been adapted to improve the dissolution characteristics of some water-insoluble drugs in drug delivery systems. Some formulation development work had been reported previously (Khan et al., 1994; Zaghloul et al., 2001). However, our ability to both monitor the co-precipitation process in real-time and link the formulation and process variables to the key quality attributes of co-precipitate (e.g., PSD) has been very limited (Wu and Hussain, 2005a; Yu et al., 2004; Wu and Khan, 2010a), which may present a challenge for co-precipitation process design space development.

On-line/in-line/at-line PAT process analyzers offer unprecedented opportunities to monitor the process in real-time, thus enable us to gain dynamic process information in real-time. Analyzing these real-time dynamic process data via appropriate techniques including MVDA would enable us to extract critical product/process information and knowledge, which is essential for rationale product and process design, and process control (Wu et al., 2007a).

A typical process kinetic study invokes the change-one-variable-at-a-time strategy, which is essentially a univariate approach. Following this approach, a selected model is used to fit the experimental data. However, possible interactions among various formulation and process variables may make the traditional univariate approach inefficient. As being discussed previously (FDA, 2004; Olivieri et al., 2006; Wu et al., 2008, 2009; Wu and Khan, 2009), a multivariate approach may provide some alterative solutions.

1.3. Focused beam reflectance measurement (FBRM): advantages and challenges

Among the PAT tools currently available, as an *in situ* real-time particle measurement technique, Lasentec FBRM has been demonstrated some unique capabilities and usefulness (Barrett and Glennon, 1999; Kougoulos et al., 2005) in the aspects of process monitoring and measurement. For example, it has capability of acquiring chord-length-distribution (CLD) data and population trends of particles in slurry on-line and in real time; there is no need for sampling, dilution, separation or isolation that may contribute to changes in particle size and distribution due to breakage or agglomeration. The measurement principle of FBRM can be found elsewhere (Greaves et al., 2008). It was shown that online CLD measurements might be enough for some processes to monitor the process dynamic behaviors especially relevant to PSD, particle shape, concentration and rheological behavior of fluid suspensions (Richmond et al., 1998). In recent years, FBRM has been widely used for various pharmaceutical crystallization studies such as process development at various scales, process assessment, polymorphic transformation process monitoring, control of crystal product quality, nucleation kinetics study, process screening, etc. (Togkalidou et al., 2004; Chew et al., 2007; Greaves et al., 2008; Howard et al., 2009; Lindenberg and Mazzotti, 2009). Those applications did adequately address the problems under investigation. However, most of these applications followed a univariate or change-one-variable-at-a-time approach. This univariate approach may not be sufficient for handling the multi-factor relationships among formulation and process variables, and quality attributes of crystal products, which are frequently encountered during the pharmaceutical process design space development. As discussed previously, an alternative multivariate approach may provide insights about the complicated relationship and possible interactions among the formulation variables, process variables, and product quality attributes, as well as identifies critical process variables which are essential for developing a pharmaceutical process design space (FDA/ICH, 2006a, 2009). This multivariate approach was used in this study.

Some models have been proposed to convert CLD to PSD (Simmons et al., 1999; Langston and Jones, 2001; Hukkanen and Brattze, 2003; Ruf et al., 2000; Li and Wilkinson, 2005). In spite of having gained success in estimating the PSD from the CLD data for some simple particulate systems in static mode, the theory behind these methods require quite a few assumptions, including that the particles perfectively backscatter light at all angles and that all particles have a known regular shape. The assumptions may not be valid for some particulate systems, especially for systems in dynamic mode, i.e., the particle evolution processes during which particle concentration, morphology and shape, and property are functions of time. A few works (Monnier et al., 1996; Yu and Erickson, 2008) highlighted the limitations of the laser reflection method and FBRM measurement. Therefore, at the present time, using a first-principles approach alone to handle either the FBRM measurement process or the relationship between CLD and PSD remains a significant scientific challenge. Given that very limited real-time PSD measurement technique is available for a direct side-by-side comparison, an alternative approach, i.e., using an integrated PAT approach for real-time FBRM process monitoring in conjunction with DOE and MVDA, then checking the results from fundamental process engineering knowledge and established theories for particulate process, may constitute a practical option for pharmaceutical particulate process characterization and process design space development.

1.4. Risk-based approach to select variables for the DOE study

To realize the advantages of a holistic approach for a complex process and make sure our limited resources are used efficiently and effectively, a risk assessment tool was applied to brainstorm and identify potential variables that can have an impact on the desired quality attribute. In this work, a Fishbone diagram or Ishikawa Diagram 1 was created to structure the process of identifying possible causes that could have impacts on co-precipitation process rate, as shown in Fig. 1.

Even though at a first glance six categories of factors (materials, process, measurement, equipment, people, and environment) could have impact on the co-precipitation process rate, in reality of a research laboratory setting, a large number of factors can actually be eliminated safely. For example, co-precipitation formulation components selection and pre-dissolution in solvent, co-precipitation vessel selection, errors due to operators and analytical side, environmental variables, and measurement

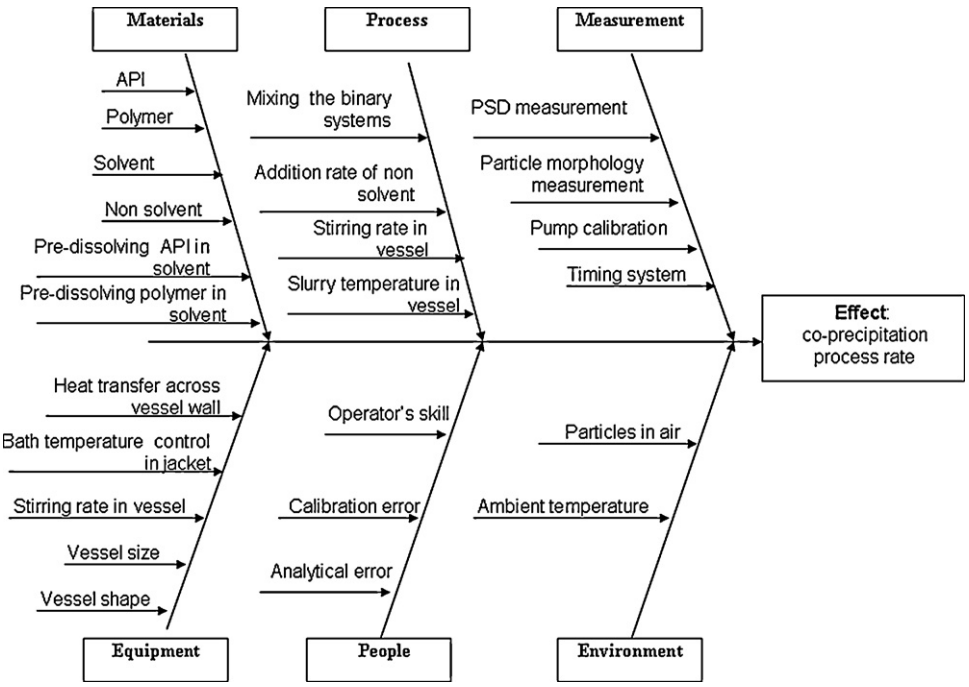

Fig. 1. Fishbone diagram for effects on the co-precipitation process rate.

variables, could be pre-specified in the study protocol based on preliminary study thus be fixed during the study. Therefore, with this risk mitigation and risk management strategy in place, we could then focus our attention on evaluating the impact of main process variables or high risk variables (being known by prior knowledge, such as slurry temperature, slurry stirring rate, and non-solvent addition rate, etc.) on process dynamics.

1.5. Rationale for co-precipitation process design space development

When developing a pharmaceutical design space, one rationale is to first identify critical process variables (Wu and Hussain, 2005b; Wu and Khan, 2009; Xie et al., 2008) and interactions, then use good science and engineering practice to manage, mitigate, eliminate, and control the risk associated with process and product variability to ensure the product quality is built in or by design. In this work, naproxen, Eudragit L100, alcohol, and water were selected as the model drug, polymer, solvent, and no solvent for the co-precipitation formulation components. A FBRM D600L system was used to monitor the co-precipitation process *in situ* at real-time during the entire course of introducing water to the ternary system of naproxen–Eudragit L100–alcohol for total of 27 batches based on a 3^3 factorial design. A 3D count-time-size plot was constructed to map the entire co-precipitation process and identify distinguishable process stages. Derived co-precipitation process kinetics was extracted to examine the impact of main process variables on the process outcomes. DOE in conjunction with analysis of variance (ANOVA) was used as a screening tool to identify the statistically significant variables and interactions for the co-precipitation process. A linkage between various formulation and process variables and the derived process rates based on the change of CLD over time was then constructed. The impact of main process variables on final CLD was also examined.

2. Experimental

2.1. Materials and experimental protocol for a dynamic co-precipitation process

Naproxen USP (lot no: NPX 368) was obtained from Albemarle Corporation (Orangeburg, SC). Eudragit L100 (lot no: 1221203048) was obtained from Röhm America Inc. (Somerset, NJ). Solvent reagent alcohol (HPLC grade, lot no: 053546) was purchased from Fisher Scientific (Fair Lawn, NJ). Non-solvent DI water was obtained from a FDA in-house facility (Silver Spring, MD) and was kept in refrigerator at 3 °C prior to use. All of these chemicals, solvent, and non-solvent were used without any further processing or purification prior to the use for this experimental work.

Our initial work examined the effects of drug/polymer ratio on co-precipitation process (Wu and Khan, 2010b). In this work, the drug/polymer ratio of 2 (4.0 g of naproxen and 2.0 g of Eudragit L100) was selected due to the fact that nice co-precipitates with regular morphology could be obtained via the dynamic co-precipitation process designed in this work. In addition, this allowed us to focus our attention on the main process variables. The naproxen and Eudragit were dissolved separately in 150 and 100 ml of reagent alcohol, respectively. These solutions were then added to the 1 L reaction vessel of Chemglass reaction kit (ChemGlass, Vineland, NJ) via a funnel. The stir speed was set via the OptiChem digital overhead stirrer. The temperature was set via the ThermoScientific Neslab RTE 740 thermoregulator. The solution in the reaction vessel was allowed to reach the pre-defined temperature prior to data acquisition. Once this temperature had been achieved, real-time monitoring began using the Lasentec FBRM system and PVM system, prior to any water addition. After establishing a baseline, the first addition of water (125 ml) kept at about 3 °C was made to the reaction vessel at the pre-defined flow rate via a flexible tubing (Tygon R-3603, Fisher Scientific), a MasterFlex solid state speed control device (Chicago, IL), and a Cole-Parmer peristaltic pump (Chicago, IL).

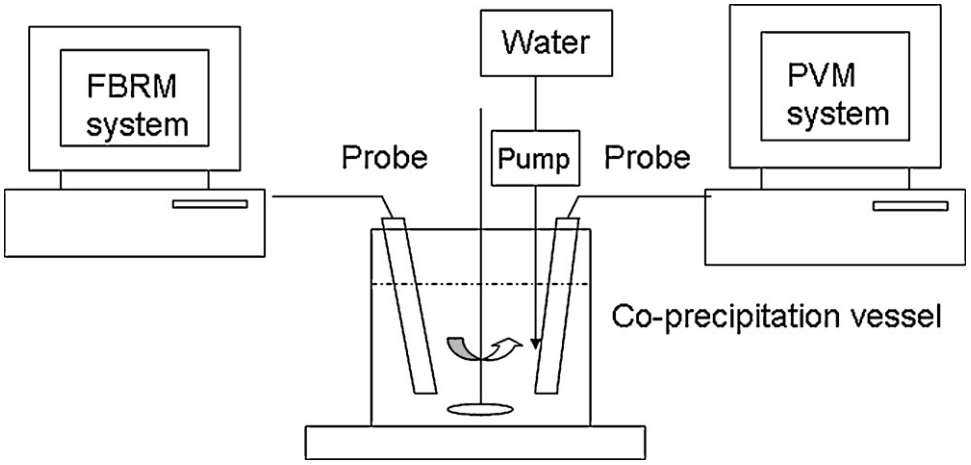

Fig. 2. Schematic of experimental set up.

Then the co-precipitation system was allowed to come to equilibrium over a period of 10 min, after which, the 2nd addition of water (50 ml) was made to the system. Depending on the temperature of the individual experiment (15 °C, 25 °C, or 35 °C), the system was introduced with water for the 3rd, 4th, and 5th addition (each time with 50 ml), respectively. The co-precipitation system was allowed to reach equilibrium over a period of 10 min after each individual addition of water. To ensure both the experimental consistency and good dispersion of water to the ternary system, the flexible tubing was kept vertically by two supporting claps attached to a lab frame. Furthermore, the end portion of the tubing was immersed in the liquid phase during the entire course of experiment. The schematic of experimental setup is shown in Fig. 2.

2.2. Co-precipitation thermodynamics and crystal growth model

A traditional thermodynamics and kinetics study is out of the scope of this work. Both naproxen and Eudragit L100 are soluble in alcohol but practically insoluble in water. The literature data for the experimental and predicted solubility of naproxen in water are 15.9 mg/L and 51.0 mg/L, respectively (http://www.drugbank.ca/drugs/DB00788, accessed on 09/03/2010). Due to the solubility difference in alcohol and water for both naproxen and Eudragit L100, introducing water to the ternary system of naproxen–Eudragit L100–alcohol gradually will lead to the movement of the overall composition point of 4 components system within the phase diagram accordingly. The super-saturation will be generated gradually once the overall composition point comes across the solid-liquid equilibrium line. Once a certain degree of super-saturation is attained, a phase change (nucleation event) will be initiated. Following this initiation, the super-saturation generated previously will be consumed by subsequent nucleation and/or co-precipitate growth process. In general, the nucleation rate per unit mass of solvent may be expressed by the following semi-empirical equation:

$$B = k_b \mu_k^j \Delta c^b \quad (1)$$

where B is nucleation rate, k_b is nucleation rate constant, μ_k is the kth moment of crystal size distribution, Δc is concentration driving force, j and b are exponent of magma density and nucleation order, respectively (Tavare, 1987).

As shown in Fig. 3, a multi-phase diffusion layer model was proposed by Nyvlt et al. (1985) to describe the crystal growth process. According to this model, the crystal growth process takes place via the following steps:

(1) Bulk transfer: transfer of solute from the bulk solution to the diffusion layer;
(2) Diffusion: diffusion of solute through the diffusion layer, whose thickness depends on the hydrodynamic conditions in the solution;
(3) Surface integration: incorporation of the solute molecules into the crystal lattice.

In this work, attempts were made to use the aforementioned thermodynamics, kinetics, and Nyvlt's multi-phase mass transfer model (Nyvlt et al., 1985) to explain the experimental observations and to link them with statistical data analysis results of the DOE data sets.

2.3. Design of experiments (DOE), real-time process monitoring, and data analysis

Based on risk assessment results from Section 1.4, three main process variables (slurry temperature (°C), slurry stirring rate (rpm), non-solvent addition rate (ml/s)) were selected as independent process variables in this study. Each factor was tested with three levels. A 3^3 full factorial design was used for experiments as shown in Table 1.

A FBRM D600L (Mettler-Toledo AutoChem, Columbia, MD) system was used to monitor the dynamic co-precipitation process *in situ* at real-time for a total of 27 batches based on a 3^3 factorial design. The FBRM D600L system was set to collect CLDs *in situ* with a laser speed of 2 m/s, applying the standard 'F-electronic' mode and using measurement duration of 2 s. A Lasentec Particle Vision Microscopy system PVM819 (Mettler-Toledo AutoChem, Columbia, MD) was used to monitor the co-precipitation process *in situ* at real-time for the selected batches, in order to obtain direct information about the morphology and size of the co-precipitates evolved during the course of adding water to the system. The PVM image was acquired every 2 s.

A 3D count-time-size map was constructed to illustrate the progress of a co-precipitation process and identify various distinguishable process stages. The impacts of main process variables on the process characteristics such as derived co-precipitation process rates and final CLD were assessed. ANOVA was conducted

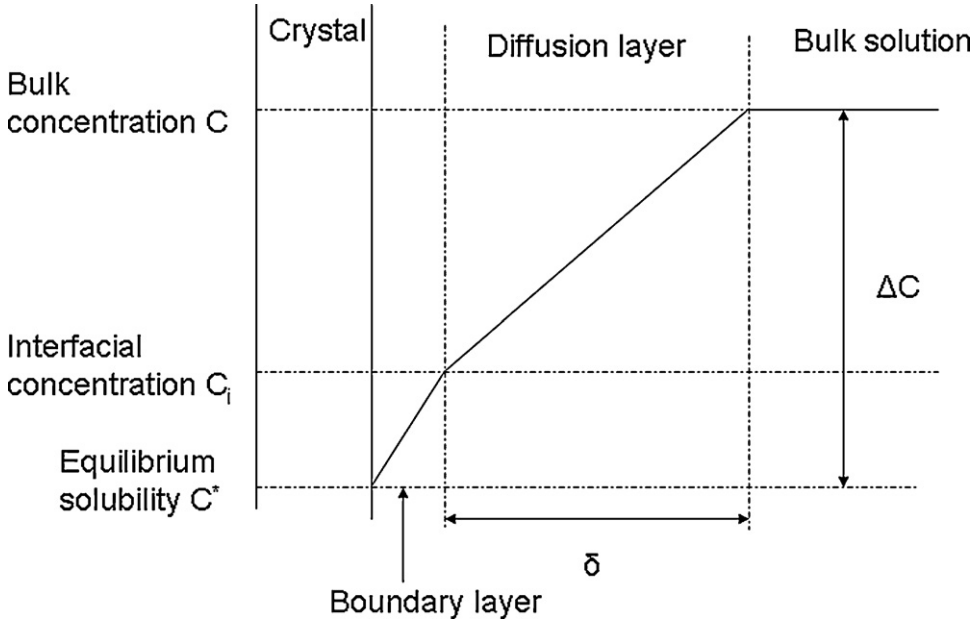

Fig. 3. Nyvlt et al. diffusion layer model (Nyvlt et al., 1985).

to identify statistical significant variables. General linear models were used for parameter estimate. Neural network algorithm was utilized to link the process variables with response variables.

3. Results and discussion

From a process engineering perspective, various process variables could impact thermodynamics, kinetics, and mass-transfer of a co-precipitation process to various extents, depending on the combination of process conditions. Therefore, the effects of main process variables on the co-precipitation process dynamics and quality attributes of final co-precipitate slurry were examined systematically.

Table 1
The 3^3 full factorial design for the co-precipitation study.

Run No	Pattern	X1 (°C)	X2 (rpm)	X3 (ml/s)
1	133	15	400	6.66
2	213	25	50	6.66
3	312	35	50	3.66
4	132	15	400	3.66
5	331	35	400	0.83
6	123	15	200	6.66
7	323	35	200	6.66
8	221	25	200	0.83
9	313	35	50	6.66
10	222	25	200	3.66
11	321	35	200	0.83
12	212	25	50	3.66
13	311	35	50	0.83
14	131	15	400	0.83
15	211	25	50	0.83
16	332	35	400	3.66
17	223	25	200	6.66
18	112	15	50	3.66
19	122	15	200	3.66
20	231	25	400	0.83
21	232	25	400	3.66
22	121	15	200	0.83
23	233	25	400	6.66
24	322	35	200	3.66
25	113	15	50	6.66
26	111	15	50	0.83
27	333	35	400	6.66

3.1. A Typical profile of FBRM counts/s in a specific chord length range vs. time during a co-precipitation process

A typical profile of FBRM counts/s in a specific chord length range vs. time was presented in Fig. 4. The following chord length (CL) range codes were used throughout this paper: CL1: 0–10 μm; CL2: 10–19.953 μm; CL3: 19.953–50.119 μm; CL4: 50.119–100 μm; CL5: 100–251.189 μm. The water was added three times with a time interval of approximately 10 min. A relative small increase of counts was observed after the initial addition. A large increase of counts was observed after the 2nd addition. However, the increases of counts at various CL ranges were not uniform. The most dominant increase occurs at CL3; the 2nd largest increase occurs at CL2 and CL4. After the 3rd addition, only a small increase of counts was observed. Therefore, the data for the 3rd addition of water were not included in further analysis. This non-uniform increase phenomenon was observed almost universally for most of the rest 26 batches with different combinations of process variables.

3.2. 3D count-time-size plot of the dynamical co-precipitation process

In order to examine the co-precipitation process progress, a 3D map of count-time-size can provide visualization of the co-precipitation process dynamics as a function of time. In addition, it can illustrate the process progress and distinguish various process stages (Wu and Khan, 2010a), as shown in Fig. 5(a) and (b) (for process runs 132 and 123).

It was found for most runs, the 3D maps consist of three stages: (1) incubation period when there was no significant particle counts detected, its time length depending on the combinations of process conditions; (2) transition period when significant nucleation and growth occurs thus significant counts were detected; (3) steady state when the number of counts per second for small chord length ranges (CL1 and CL2) remains relatively constant throughout the remainder of the process, which indicates no significant nucleation occurs. In the CL3 range, there is a slight increase after the 2nd addition of water and a slight decrease after the 3rd addition of water for run 132; there is a

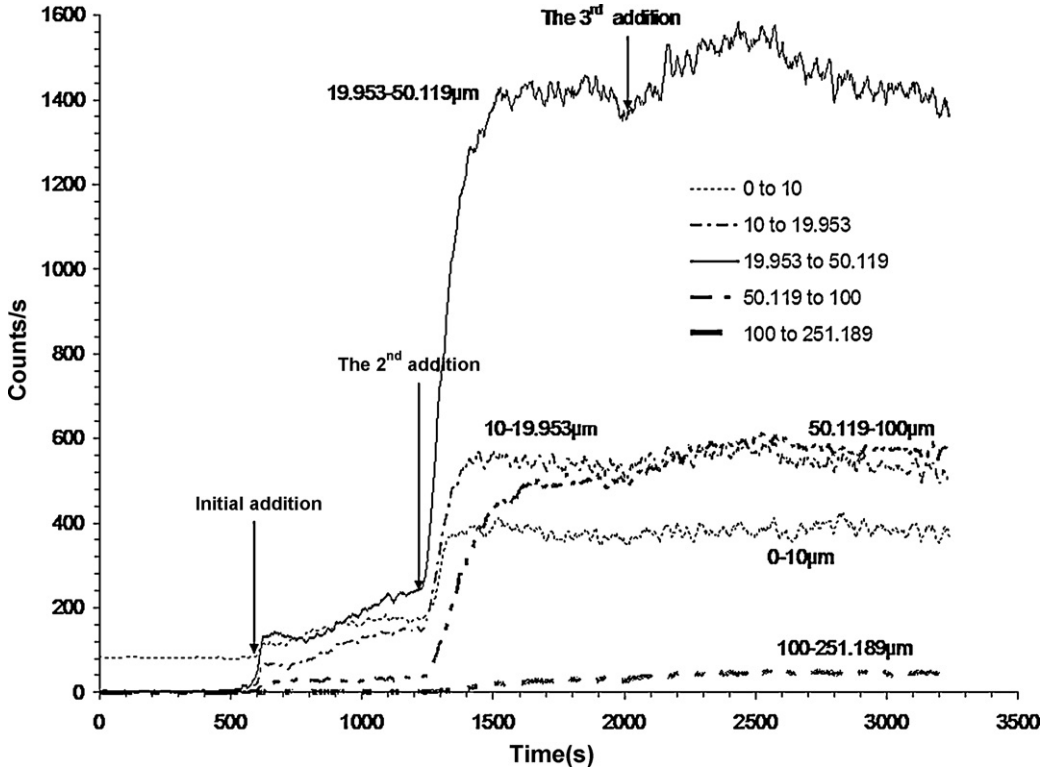

Fig. 4. A typical profile of CLD vs. time for batch run 111.

slight decrease after the 2nd and 3rd addition of water for run 123.

3.3. Derived process rates for a co-precipitation process during the transition period

Although CLD is a relative value, it can be used to represent the process characteristics of a particulate system because it depends on both PSD and shape of the particles (both are a function of process time). A recent report (Ma et al., 2009) demonstrated that the crystal mean chord length (MCL) measured by FBRM can be converted successfully to the crystal mean size (MS) measured by a digital photo-technique. It verifies the reliability of the FBRM data.

Based on the profile of the FBRM counts/s vs. time at specific CL range during the transition period, a new procedure was developed in this work to derive the co-precipitation process rates from FBRM counts/s vs. time profiles. The slopes were based on the FBRM counts/s over time in five different CL ranges. They were calculated by (1) first identifying the general time range where the maximum slope occurred for a given size range during a single addition; (2) then calculating the slope of that CL range over the preceding 20 s for each individual time point within that CL range; (3) the maximum of these calculated slope values for each addition and each CL range was used as the derived co-precipitation process rate for the corresponding process conditions at specific single addition and CL range. Following this procedure, the slope of BBRM counts/s over time for each run after each addition of water was obtained. The following derived process rate codes were used throughout the paper: Rate1, Rate2, Rate3, Rate4 are corresponding to the derived rates at CL1, CL2, CL3, and CL4 after the 1st water addition; Rate5, Rate6, Rate7, Rate8 are corresponding to the derived rates at CL1, CL2, CL3, and CL4 after the 2nd water addition. These derived process rate data as the designated response variable values were listed in Table 2. For each run and addition, a plot of the derived process rate vs. CL range was made. Fig. 6 is one example of such a plot, which represents a typical bell-shape distribution. As discussed below, this newly derived parameter was used to characterize the dynamic transition behaviors of nucleation and growth of co-precipitates after water addition, based on 2D count-time information for each CL range.

3.4. Exceptions for the derived process rate distributions over CL for runs 132 and 133

It reveals that for the majority of the 27 runs the plot of derived process rate vs. CL range has a bell-shape distribution with the maximum derived process rate occurring at the CL3. However, there are two exceptions for runs 132 and 133. At the 2nd addition of water, as shown in Fig. 7, these two runs first displayed a plateau of derived process rate over the CL range of [0, 50.12] μm, then followed by a sharp decreasing over the CL range of [50.12, 251.18] μm.

In these two cases, a combination of a lowest temperature (15 °C), highest stirring rate (400 rpm), and moderate to high addition rate of water was used. The lowest temperature may provide sufficient super-saturation as the thermodynamic driving force, which eliminates the possibility of the process being thermodynamic control. From the transport process perspective, the highest stirring rate may facilitate two sequential mass-transfer steps: bulk transfer and diffusion through the diffusion layer whose thickness is inversely proportional to the stirring intensity. Therefore, according to Nyvlt diffusion layer model (Nyvlt et al., 1985), the most likely rate-limiting step for this multi-phase transfer process could be the surface integration, i.e., the integration of species onto the nuclei or crystal's surface, which is unlikely depending on the particle size. Therefore a more uniform derived process rates were obtained for runs 132 and 133 at the 2nd addition of water, compared to the bell-shape rate distributions for runs with other combinations of process variables.

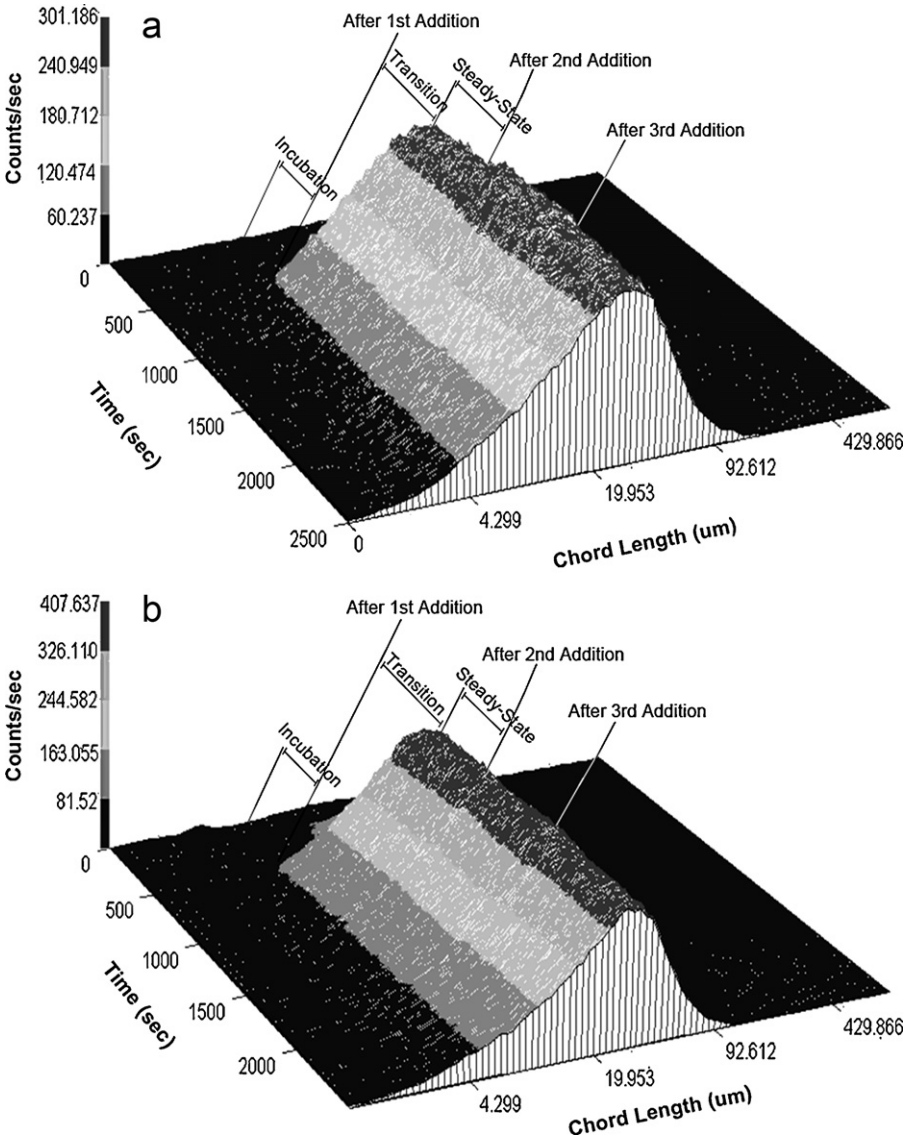

Fig. 5. 3D counts-chord length-time map for a co-precipitation process. (a) Run 132; and (b) run 123.

3.5. CLD at the end of steady state period

To characterize the effects of main process variables on the steady state, the CLDs at the end of steady state of the 27 runs were plotted. Fig. 8 is a representative example for runs 131, 231, and 331. All of the CLDs have bell shapes. As characteristic parameters of this bell-shaped CLD, its mode and peak frequency were taken as additional designated response variables and listed in the last section of Table 2.

3.6. Analysis of variance (ANOVA) and general linear modeling for both transition period and steady state period

Statistical data analysis and general linear modeling were performed on data from both transition period and steady state period to determine: (1) if there are any relationships between the process variables and the slopes at various CL ranges between [0, 100] μm during the transition period; and (2) if there are any relationships between the process variables and the quality attributes of the final co-precipitate slurry at the end of steady state period.

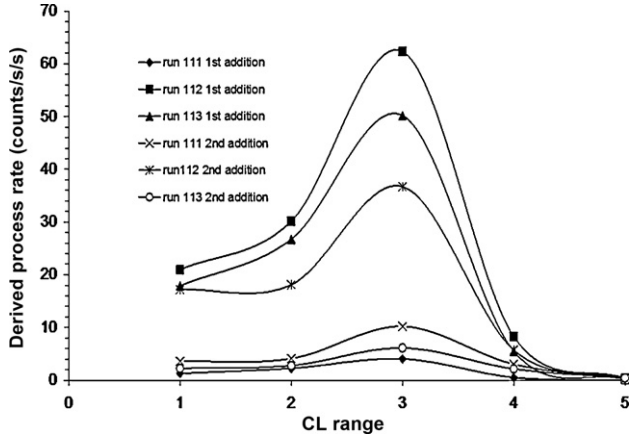

Fig. 6. Typical profiles of the derived process rate vs. chord length range for batch runs 111, 112, and 113.

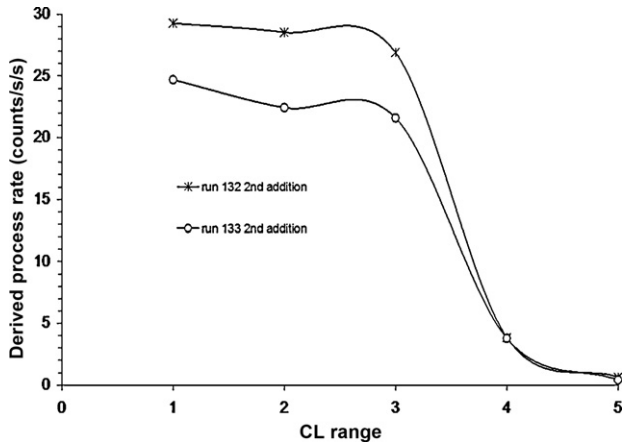

Fig. 7. Profiles of the derived process rate vs. chord length range for the 2nd water addition of runs 132 and 133.

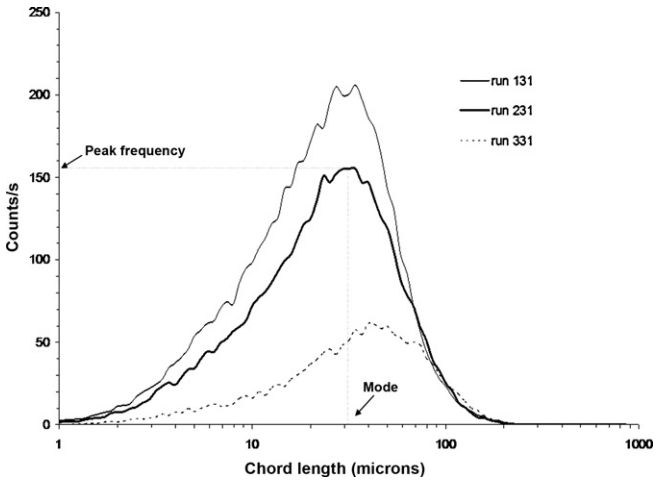

Fig. 8. A typical CLD at the end of steady state period for runs 131, 123, and 331.

3.6.1. ANOVA for the DOE results of the co-precipitation process: identifying critical process variables and interactions

Let X1, X2, and X3 be the slurry temperature, stirring rate, and water addition rate for the 3^3 full factorial design, respectively. With 27 runs of 3^3 factorial design of experiment, the ANOVA indicates that all main effects and two-way interactions among these main effects are estimable and the three-way interaction is used to estimate the underlying error. Therefore, the following statistical model was used:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + \varepsilon_{ijk}$$
$$i = 1, 2, 3, \quad j = 1, 2, 3, \quad k = 1, 2, 3 \qquad (2)$$

where y_{ijk} is the observation (either derived process rate value, mode, or peak frequency) in the ith level of variable X1 and the jth level of variable X2 and the kth level of variable X3, μ is the overall mean, α_i is the ith X1 effect, β_j is the jth X2 effect, γ_k is the kth X3 effect, $(\alpha\beta)_{ij}$ is the two-way interaction between the ith X1 effect and the jth X2 effect, $(\alpha\gamma)_{ik}$ is the two-way interaction between the ith X1 effect and the kth X3 effect, $(\beta\gamma)_{jk}$ is the two-way interaction between the jth X2 effect and the kth X3 effect, ε_{ijk}, the random error, is normally independently identically distributed with mean 0 and variance σ^2, which is written as $\varepsilon_{ijk} \sim iid\ N(0, \sigma^2)$.

Based on this statistical model, ANOVA was conducted to screen significant factors and interactions. The screening results were listed in Table 3. As shown in the last column of Table 3, at significance level $\alpha = 0.05$, X1 and X3 are statistically significant for both Rate1 and Rate7; X1, X2, and X3 are statistically significant for Rate5, Rate6, Rate8, and peak frequency; X1 and X2 are statis-

Table 2
Derived process rates during transition and final CLD parameters at the end of steady state for the 27 co-precipitation runs based on real-time FBRM data (CL1: 0–10 μm; CL2: 10–19.953 μm; CL3: 19.953–50.119 μm;CL4: 50.119–100 μm).

Run #	Pattern	During the transition period (Rate: counts/s/s)								At the end of steady state period	
		After the 1st addition of water				After the 2nd addition of water				Mode (μm)	Peak frequency (counts/s)
		Rate1	Rate2	Rate3	Rate4	Rate5	Rate6	Rate7	Rate8		
1	133	10.5736	10.7710	10.5172	1.6980	24.6889	22.4311	21.5892	3.7845	27.123	305.72
2	213	2.2180	3.9841	8.4474	0.8013	1.4502	3.1170	8.1841	2.4616	46.416	91.581
3	312	0.3826	0.4129	0.9374	0.6608	5.2500	2.9591	4.9887	2.3592	54.117	46.169
4	132	2.7907	3.4472	4.2661	1.1092	29.2702	28.5301	26.8955	3.8570	25.119	270.67
5	331	1.6878	1.2320	2.7800	1.3865	2.8786	1.8109	3.3692	1.5301	42.987	60.755
6	123	17.4587	14.5936	12.1450	0.9303	31.2612	38.0034	55.6426	5.0592	31.623	392.19
7	323	1.2708	2.1529	3.9911	2.0179	6.3940	7.0410	11.3126	4.9157	39.811	109.45
8	221	0.8667	1.2557	3.3192	1.2608	0.7273	0.9842	0.4499	1.5104	46.416	92.221
9	313	0.3736	0.6492	0.9462	0.7991	1.8012	2.9355	5.7506	2.0944	54.117	59.715
10	222	4.7627	5.3115	5.7942	0.3398	12.8510	17.5157	35.4949	9.5555	36.869	254.68
11	321	1.0028	1.4100	3.4853	1.6461	1.7586	3.0196	4.7413	2.2973	39.811	87.882
12	212	16.8696	28.8235	59.4750	4.7923	10.3219	10.9985	27.4774	4.0631	36.869	260.91
13	311	0.5023	0.5972	1.4905	0.9732	0.9189	1.2675	2.4570	1.2153	50.119	34.156
14	131	2.4071	3.3990	3.2181	0.5674	9.6399	9.4964	15.6298	3.2423	34.145	205.99
15	211	0.2351	0.3937	0.9423	0.2103	1.0915	1.8768	5.8460	2.7091	50.119	88.607
16	332	4.2665	5.7420	8.6329	2.5536	5.8941	6.4683	11.9191	4.6014	34.145	127.26
17	223	5.4219	6.9866	8.5837	0.4530	10.4084	15.8936	34.5608	7.6965	34.145	252.77
18	112	20.9567	30.1534	62.3363	8.2928	17.1882	18.1251	36.7147	5.7329	31.623	285.74
19	122	3.8570	0.1250	0.1682	0.0324	14.4302	22.9464	35.9995	9.6199	34.145	292.63
20	231	1.0080	0.6780	1.0058	0.7622	9.3001	10.1723	17.7417	5.2822	31.623	155.29
21	232	0.8753	0.5902	0.8450	1.0472	12.5681	14.7266	26.3863	7.0553	31.623	171.35
22	121	4.3244	6.5958	10.1978	1.0433	5.3310	6.1365	7.4282	2.6867	31.523	243.11
23	233	1.5458	1.4770	1.7359	0.4937	11.0354	14.8249	31.6042	7.8285	29.286	203.1
24	322	0.8057	1.4403	2.7305	1.3534	3.8065	4.0625	9.3274	3.4756	39.811	82.84
25	113	17.8998	26.7439	50.1879	5.4916	2.2388	2.7659	6.1132	2.1597	36.869	201.59
26	111	1.2927	2.2658	4.0293	0.5283	3.5820	4.1017	10.2035	3.0273	34.145	131.14
27	333	11.7237	16.2410	26.0520	3.6388	5.0155	4.0808	8.1521	4.4261	36.869	159.24

Table 3
3^3 full factorial design analysis results using ANOVA.

Process period	Dependent	Hypothesis type	Source	Degree of freedom	Mean sum of squares	*F*-value	*P*-value	Statistically significant variables or interactions ($\alpha = 0.05$)
During the transition period after the 1st water addition	Rate1 (counts/s/s)	3	X1	2	110.47	7.07	0.0171	X1, X3
		3	X2	2	18.83	1.20	0.3491	
		3	X1 × X2	4	39.68	2.54	0.1222	
		3	X3	2	92.47	5.91	0.0265	
		3	X1 × X3	4	36.30	2.32	0.1445	
		3	X2 × X3	4	40.59	2.60	0.1169	
			Error	8	15.64			
	Rate2 (counts/s/s)	3	X1	2	137.04	3.28	0.0909	
		3	X2	2	101.69	2.44	0.1491	
		3	X1 × X2	4	99.44	2.38	0.1376	
		3	X3	2	143.94	3.45	0.0831	
		3	X1 × X3	4	51.22	1.23	0.3718	
		3	X2 × X3	4	98.39	2.36	0.1403	
			Error	8	41.73			
	Rate3 (counts/s/s)	3	X1	2	319.39	1.99	0.1988	
		3	X2	2	667.69	4.16	0.0577	
		3	X1 × X2	4	433.02	2.70	0.1083	
		3	X3	2	410.36	2.56	0.1385	
		3	X1 × X3	4	144.15	0.90	0.5079	
		3	X2 × X3	4	415.61	2.59	0.1174	
			Error	8	160.48			
	Rate4 (counts/s/s)	3	X1	2	2.52	1.42	0.2965	
		3	X2	2	5.28	2.97	0.1083	
		3	X1 × X2	4	6.78	3.81	0.0508	
		3	X3	2	4.02	2.26	0.1662	
		3	X1 × X3	4	1.77	0.99	0.4635	
		3	X2 × X3	4	4.68	2.63	0.1135	
			Error	8	1.78			
During the transition period after the 2nd water addition	Rate5 (counts/s/s)	3	X1	2	309.33	14.60	0.0021	X1, X2, X3
		3	X2	2	126.28	5.96	0.0260	
		3	X1 × X2	4	27.09	1.28	0.3545	
		3	X3	2	178.10	8.40	0.0108	
		3	X1 × X3	4	34.02	1.61	0.2633	
		3	X2 × X3	4	50.11	2.36	0.1396	
			Error	8	21.19			
	Rate6 (counts/s/s)	3	X1	2	392.97	15.94	0.0016	X1, X2, X3
		3	X2	2	161.23	6.54	0.0208	
		3	X1 × X2	4	32.78	1.33	0.3382	
		3	X3	2	242.58	9.84	0.0070	
		3	X1 × X3	4	44.51	1.81	0.2211	
		3	X2 × X3	4	58.02	2.35	0.1409	
			Error	8	24.66			
	Rate7 (counts/s/s)	3	X1	2	748.06	10.27	0.0062	X1, X3
		3	X2	2	216.60	2.97	0.1082	
		3	X1 × X2	4	51.73	0.71	0.6075	
		3	X3	2	666.40	9.15	0.0086	
		3	X1 × X3	4	74.00	1.02	0.4540	
		3	X2 × X3	4	205.60	2.82	0.0988	
			Error	8	72.84			
	Rate8 (counts/s/s)	3	X1	2	12.64	7.27	0.0159	X1, X2, X3
		3	X2	2	13.28	7.63	0.0140	
		3	X1 × X2	4	2.94	1.69	0.2440	
		3	X3	2	20.44	11.75	0.0042	
		3	X1 × X3	4	2.33	1.34	0.3355	
		3	X2 × X3	4	4.60	2.64	0.1129	
			Error	8	1.74			
At the end of steady state (end of the co-precipitation process)	Mode (μm)	3	X1	2	309.70	19.14	0.0009	X1, X2
		3	X2	2	289.37	17.88	0.0011	
		3	X1 × X2	4	35.00	2.16	0.1640	
		3	X3	2	38.63	2.39	0.1538	

Table 3 (*Continued*)

Process period	Dependent	Hypothesis type	Source	Degree of freedom	Mean sum of squares	*F*-value	*P*-value	Statistically significant variables or interactions ($\alpha = 0.05$)
		3	X1 × X3	4	9.30	0.57	0.6892	
		3	X2 × X3	4	12.76	0.79	0.5640	
			Error	8	16.18			
	Peak frequency (counts/s)	3	X1	2	67733.69	38.49	0.0001	X1, X2, X3
		3	X2	2	11171.37	6.35	0.0223	
		3	X1 × X2	4	1324.77	0.75	0.5834	
		3	X3	2	17368.57	9.87	0.0069	
		3	X1 × X3	4	2307.88	1.31	0.3438	
		3	X2 × X3	4	3597.69	2.04	0.1807	
			Error	8	1759.65			

tically significant for mode. No statistically significant variables or interactions were found for Rate2, Rate3, and Rate4.

To fully realize the values of DOE and ANOVA methodologies and overcome the inherent limitations of empirical data-driven models, fundamental formulation and process domain knowledge were used to facilitate the interpretations of the statistical data analysis results. Doing so may help to bridge the gaps between the data-driven models, mechanistic models, and first-principle models of the process. As a case study, effort was made throughout this work to link the ANOVA results to particulate process engineering principles.

For the initial addition during transition period, only X1 and X3 stand out as critical process variables to the derived co-precipitation process rate for the smallest CL range of [0, 10] μm which is most likely related to the nucleation process. Since X1 is directly related to the degree of super-saturation (S) as discussed in Section 3.6.2.1, X3 is directly related to the creation of S in the co-precipitation vessel, the ANOVA results for the 1st water addition suggest that the initial nucleation process is thermodynamically controlled. No statistically significant process variables and interactions were found from ANOVA at other CL ranges after the 1st water addition. This suggests that the competition among various process mechanisms (such as thermodynamics for super-saturation creation, nucleation kinetics, and multi-phase mass transfer, etc.) is at such state that none of them becomes dominant.

For the 2nd addition during the transition period, all of three process variables (X1, X2, and X3) stand out as critical variables to Rate5, Rate6, and Rate8. This may suggest all of three process variables can contribute directly to the co-precipitation process rates at CL5, CL6, and CL8. X1 and X3 stand put as critical process variables to Rate7, which may suggest both the thermodynamic factor (X1) and a mass-transfer factor (X3) are playing equally important role.

At the end of steady state of the co-precipitation process, X1 and X2 are critical process variables to the mode of the final CLD. This suggests both the thermodynamic factor (X1) and mass-transfer factor (X2) are most influential to the particle size of the final co-precipitate slurry. X1, X2, and X3 are critical process variables to the peak frequency of the final CLD. This suggests that all of the main process variables directly contribute to the number of counts generated. Therefore, the ANOVA results for the characteristic parameters of the final CLD demonstrated the critical importance of process control (i.e., controlling process parameters as studied in this work) to achieve the desired product quality attributes at the end of the co-precipitation process.

3.6.2. Development of statistical predictive models using general linear modeling

Based on the screening results of statistically significant process variables as highlighted in the last column in Table 3, further statistical analyses were carried out using general linear models (GLM). The general linear models and parameters' estimations are listed in Table 4.

To illustrate the GLM model formulation, let us discuss the linear model for Rate1.

$$(\text{Rate1})_i = \beta_0 + \beta_1(\text{X1})_i + \beta_3(\text{X3})_i + \varepsilon_i, \quad i = 1, \ldots, 27, \tag{3}$$

where $(\text{Rate1})_i$ is the ith observation of Rate1, $(\text{X1})_i$ and $(\text{X3})_i$ are the ith values for X1 and X3, respectively, error term ε_i is normally independently identically distributed with mean 0 and variance σ^2, which is written as $\varepsilon_i \sim iid\ N(0, \sigma^2)$, $i = 1, \ldots, 27$. The GLM modeling was carried out using SAS9.2. The modeling results were summarized in Table 4. From the parameter estimate listed in the 5th and 6th columns, it is easily appreciated that the effects of main process variables on derived process rates, mode and peak frequency of the final CLD have certain trends as discussed below.

3.6.2.1. Slurry temperature (X1). It has a significant effect (at significance level $\alpha = 0.05$) on the derived process Rate1, Rate5, Rate6, Rate7, and Rate8. Furthermore, since parameter β_1 estimate is negative for these rates, the lower slurry temperature will result in greater derived process rates. This could be explained from a thermodynamic point of view. Theoretically, X1 could directly impact the super-saturation within reaction vessel. Since the solubility of co-precipitate is a function of X1, change of X1 will lead to a change of solubility. Consequentially, the degree of super-saturation (S) will be changed accordingly:

$$S = \frac{C - C_s}{C_s} \tag{4}$$

where C stands for the actual concentration of a substance or species in the solution, C_s the saturated concentration determined by its solubility. Thus, when C remains constant, a lower X1 will lead to a smaller C_s (when the temperature coefficient of C_s is positive, as in most cases) or a larger C_s (when the temperature coefficient of C_s is negative, as occasional case), which in turn will lead to a greater thermodynamic driving force $\Delta C = C - C_s$. According to Eq. (1), it will lead to a greater co-precipitation rate, thus more counts/s will be generated.

On the other hand, X1 was a significant factor (at significance level $\alpha = 0.05$) on mode of final CLD at the end of steady state. However, parameter β_1 estimate is positive for mode. Therefore, the lower slurry temperature will result in smaller mode or finer co-precipitate crystal.

3.6.2.2. Slurry stirring rate (X2). It was a significant factor (at significance level $\alpha = 0.05$) for Rate5 and Rate6. The parameter β_2 estimate is positive for Rate5, Rate6, and Rate8. This indicates that a higher X2 will result in a higher Rate5, Rate6, and Rate8. This trend is likely attributed to the possible mixing effect caused by stirring, as discussed below.

Table 4
Development of statistical predictive models using general linear modeling.

Process period	Response variable	Model	Coefficient of correlation, r	Parameter	Estimate	Standard error	t-value	P-value
Transition period after the 1st water addition	Rate1	$(\text{Rate1})_i = \beta_0 + \beta_1(\text{X1})_i + \beta_3(\text{X3})_i + \varepsilon_i$ $\varepsilon_i \sim N(0, \sigma^2)$	0.595	β_0	5.09	1.02	5.01	<0.0001
				β_1	−3.31	1.24	−2.66	0.0138
				β_3	3.06	1.24	2.46	0.0214
Transition period after the 2nd water addition	Rate5	$(\text{Rate5})_i = \beta_0 + \beta_1(\text{X1})_i + \beta_2(\text{X2})_i + \beta_3(\text{X3})_i + \varepsilon_i$ $\varepsilon_i \sim N(0, \sigma^2)$	0.751	β_0	8.93	1.14	7.85	<0.0001
				β_1	−5.77	1.39	−4.15	0.0004
				β_2	3.69	1.39	2.65	0.0143
				β_3	3.28	1.39	2.36	0.0273
	Rate6	$(\text{Rate6})_i = \beta_0 + \beta_1(\text{X1})_i + \beta_2(\text{X2})_i + \beta_3(\text{X3})_i + \varepsilon_i$ $\varepsilon_i \sim N(0, \sigma^2)$	0.748	β_0	10.23	1.29	7.96	<0.0001
				β_1	−6.61	1.57	−4.19	0.0003
				β_2	3.58	1.57	2.27	0.0328
				β_3	4.01	1.57	2.55	0.0180
	Rate7	$(\text{Rate7})_i = \beta_0 + \beta_1(\text{X1})_i + \beta_3(\text{X3})_i + \varepsilon_i$ $\varepsilon_i \sim N(0, \sigma^2)$	0.631	β_0	17.26	2.19	7.87	<0.0001
				β_1	−8.57	2.68	−3.19	0.0039
				β_3	6.39	2.68	2.38	0.0256
	Rate8	$(\text{Rate8})_i = \beta_0 + \beta_1(\text{X1})_i + \beta_2(\text{X2})_i + \beta_3(\text{X3})_i + \varepsilon_i$ $\varepsilon_i \sim N(0, \sigma^2)$	0.511	β_0	4.23	0.42	10.15	<0.0001
				β_1	−0.68	0.51	−1.33	0.1956
				β_2	0.88	0.51	1.72	0.0994
				β_3	0.94	0.51	1.84	0.0786
At the end of steady state period	Mode	$(\text{mode})_i = \beta_0 + \beta_1(\text{X1})_i + \beta_2(\text{X2})_i + \varepsilon_i$ $\varepsilon_i \sim N(0, \sigma^2)$	0.854	β_0	37.83	0.83	45.76	<0.0001
				β_1	5.86	1.01	5.79	<0.0001
				β_2	−5.64	1.01	−5.57	<0.0001
	Peak frequency	$(\text{peak})_i = \beta_0 + \beta_1(\text{X1})_i + \beta_2(\text{X2})_i + \beta_3(\text{X3})_i + \varepsilon_i$ $\varepsilon_i \sim N(0, \sigma^2)$	0.856	β_0	172.84	10.07	17.16	<0.0001
				β_1	−86.74	12.33	−7.03	<0.0001
				β_2	25.54	12.33	2.07	0.0498
				β_3	37.57	12.33	3.05	0.0057

First of all, X2 is directly linked to the fluid hydrodynamics and flow field within the vessel which could impact both the concentration distributions of the drug and polymer and mass-transfer process in the co-precipitation vessel. Secondly, it impacts the magnitude and functionality of the coagulation kernel (Randolph and Larson, 1988) that describes the rate at which particles collide and coalesce, which will be entered into the particle population balance of the co-precipitation process. From a fluid dynamics perspective, local and mean values of the specific power input caused by the rotation speed of the impeller determine the degree of macro-mixing and micro-mixing. The bulk mixing inside the co-precipitation vessel must be ensured that its elements participate equally in the mass transfer process and hence the entire volume is sufficiently utilized. In addition to this bulk mixing, sufficient micro-mixing has to be present in the vicinity of contacting zone where water is introduced and fed to the slurry via the feed inlet, in order to even out quickly any local super-saturation peaks generated due to relatively fast addition of water and relatively slow stirring rate. Thermodynamically, during the initial addition of water, super-saturation is expected to be generated in the vicinity of contacting zone where water is introduced into the solution and is in contact with the solution, as soon as the overall composition point of the 4 component system comes across the meta-stable line. Visual evidence along with the results from online turbidity monitoring (Wu and Khan, 2009) of the co-precipitation process have supported this theoretical analysis. Therefore, a lower X2 may not be strong enough to guarantee an effective mixing at both macro and micro level. On the other hand, a higher X2 could promote micro-mixing in the vicinity of contact zone thus reduce the super-saturation there. Furthermore, a higher X2 can facilitate mass transfer process including both bulk transfer and diffusion across the diffusion layer. Therefore, according to the Nyvlt's model (Nyvlt et al., 1985), the apparent overall derived process rates covering the three steps of multi-phase mass transfer process are faster.

X2 was a significant factor (at significance level $\alpha = 0.05$) for both mode and peak frequency of the final CLD at the end of steady state.

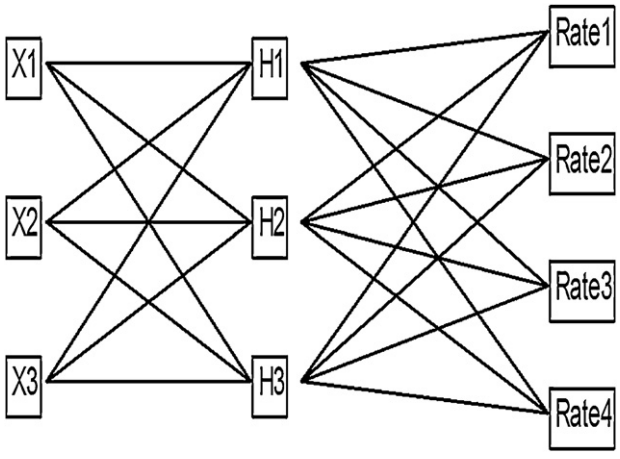

Fig. 9. Neural network modeling framework for Rate1–4.

Table 5
Neural Network models' performance summary.

Process period	Response variable	SSE	RMSE	SSE scaled	RMSE scaled	R^2
After the 1st water addition during the transition period	Rate1 (counts/s/s)	19.7528	0.9267	0.4963	0.1469	0.9809
	Rate2 (counts/s/s)	54.0231	1.5326	0.6703	0.1707	0.9742
	Rate3 (counts/s/s)	151.4385	2.5660	0.4891	0.1458	0.9812
	Rate4 (counts/s/s)	6.2611	0.5217	1.7930	0.2792	0.9310
After the 2nd water addition during the transition period	Rate5 (counts/s/s)	86.5533	1.9399	1.2218	0.2305	0.9530
	Rate6 (counts/s/s)	121.0679	2.2943	1.3498	0.2422	0.9481
	Rate7 (counts/s/s)	587.4973	5.0540	2.9544	0.3584	0.8864
	Rate8 (counts/s/s)	11.6128	0.7106	2.0667	0.2998	0.9205
At the end of steady state period	Mode (μm)	99.5207	2.0801	1.5845	0.2625	0.9391
	Peak frequency (counts/s)	14200.851	24.8481	1.5675	0.2611	0.9397

Where: SSE: residual sum of squares error; MSE: mean squared error, which is the estimate of the variance of the residual error and equals to the SSE divided by the degree of freedom; and RMSE: square root of the MSE and estimates the standard deviation of the residual error.

However, the parameter β_2 estimate is negative for mode but positive for peak frequency. This indicates that higher stirring rate will lead to smaller particles but more counts in its final CLD.

3.6.2.3. Addition rate of water (X3). It has a significant effect (at significance level $\alpha = 0.05$) on Rate1, Rate5, Rate6, and Rate7, with a higher X3 resulting in a greater derived process rate, as indicated by the GLM modeling results (the parameter β_3 estimate is positive). Similar trend was observed for the derived process Rate8 although it did not meet the significance criteria. This trend could be explained by the generation of super-saturation and the re-distribution of various species in the vicinity of contacting zone. Firstly, the addition of water into the slurry system will directly change the concentrations of various species, which in turn will change the degree of super-saturation in the vicinity of contacting zone; secondly, it indirectly impacts the re-distribution of various species within the vessel. The generation rate of super-saturation in the vicinity of contacting zone due to a faster X3 could be larger than the consumption rate of the super-saturation due to micro-mixing caused by the stirring. Therefore, a trend of a higher X3 leading to a greater derived process rate was observed.

Similar trend was observed for peak frequency of the final CLD at the end of steady state of the co-precipitation

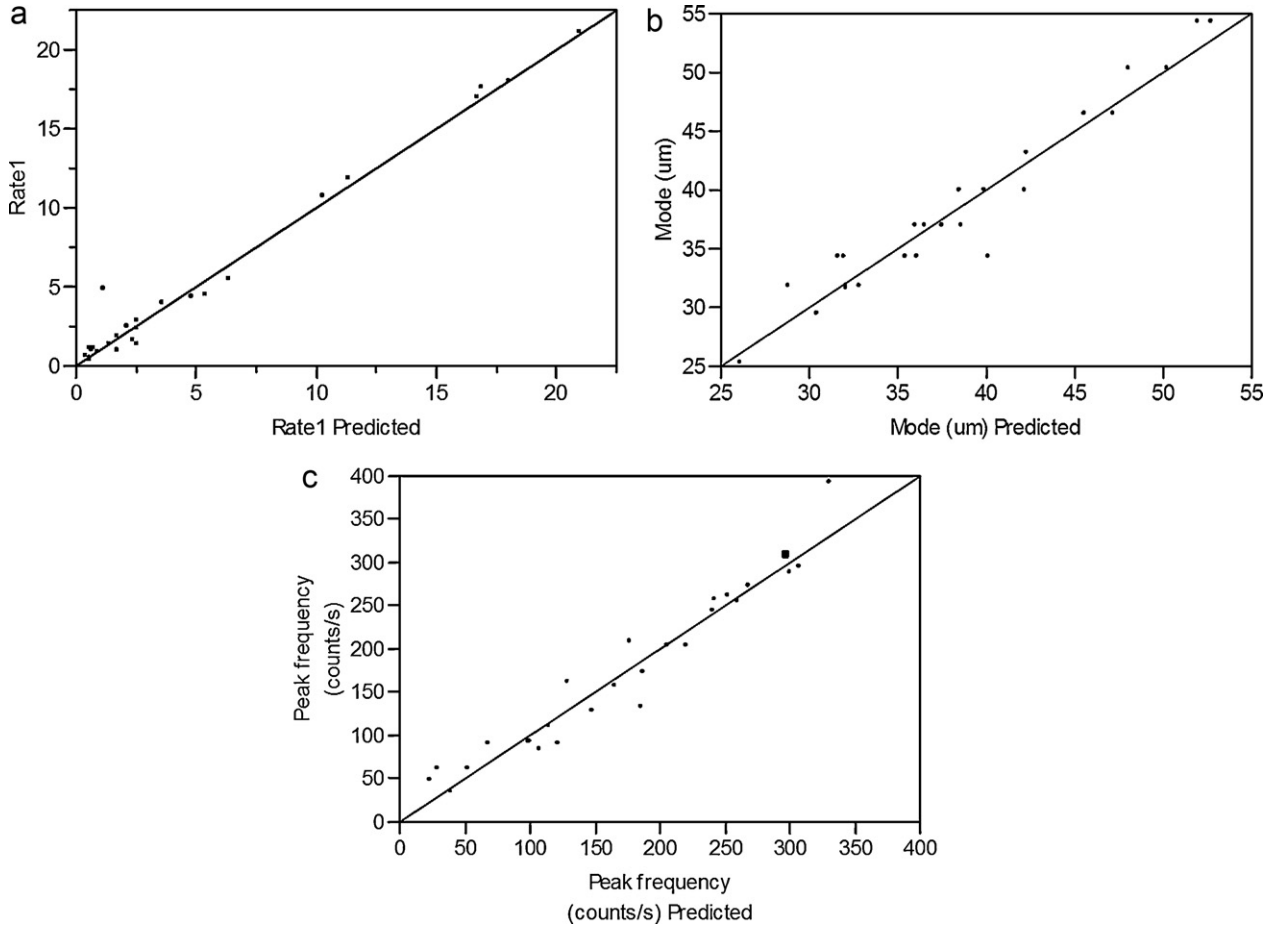

Fig. 10. Plot of actual responsible variable values vs. predicted response variable values via neural network models. (a) Rate1; (b) mode; and (c) peak frequency.

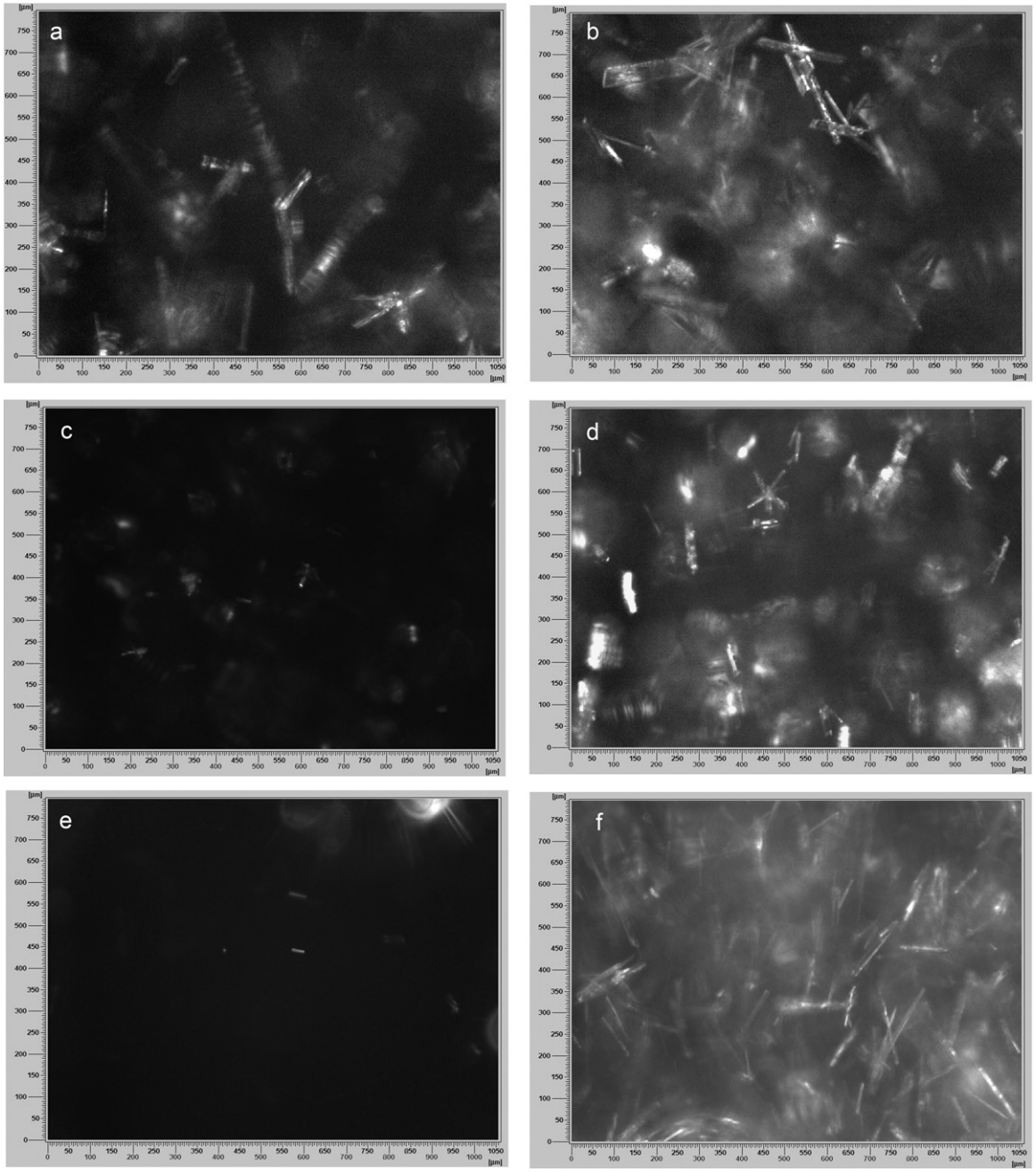

Fig. 11. a–f PVM images right before and right after the 2nd addition of water to the co-precipitation system for various runs. (a) Run 111 (before); (b) run 111 (after); (c) run 113 (before); (d) run 113 (after); (e) run 131 (before); and (f) run 131 (after).

process, although it did not meet the significance criteria.

3.7. Neural network modeling for both transition period and steady state period

Table 4 shows that GLM modeling based on selection results of statistically significant process variables provided a good parameter estimate. The GLS parameter estimate enabled us to find out the trends about the effects of each critical process variable on both the derived process rates during transition period and characteristic quality parameters of CLD at the end of steady state period. It was also able to explain the variability for each response variable to a large extent, with coefficient of correlation r from 0.511 to 0.856. As shown by screening results in Table 3, there are certain two-way interactions existing, although not necessarily statistically significant at significance level $\alpha = 0.05$. Therefore, there is certain nonlinearity for the co-precipitation process. This is expected due to the fact that the dynamic co-precipitation process involves process thermodynamics, kinetics, mass-transfer, and other transfer phenomena, many process factors and material attributes could contribute to the overall process kinetics in a very complicated fashion. The ANOVA and GLM results in this work demonstrated that, while only focusing on statistically significant variables and interactions, it is expected that GLM may generate coefficient of correlation r with some deviation from 1.0 due to the complexity of the co-precipitation process.

Among various multivariate techniques, neural network algorithm has the capability of handling nonlinearity (Despagne and

Massart, 1998) between the independent variables and response variables. Therefore, a neural network (NN) algorithm implemented in JMP 7.0 software was utilized to examine if the modeling's R^2 values can be improved. Furthermore, neural network algorithm was applied to construct the linkage between the three normalized process variables and response variables that include derived process rates at transition stage and mode and peak frequency of CLD at the end of steady state.

As an example, the NN framework was shown in Fig. 9 for Rate1. Our modeling results show that the NN was able to generate quantitative models with decent correlation coefficients, as shown in Table 5. Most of the R^2 values are pretty high (0.88–0.98), Furthermore, the predicted response variable values agree well with the actual response variable values, as demonstrated in Fig. 10(a)–(c) which are examples for derived process Rate1 at transition period, mode and peak frequency of CLD at the end of steady state period, respectively.

In summary, our NN modeling results suggest that quantitative correlations between the process variables and the derived process rate based on Lasentec FBRM CLD vs. time profile do exist. Therefore, it is possible to utilize these quantitative correlations for better process design and process control.

3.8. PVM images evidenced nucleation and crystal growth during the dynamic course of adding water

The profile of FBRM counts/s vs. time stands for the real-time chord length statistics for slurry population under scrutiny of the FBRM probe. The PVM images provide direct evidences on morphology and size of individual particles under scrutiny of the PVM probe at corresponding time point. Therefore, statistically speaking, the PVM data should reveal some sample information of the slurry population under investigation. A recent report (Greaves et al., 2008) had shown that the PVM can be used as a direct visual method to test the reliability of the FBRM results. In addition, data acquired via those two techniques should be complementary to each other. Therefore, it would be interesting to examine the linkage between the FBRM data and PVM images.

In this work, 6 co-precipitation runs were selected for *in-situ* real-time monitoring using PVM during the entire course of co-precipitation process. First of all, it demonstrated that for most cases, no appreciable sizes of particles were detected via PVM before and immediate after the 1st addition of water to the ternary system. However, about 10 min after the 1st addition of water (the time interval between each addition was *ca.* 10 min), right before the 2nd addition of water, co-precipitates with appreciable sizes were detected via PVM. This is qualitatively consistent with what FBRM data revealed that relative fewer counts of fine particles were detected after the 1st addition of water. Secondly, significant increases were observed via PVM for both the number and dimensions of the co-precipitates before and right after the 2nd addition of water, as shown in Fig. 11(a)–(f). Again, this is qualitatively consistent with what FBRM indicated that significant counts at various CL ranges were detected after the 2nd addition of water. In addition to this enlargement phenomenon which was expected, the following interesting facts were also observed: (1) for run with stirring rate of 400 rpm, the co-precipitates have needle-like morphology; (2) for runs with stirring rate of 50 rpm, the co-precipitates have both needle-like and rectangular prism morphologies. Some co-precipitates are attached to each other. The rectangular prisms usually are shorter than those needle-like co-precipitates in length.

During each addition of water, size changes of the co-precipitate on both length and width dimensions are appreciable via PVM images. As a first approximation, estimation was made based on the surrogated data from the limited PVM images, as shown in Fig. 12. It demonstrated the trends of the dimension changes of co-precipitates at both length and width dimensions during the course of a co-precipitation process.

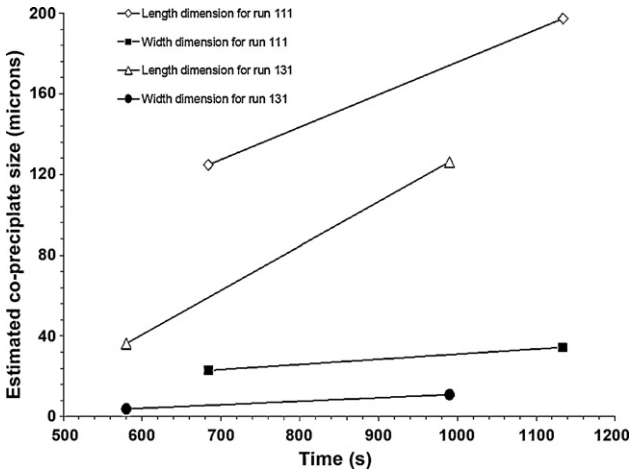

Fig. 12. Estimated average co-precipitate sizes for before and after the 2nd addition of water to the co-precipitation system based on PVM images.

3.9. Discussion from a regulatory science perspective

ICH Q8(R2) (FDA/ICH, 2009) defines the design space as "The multidimensional combination and interaction of input variables (e.g., material attributes) and process parameters that have been demonstrated to provide assurance of quality. Working within the design space is not considered as a change. Movement out of the design space is considered to be a change and would normally initiate a regulatory post-approval change process." In addition to identifying the main process variable effects, multi-factorial DOE study as conducted in this work permits the quantification of variable interactions. The information on the presence or absence of interactions might be valuable for process optimization, identifying cause of variability, or informing a failure modes and effects analysis (FMEA) for defining a design space. "A design space can be developed at any scale. The applicant should justify the relevance of a design space developed at small or pilot scale to the proposed production scale manufacturing process and discuss the potential risks in the scale-up operation. If the applicant proposes the design space to be applicable to multiple operational scales, the design space should be described in terms of relevant scale-independent parameters." (FDA/ICH, 2009). As a proof-of-concept, this simplified case study was mainly focused on the effects of process variables while keeping other variables fixed based on risk analysis and initial formulation development results. It is important to keep in mind that other variables such as formulation variables can certainly play vital roles during process characterization and design space development. Based on the GLM modeling results, both 2D and 3D contour plots were made for certain response variables of the dynamic co-precipitation process, such as Rate7 for the transition period and mode obtained at the end of steady state period, as shown in Figs. 13(a), (b) and 14(a), (b), respectively. Fig. 13(a) and (b) demonstrated that maximum Rate7 was obtained around the following process window: $-1 < X1 < 0.1$, $-0.25 < X3 < 0.95$. Fig. 14(a) and (b) demonstrated that maximum mode was obtained around the following process windows: $0.5 < X1 < 1$, $-1 < X2 < -0.5$.

As demonstrated in this and previous works, process knowledge and in-depth process understanding are essential to establish appropriate process control strategy and operational space, such that co-precipitation product with desired quality attributes and properties (such as PSD, morphology, etc.) could be achieved eventually. Therefore, they are key elements for developing appropriate

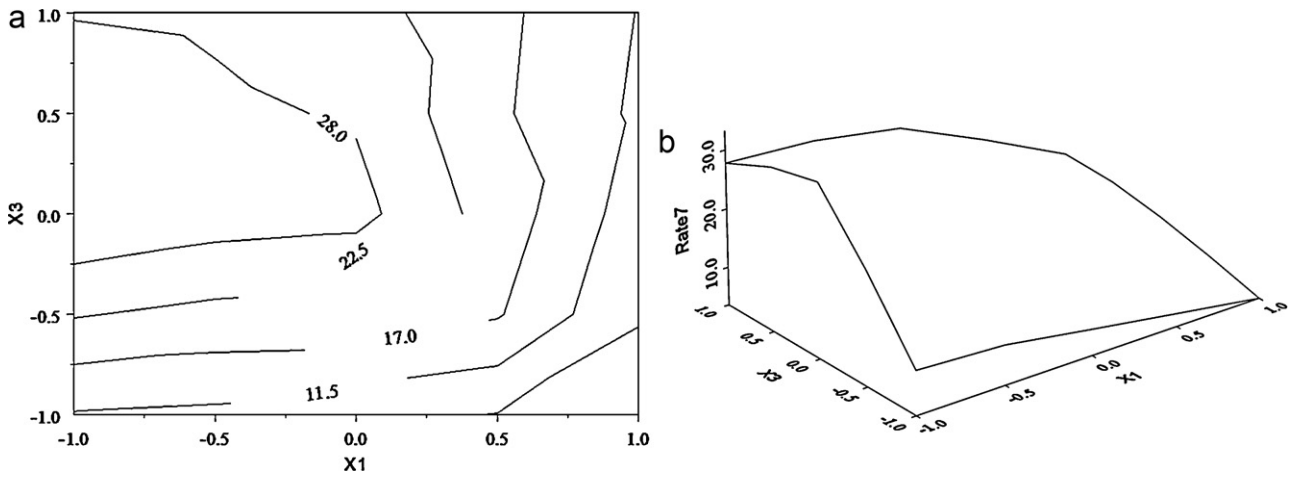

Fig. 13. Contour plots for Rate7 at the transition period as a function of critical variables X1 and X3, based on GLM modeling results. (a) 2D contour plot; and (b) 3D contour plot.

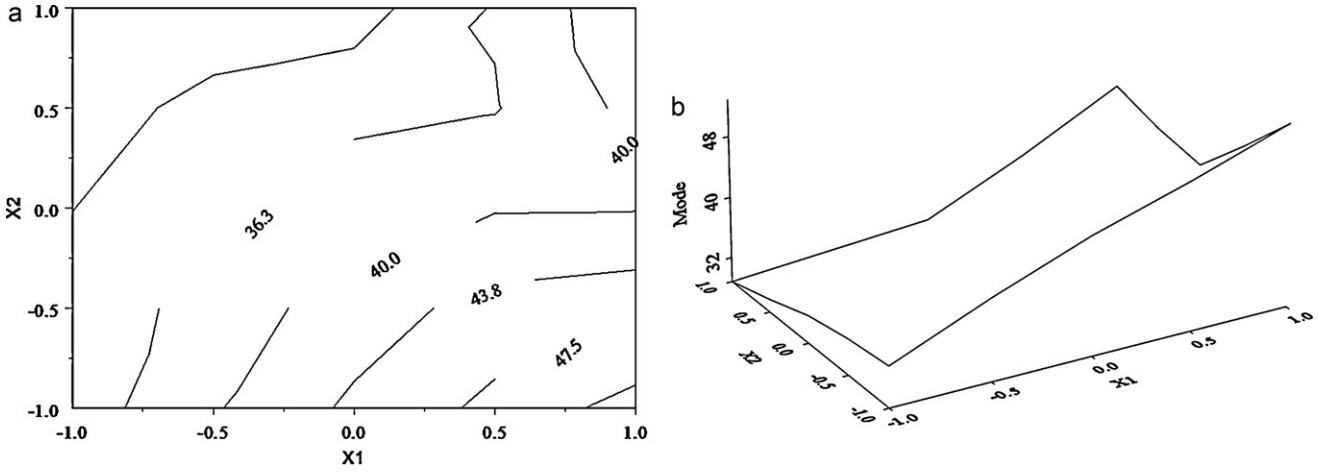

Fig. 14. Contour plots for mode at the end of steady state as a function of critical variables X1 and X2, based on GLM modeling results. (a) 2D contour plot; and (b) 3D contour plot.

process design space. Risk analysis, identification of critical formulation/process variables and interaction, understanding the effects of these critical variables and interactions on the key product quality attributes are essential steps to achieve enhanced process understanding. Validating and challenging design space are important steps to ensure a robust design space developed for a pharmaceutical manufacturing process.

4. Conclusions

In this work, a dynamic pharmaceutical co-precipitation process of gradually introducing water (non-solvent) to the ternary system of naproxen (drug)–Eudragit L100 (polymer)–alcohol (solvent) was monitored in real-time and *in situ* by using Lasentec FBRM and PVM technology. Based on the Fishbone analysis result, the impact of high risk process variables such as co-precipitation slurry temperature, slurry stirring rate, and non-solvent addition rate on both the FBRM counts/s vs. time profile and CLD at the end of co-precipitation process were examined systematically, using a 3^3 full factorial design. A 3D count-time-size plot was constructed to map the entire co-precipitation process and distinguish various process stages: incubation, transition, and steady state. A new method was developed to extract a derived process rate from the Lasentec FBRM counts/s vs. time profile by calculating the maximum slope of the FBRM counts/s vs. time after each addition of water. Critical process variables were identified via ANOVA for both transition and steady state periods. GLM were used for parameter estimate of each critical variable. The GLM results demonstrated clear trends about the effects of each critical variable, which can be interpreted using principles of fundamental process phenomena and Nyvlt's diffusion layer model. Furthermore, contour plots demonstrated the process windows bordered by ranges of critical process variables where maximum Rate7 and maximum mode can be obtained at the transition period and at the end of steady state, respectively. Neural network models were able to link process variables with response variables both at transition period and steady state with R^2 of 0.90–0.98. PVM images evidenced nucleation and crystal growth during the dynamic co-precipitation process.

In summary, this work demonstrated the critical importance and technical feasibility of obtaining in-depth process understanding using real-time PAT process monitoring, risk analysis, DOE, multivariate statistical modeling, and fundamental process knowledge during the course of process characterization and process design space development under the ICH Q8(R2) framework.

Acknowledgements

The following persons and organizations are acknowledged: Dr. Meiyu Shen in CDER of FDA for strong statistical supports on DOE analysis and statistical modeling; Dr. Vincent Vilker at the Office

of Testing and Research, CDER, FDA for insightful internal review; Dr. Robbe Lyon at the Division of Product Quality Research, Office of Testing and Research, CDER, FDA (currently retired) for mentorship to H. Wu in FDA; Mr. Eric Dycus at the Mettler Toledo AutoChem Inc. (Columbia, MD) for technical support to our Lasentec systems; Dr. San Kiang at the Bristol-Myers Squibb Company (New Brunswick, NJ), Mr. Steve Ware and Mr. Brendan Simon at the Chemglass Inc. (Vineland, NJ) for their technical discussion and assistance to our crystallization system; Dr. Richard D. Braatz at MIT (Boston, MA) for technical discussion on topics of crystallization; Dr. Ajaz Hussain at Phillips Morris (Washington, DC) for inspiration in the area of PAT as well as encouragement and supports to our crystallization research during his tenure in FDA; and The FDA BioSciences Library for literature supports. The internship for Maury White from CDER of FDA is acknowledged.

References

Am Ende, D.J., Rose, P.R., 2006. Strategies to achieve particle size of active pharmaceutical ingredients. In: Ahmed, F., Abdel-Magid, Stéphane Caron (Eds.), Fundamentals of Early Clinical Drug Development: From Synthesis Design to Formulation. John Wiley & Sons, Inc., pp. 247–267.

Barrett, P., Glennon, B., 1999. In-line FBRM monitoring of particle size in dilute agitated suspensions. Part. Syst. Charact. 16, 207–211.

Chew, J.W., Chow, P.S., Tan, R.B.H., 2007. Automated in-line technique using FBRM to achieve consistent product quality in cooling crystallization. Cryst. Growth Des. 7, 1416–1422.

Despagne, F., Massart, D.L., 1998. Neural networks in multivariate calibration. Analyst 123, 157R–178R.

FDA, 2004. "Guidance for Industry. PAT—A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance". Available at: http://www.fda.gov/cder/guidance/6419fnl.pdf.

FDA/ICH, 2006a. Guidance for Industry. Q8 Pharmaceutical Development. Available at: http://www.fda.gov/cder/guidance/6746fnl.pdf.

FDA/ICH, 2006b. "Guidance for Industry. Q9 Quality Risk Management". Available at: http://www.fda.gov/cder/guidance/7153fnl.pdf.

FDA/ICH, 2007. Q10. Pharmaceutical Quality System. Draft Consensus Guideline. Available at: http://www.fda.gov/cder/guidance/7891dft.pdf.

FDA/ICH, 2009. Guidance for Industry, Q8 (R2) Pharmaceutical Development. Draft available at: http://www.fda.gov/downloads/Drugs/GuidanceCompliance-RegulatoryInformation/Guidances/ucm073507.pdf.

Greaves, D., Boxall, J., Mulligan, J., Montesi, A., Creek, J., Sloan, E.D., Koh, C.A., 2008. Measuring the particle size of a known distribution using the focused beam reflectance measurement technique. Chem. Eng. Sci. 63, 5410–5419.

Howard, K.S., Nagy, Z.K., Saha, B., Robertson, A.L., Steele, G., 2009. Combined PAT-solid state analytical approach for the detection and study of sodium benzoate hydrate. Org. Process Res. Dev. 13, 590–597.

Huang, J., Kaul, G., Cai, C., Chatlapalli, R., Hernandez-Abad, P., Ghosh, K., Nagi, A., 2009. Quality by design case study: an integrated multivariate approach to drug product and process development. Int. J. Pharm. 382, 23–32.

Hukkanen, E.J., Brattze, R.D., 2003. Measurement of particle size distribution in suspension polymerization using in situ laser backscattering. Sens. Actuators, B 96, 451–459.

Hwang, R.C., Geemoules, M.K., Ramlose, D.S., Thomason, C.E., 1998. A systematic formulation optimization process for a generic pharmaceutical tablet. Pharm. Technol. 22, 48–64.

Johnson, R.A., Wichern, D.W., 2007. Applied Multivariate Statistical Analysis, 6th edition. Prentice Hall.

Khan, M.A., Bolton, S., Kislalioglu, M.S., 1994. Optimization of process variables for the preparation of ibuprofen coprecipitates with Eudragit S100. Int. J. Pharm. 102, 185–192.

Kougoulos, E., Jones, A.G., Wood-Kaczmar, M.W., 2005. Modeling particle disruption of an organic fine chemical compound using Lasentec focused beam reflectance monitoring (FBRM) in agitated suspensions. Powder Technol. 155, 153–158.

Langston, P.A., Jones, T.F., 2001. Non-spherical 2-dimensional particle size analysis from chord measurements using Bayes' theorem. Part. Part. Syst. Char. 18, 12–21.

Li, M., Wilkinson, D., 2005. Determination of non-spherical particle size distribution from chord length measurements. Part 1: Theoretical analysis. Chem. Eng. Sci. 60, 3251–3265.

Lindenberg, C., Mazzotti, M., 2009. Effect of temperature on the nucleation kinetics of α L-glutamic acid. J. Cryst. Growth 311, 1178–1184.

Ma, Y., Zhu, J., Chen, K., Wu, Y., Chen, A., 2009. Study on growth kinetics of phosphoric acid hemihydrate using FBRM. J. Cryst. Growth 312, 109–113.

Monnier, O., Klein, J.-P., Hoff, C., Ratsimba, B., 1996. Particle size determination by laser reflection: methodology and problems. Part. Part. Syst. Char. 13, 10–17.

Montgomery, D.C., 2000. Design and Analysis of Experiments, 5th edition. John Wiley and Sons, Chichester.

Naelapaa, K., Alleso, M., Kristensen, H.G., Bro, R., Rantanen, J., Bertelsen, P., 2008. Increasing process understanding by analyzing complex interactions in experimental data. J. Pharm. Sci. 98, 1852–1861.

Nyvlt, J., Sohnel, O., Matuchova, M., Broul, M., 1985. The Kinetics of Industrial Crystallization. Elsevier, New York.

Olivieri, A.C., Faber, N.M., Ferré, J., Boqué, R., Kalivas, J.H., Mark, H., 2006. Uncertainty estimation and figures of merit for multivariate calibration. Pure Appl. Chem. 78, 633–661.

Randolph, A.D., Larson, M.A., 1988. Theory of Particulate Processes: Analysis and Techniques of Continuous Crystallization. Academic Press, Inc. Harcourt Brace Jovanovich, Publishers, New York, p.71.

Rathore, A.S., Winkle, H., 2009. Quality by design for biopharmaceuticals. Nat. Biotechnol. 27, 26–34.

Richmond, W.R., Jones, R.L., Fawell, P.D., 1998. The relationship between particle aggregation and rheology in mixed silica-titania suspensions. Chem. Eng. J. 71, 67–75.

Ruf, A., Worlischek, J., Mazzotti, M., 2000. Modeling and experimental analysis of PSD measurements through FBRM. Part. Part. Syst. Char. 17, 167–179.

Simmons, M.J.H., Langston, P.A., Burbidge, A.S., 1999. Particle and droplet size analysis from chord distributions. Powder Technol. 102, 75–83.

Smith, J., Moore, C., Nasr, M., 2009. Quality trio takes final shape. Pharm. Technol. 33, 66.

Tavare, N.S., 1987. Batch crystallizations: a review. Chem. Eng. Commun. 61, 259–318.

Togkalidou, T., Tung, H.-H., Sun, Y., Andrews, A.T., Braatz, R.D., 2004. Parameter estimation and optimization of a loosely bound aggregating pharmaceutical crystallization using in situ infrared and laser backscattering measurements. Ind. Eng. Chem. Res. 43, 6168–6181.

Woodcock, J., 2004. The concept of pharmaceutical quality. Am. Pharm. Rev. 7, 10–15.

Wu, H., Hussain, A.S., 2005a. Use of PAT for active pharmaceutical ingredients crystallization process control. In: Horacek, P., Simandl, M., Zitek, P. (Eds.), Conference Proceedings of the 16th International Federation of Automatic Control (IFAC) World Congress (CD-ROM). Prague, Czech Republic, July 4–9, 2005.

Wu, H., Hussain, A.S., 2005b. Integration of multivariate statistics and design of experiments to identify critical process variables for pharmaceutical process analytical technology (PAT) applications. In: American Statistical Association 2005 Proceedings of Joint Statistical Meetings , Minneapolis, Minnesota, August 7–11, 2005. Mira Digital Publishing, CD-ROM.

Wu, H., Hussain, A.S., Khan, M.A., 2007a. Process control perspective for process analytical technology: integration of chemical engineering practice into semiconductor and pharmaceutical industries. Chem. Eng. Commun. 194, 760–779.

Wu, H., Heilweil, E.J., Hussain, A.S., Khan, M.A., 2008. Process analytical technology (PAT): quantification approaches in terahertz spectroscopy for pharmaceutical application. J. Pharm. Sci. 97, 970–984.

Wu, H., Khan, M.A., 2009. Quality-by-design (QbD): an integrated approach for evaluation of powder blending process kinetics and determination of powder blending end-point. J. Pharm. Sci. 98, 2784–2798.

Wu, H., Tawakkul, M., White, M., Khan, M.A., 2009. Quality-by-design (QbD): an integrated multivariate approach for the component quantification in powder blends. Int. J. Pharm. 372, 39–48.

Wu, H., Khan, M.A., 2010a. Quality-by-design (QbD): an integrated process analytical technology (PAT) approach for real-time monitoring and mapping the state of a pharmaceutical co-precipitation process. J. Pharm. Sci. 99, 1516–1534.

Wu, H., Khan, M.A., 2010b. Quality-by-design (QbD): an integrated process analytical technology (PAT) approach to determine the nucleation and growth mechanisms during a dynamic pharmaceutical co-precipitation process. J. Pharm. Sci., doi:10.1002/jps22430.

Xie, L., Wu, H., Shen, M., Augsburger, L., Lyon, R.C., Khan, M.A., Hussain, A.S., Hoag, S.W., 2008. Quality-by-design (QbD): effects of testing parameters and formulation variables on the segregation tendency of pharmaceutical powder measured by the ASTM D 6940-04 segregation tester. J. Pharm. Sci. 97, 4485–4497.

Yu, L., Lionberger, R.A., Raw, A.S., D'Costa, R., Wu, H., Hussain, A.S., 2004. Applications of process analytical technology to crystallization processes. Adv. Drug Deliv. Rev. 56, 349–369.

Yu, W., Erickson, E., 2008. Chord length characterization using focused beam reflectance measurement probe—methodologies and pitfalls. Powder Technol. 185, 24–30.

Yu, L.X., 2008. Pharmaceutical quality by design: product and process development, understanding, and control. Pharm. Res. 25, 781–791.

Zaghloul, A.A., Faltinek, J., Vaithiyalingam, S.R., Reddy, I.K., Khan, M.A., 2001. Naproxen–Eudragit microspheres: screening of process and formulation variables for the preparation of extended release tablets. Pharmazie 56, 321–324.