2017

Quality by Design: Process Trajectory Development for a Dynamic Pharmaceutical Coprecipitation Process Based on an Integrated Real-Time Process Monitoring Strategy

Huiquan Wu
Food and Drug Administration

Mansoor A. Khan
Texas A&M University Health Science Center

9

Quality by Design: Process Trajectory Development for a Dynamic Pharmaceutical Coprecipitation Process Based on an Integrated Real-Time Process Monitoring Strategy

Huiquan Wu* and Mansoor A. Khan**

Division of Product Quality Research (DPQR, HFD-940), OTR, Office of Pharmaceutical Quality, Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, MD, USA
*Current affiliation: Process Assessment Branch II, Division of Process Assessment 1, Office of Process and Facilities, Office of Pharmaceutical Quality, Center for Drug Evaluation and Research, FDA, Silver Spring, MD, USA
**Current affiliation: Rangel College of Pharmacy, Texas A&M University Health Science Center, College Station, TX, USA

9.1　Introduction

Increasing prevalence of poorly water-soluble drugs in pharmaceutical development provides notable risks of new products demonstrating low and erratic bioavailability. This dissolution-limited bioavailability may have consequences for safety and efficacy, particularly for drugs delivered by the oral route of administration. Several novel drug delivery technologies have been developed to improve drug solubility, dissolution rates, and bioavailability. Among those are solid dispersion, nanotechnology, supercritical fluid technology, lipid-based technology, and crystal engineering. Although these strategies are available for enhancing the bioavailability of drugs with low aqueous solubility, the success of these approaches is not yet guaranteed and is greatly dependent on the physical and chemical nature of the molecules being developed. On the other hand, crystal engineering [1] offers a number of routes such as cocrystallization [2, 3] and coprecipitation [4–6] to improve solubility and dissolution rate. Coprecipitation of poorly soluble drugs with polymers, an important technique for improving the dissolution and absorption of drugs, has been modified in recent years to prepare extended-release preparations. Previous work on coprecipitation was largely focused on formulation development and product characterization [4–6], for example, optimization of process variables for the preparation of ibuprofen coprecipitates with Eudragit S100, screening of process and formulation variables for

the preparation of extended-release naproxen tablets with Eudragit L100-55, and preparation and characterization of coprecipitate of ibuprofen using different acrylate polymers. It was not until recently that the process analytical technology (PAT) approach [7–11] was explored to gain insights about the coprecipitation process and process monitoring.

PAT [7] offered unprecedented opportunities for the pharmaceutical community to take advantage of the availabilities of modern process analyzers and wealth of experience from other industry sectors. Successful implementation of PAT can enable real-time process data acquisition and extract process information and knowledge from real-time data for better process understanding and better process control. One of the frequently used technologies for pharmaceutical PAT applications is near-infrared (NIR) spectroscopy (NIRS). NIRS is capable of monitoring many pharmaceutical unit operations [9, 10, 12], in addition to its well-established pharmaceutical applications in quality control (QC) and quality assurance (QA) areas [13]. It is a well-known fact that NIR spectrum of pharmaceutical material or dosage form is information rich and it embraces both the physical and chemical information of pharmaceutical material or dosage form [10].

One of the technical challenges for implementing PAT in the pharmaceutical sector is how to handle the dense flow of process data, including spectra acquired by real-time online/in-line/at-line process analyzers and process data acquired by other process measurement techniques. To extract meaningful information from real-time process data for enhanced process understanding and ultimately process control, one must use chemometrics and process modeling techniques. Depending on specific systems studied and real-time process monitoring techniques implemented, various data analysis and modeling techniques can be used. Multivariate statistical modeling techniques such as principal component analysis (PCA) [10, 14, 15], principal component regression (PCR), partial least squares (PLS) [14–16], and artificial neural network (ANN) [17] are often used to develop a predictive model for depleting raw material components and/or increasing product components, multivariate process trajectory, and process map for robust production with optimized process conditions, productivity, and yield.

Naproxen is a potent nonsteroidal anti-inflammatory drug used in the treatment of rheumatoid arthritis, osteoarthritis, and acute gout and as analgesic and antipyretic. However, its use is frequently limited due to significant gastrointestinal side effects. In this study, naproxen was selected as a model drug as it has a short half-life, has gastrointestinal effects, is soluble in alcohol, and is practically insoluble in water. The literature data for the experimental and predicted solubility of naproxen in water are 15.9 and 51.0 mg/l, respectively (http://www.drugbank.ca/drugs/DB00788, accessed on April 16, 2010). Eudragit L100 is the commercial name of an enteric polymer of methacrylic acid–methylmethacrylate that belongs to a class of

reversible soluble/insoluble polymers. Eudragit L100 was used as it yields coprecipitates with naproxen without any need of additives. The dynamic coprecipitation process designed in this work was proceeded via gradually introducing water to the ternary system of naproxen–Eudragit L100–alcohol at controlled temperature of 25°C. When water is introduced into the initially transparent ternary solution sequentially, the overall composition point of naproxen–Eudragit L100–alcohol–water system will be moved accordingly within the phase diagram. Due to significant difference of solubility in alcohol and water for both naproxen and Eudragit L100, the naproxen–Eudragit L100 coprecipitates out from the solution when water is introduced. As a result of the coprecipitation phenomena, the initial transparent solution of the four-component system will become cloudy at the onset of nucleation. In this work, an integrated PAT real-time monitoring strategy was developed to follow the dynamics of the coprecipitation process. As discussed previously, the real-time in-line data acquired can help to illustrate the progress of a multiphase process and to elaborate the sequential events that take place during the process [9]. In this work, multivariate process trajectory was constructed based on the real-time process monitoring data sets, which would be helpful for achieving rational particulate process design and ultimately particulate process control.

9.2 Experimental

9.2.1 Materials

Naproxen USP was obtained from Albemarle Corporation (Orangeburg, SC). It is an odorless, white to off-white crystalline substance. Eudragit L100 was obtained from Röhm America Inc. (Somerset, NJ). It is a white powder with a faint characteristic odor. Solvent reagent alcohol (HPLC grade) was purchased from Fisher Scientific. Nonsolvent DI water was obtained from a FDA in-house facility and was kept in refrigerator at 4°C prior to use. All of these chemicals and solvents were used without any further processing or purification prior to the use for this experimental work.

9.2.2 Equipment and Instruments

Process NIR spectra were acquired with a Luminar™ acousto-optic tunable filter (AOTF)-based NIR process spectrometer (Brimrose Corporation of America, Baltimore, MD), equipped with a transflectance probe. This AOTF NIR process spectrometer has the capability of in-line monitoring both the changes in the concentration of the key components and phase changes in the coprecipitation system or subsystems. The acquisition parameters of Brimrose NIRS being used in experiment included the following: the number of spectra

average was 50; no background correction was applied; normal scan type; and the gain was 2. One spectrum was recorded and saved automatically every 6 s. A 1.0 cm probe extension was attached to the 679B dark-field diffuse reflectance/transmission probe for the actual process monitoring.

Online turbidity measurements were made using a Model 2100AN Laboratory Turbidimeter (Hach Company, Loveland, Colorado) with a Masterflex® L/S® model 7518-10 digital fluid pump (Cole Parmer Instrument Company, Chicago, IL). A flexible tubing connection between the coprecipitation vessel and the sample flow cell of the turbidimeter was made, such that the fluid could be pumped from the coprecipitation vessel to the sample flow cell of the turbidimeter for online turbidity measurement and then be returned to the coprecipitation vessel in the reverse direction. Prior to its use for this work, the 2100AN turbidimeter was calibrated using formazin stock solution as recommended by the supplier. Data were reported using nephelometric turbidity units (NTU), which are specifically compliant for a 90° measurement technique.

Pure components of naproxen and Eudragit L100 were scanned by FOSS NIRsystems (Foss NIRsystems, Silver Spring, MD) offline. The schematic of the experimental setup is shown in Figure 9.1. As shown in Figure 9.1, among many important formulation and process variables, only the drug/polymer ratio was changed during the coprecipitation experiment of this study, while
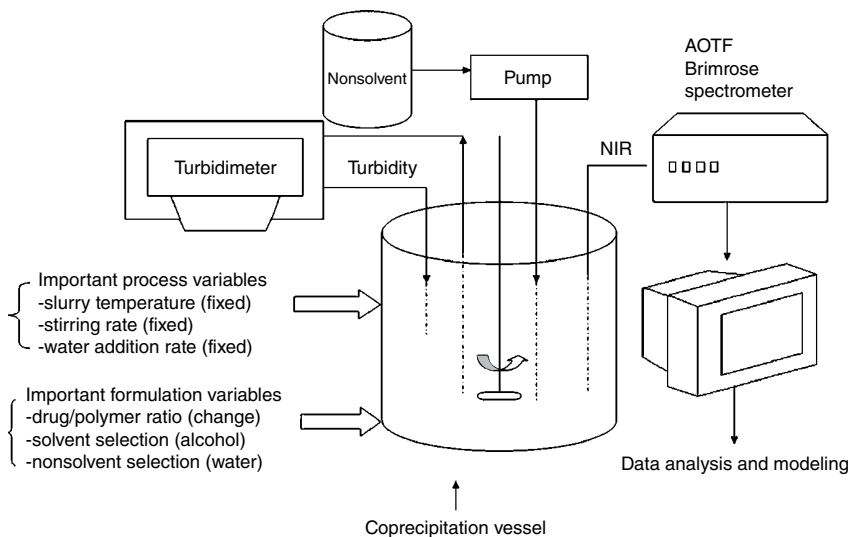

Figure 9.1 Schematic diagram of experimental setup and process flow with an integrated PAT monitoring system for API (naproxen)/polymer coprecipitation process using in-line process NIR and online turbidity measurement simultaneously.

others were fixed. For terminologies about in-line/at-line/online definitions can be found in the FDA PAT guidance [7].

9.3   Data Analysis Methods

9.3.1   PCA and Process Trajectory

PCA is a very useful tool as it can reduce the dimensionality of a data set consisting of a large number of interrelated variables while retaining as much as possible the variation present in the data set. PCA has been extensively applied in almost every discipline, such as chemistry, biology, engineering, meteorology, and pharmaceutical. It can be used in process monitoring, QC, data visualization, batch trajectories, and other areas. For pharmaceutical QbD and PAT applications where process analyzers are installed, frequently a huge amount of interrelated process data has to be handled, in a way that critical process information and knowledge can be extracted for both process and QC purposes. Therefore, PCA is a powerful chemometric technique for such applications. With the help of PCA, the following aspects could be examined: in what aspect one sample is different from another, which variables contribute most to this difference, and whether those variables contribute in the same way (i.e., are correlated) or independently of each other. It can also help to detect patterns and to quantify the amount of useful information, as opposed to noise or meaningless variation, contained in the data set. In this work, plots of a PC versus another (e.g., PC1 vs. PC2), called score plot, are used to depict the process trajectory. Process trajectory method has been used for batch process supervision [18], process monitoring, and diagnosis [19] in chemical and biotech sectors. However, relatively few applications have been reported for pharmaceuticals, especially for small molecule drug manufacturing process monitoring and process control [20]. In this work, all of the NIR spectra were acquired for the wavelength range of [1100, 2300] nm. PCA was conducted on raw NIR spectra without any preprocessing in all cases.

9.3.2   Singular Points of a Signal

It has been noticed that for a time-varying signal in a dynamic system, the information content is not homogeneously distributed throughout [21]. Some landmarks such as extreme values and shape changes in the data, termed as singular points (SPs), in the process trajectory contain more information about the dynamic behavior than others. Mathematically, a singularity is in general a point at which a given mathematical object is not defined or a point of an exceptional set where it fails to be well behaved in some particular way, such as differentiability. Therefore, for a differential equation, an SP is a point that is a singularity for at least one of the known functions appearing in the equation. Geometrically,

an SP is a point (ii) on a curve at which the curve possesses no smoothly turning tangent, or crosses or touches itself, or has a cusp or isolated point, or (ii) on a surface whose coordinates, x, y, and z, depend on the parameters u and v, at which the Jacobians $D(x, y)/D(u, v)$, $D(y, z)/D(u, v)$, and $D(z, x)/D(u, v)$ all vanish. The property of singularity and SP could be very useful for many important engineering applications such as rational process design and process control. For example, SPs are used to segment the process signal into regions with homogeneous properties. Because SPs have physical meaning such as beginning or ending of a process event, they can be directly used for state identification, process monitoring, and process supervision. Examples of SPs include points of discontinuities, trend changes, and extrema. SPs were used to detect phase shifts during rifamycin B fermentation experiments [22] and to characterize microscopic flows [23]. In our previous works, SPs were used to determine powder blending process end point [15] and phase change during a coprecipitation process [9]. In this work, SPs are used to detect the onset of nucleation and crystal growth and to establish the process transition window.

In the area of multivariate data analysis and modeling under the pharmaceutical QbD/PAT framework, it is important to recognize the difference between SP and outlier. In multivariate statistics, an outlier is an observation that is numerically distant from the rest of the data. Outliers can occur by chance in any distribution, but they are often indicative either of measurement error or that the population has a heavy-tailed distribution. Outliers may be indicative of data points that belong to a different population than the rest of sample set. The commonly used method for identifying outliers in multivariate analysis is based on the squared Mahalanobis distance (MD) [24]. Points for which MD^2 value is large are identified as atypical or outliers and evaluated using the χ^2 distribution with the appropriate degrees of freedom. SP, as discussed earlier, is in general a point at which a given mathematical object is not defined or a point of an exceptional set where it fails to be well behaved in some particular way, such as differentiability. Therefore, mathematically outlier and SP are two totally different concepts; physically, they have different meanings that are related to different properties.

9.4 Results and Discussion

To investigate the utility of process trajectory for pharmaceutical process monitoring, a step-by-step approach was taken. First, we examined simple cases of binary systems for which either two dry powder components or two liquid components are involved. The NIRS was used for process assessment and process monitoring of the binary systems. Then, we examined a complicated system of four-component coprecipitation process (both solid and liquid phases are involved) for which NIRS was used for real-time in-line process monitoring.

9.4.1 Using Offline NIR Measurement to Characterize the Naproxen–Eudragit L100 Binary Powder Mixing Process

The binary systems of naproxen–Eudragit L100 with various ratios of Eudragit L100 over Naproxen (E/N ratio) were mixed and assessed by FOSS NIRS. The blending end point was determined by the STDEV method described in our previous work [15]. Care was exercised to ensure the sufficient homogeneity for the blends and the sampling representativeness. Only the NIR spectra of the final well-blended binary mixtures at each E/N ratio were used for the PCA. It was found that PC1 and PC2 can explain 98 and 2% of the total variability captured by the NIR spectra, separately. As shown in Figure 9.2, process trajectory depicted by the PCA score plot demonstrated three distinct segments or clusters that can be approximated by three straight lines:

1) The leftmost segment starts with pure component of naproxen and is a cluster of binary blends dominated with naproxen. In addition, the plot of PC2 score versus PC1 score has a positive slope. Furthermore, the higher the Eudragit L100–naproxen ratio, the larger the scores of both PC1 and PC2. However, the situation in the vicinity of the transition is complicated. As expected, the PC2 score value for point $E/N = 0.257$ is larger than that for point $E/N = 0.206$. This follows the trend displayed by other three points ($E/N = 0$, 0.098, 0.113).
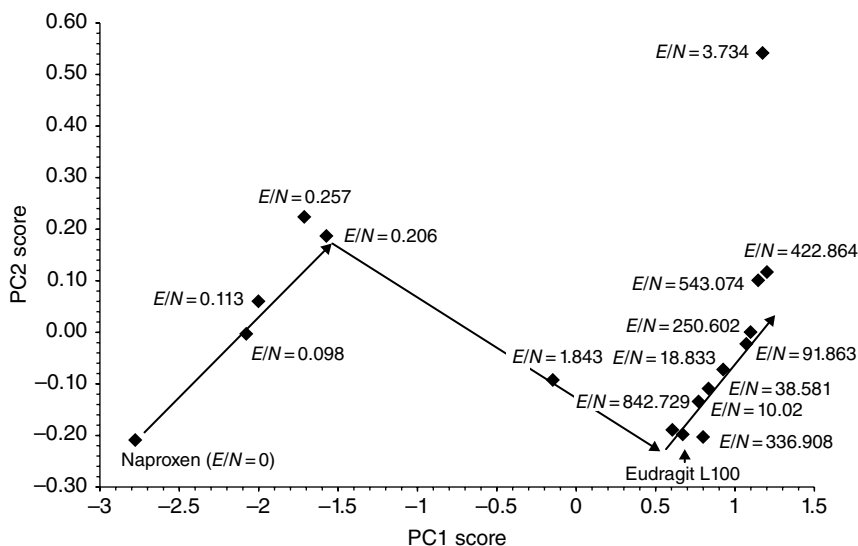

Figure 9.2 PCA-based process trajectory for binary powder blends of naproxen and Eudragit L100. The binary powder blends were made by mixing naproxen and Eudragit L100 with various weight ratios as shown in the figure by *E/N* values. After mixing well, the blends were transferred to vials that were then scanned by FOSS NIRS offline. The PCA was conducted on NIR spectra of the binary blends with various *E/N* values.

However, the absolute PC1 score value for point $E/N = 0.257$ is a bit smaller than that for point $E/N = 0.206$. This does not follow the trend exhibited by other three points. Possible factors that may attribute to this situation include (i) abnormal powder behavior in the vicinity of transition point, such as local inhomogeneity due to back-mixing, and (ii) experimental error.

2) The second segment links the leftmost segment and the rightmost segment, which is a cluster of binary blends dominated by either naproxen or Eudragit L100, respectively. In addition, the plot of PC2 score versus PC1 score has a negative slope.

3) The third segment is on the rightmost, starts with the pure component of Eudragit L100, and is a cluster of binary blends dominated by Eudragit L100. The plot of PC2 score versus PC1 score has a positive slope. It was found that apparently the sign change of the slope for the plot of PC2 score versus PC1 score might be linked to the NIR detection limit of each component in the binary mixture. That is, when the actual weight (mole) fraction of one component is decreased to such an extent that it becomes a minor component in the binary mixture, the signal of the minor component becomes so weak that the presence of the minor component cannot be evidenced solely and directly by examining the NIR raw spectra of the blends. As demonstrated in this case, the process trajectory based on PCA score plots has the capability of differentiating the information embedded. For example, it is able to distinguish when the dominant source of contributing component is switching, as demonstrated by the second segment in Figure 9.2. The second segment serves as a process transition window.

9.4.2 Using In-Line NIR Spectroscopy to Monitor the Alcohol–Water Binary Liquid Mixing Process

In-line NIRS was used to monitor the binary mixing process of alcohol and water in real time. Alcohol and water with various volume ratios were mixed and homogenized. The mixing process was monitored using in-line NIRS. The NIR spectra of the well-mixed binary mixture were complied as an NIR data set. PCA was then applied to this data set. It was found that PC1 and PC2 account for 98 and 2%, respectively, of the total variability embedded with the NIR spectra of the binary mixtures. As shown in Figure 9.3, process trajectory depicted by the PCA score plot demonstrated that there is a process transition window where the slope of the plot of PC2 score versus PC1 score changes from positive to negative. The left segment starts with pure component of alcohol and consists of a cluster of alcohol-dominant binary solutions. When water is gradually introduced into alcohol, the scores of both PC1 and PC2 increase. In contrast, the right segment starts with pure component water and consists of a cluster of water-dominant binary solutions. When alcohol is gradually introduced into water, the PC1 score is decreased, while the PC2 score is increased. Furthermore,
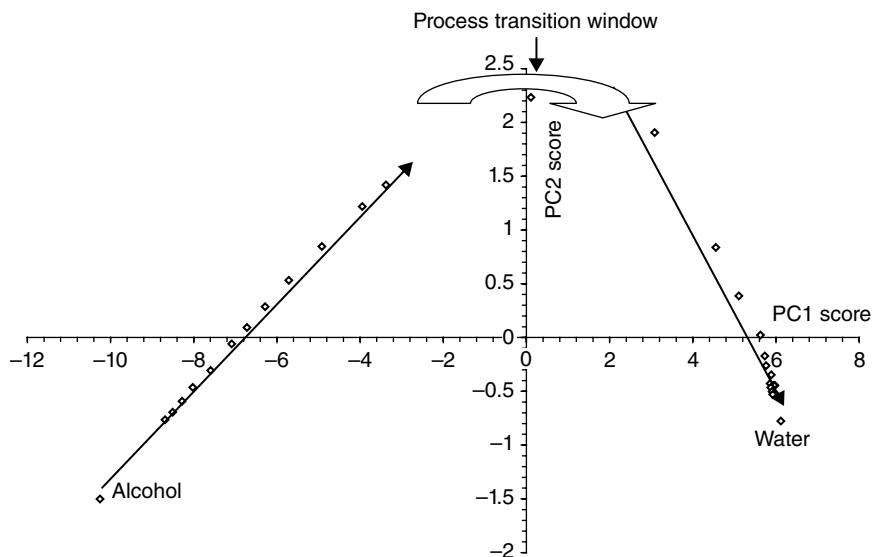
Figure 9.3 PCA-based process trajectory for binary liquid mixtures of alcohol and water with various volume ratios at room temperature. The binary liquid mixtures were made by mixing alcohol and water with various volume ratios. After mixing well, the blends were measured by real-time Brimrose in-line NIR spectroscopy. The PCA was conducted on NIR spectra of the binary blends with various alcohol/water ratios.

it was found that apparently the sign change of the slope for the plot of PC2 score versus PC1 score is linked to the NIR detection limit of each component in the binary mixture. When the actual weight (mole) fraction of one component is decreased to such an extent that it becomes a minor component in the binary mixture, the signal of the minor component becomes so weak that the presence of the minor component cannot be evidenced solely and directly by examining the NIR raw spectra of the blends. However, apparently the process trajectory based on PCA score plots helps to differentiate the information embedded. In other words, the switching of the dominant source of contributing component is demonstrated by the process transition window in Figure 9.3. As in this case, the inhomogeneous nature of the information embedded as well as the discontinuity exemplified by the sign change of the slope in the process trajectory collectively evidenced the existence of an SP.

9.4.3 Real-Time Integrated PAT Monitoring of the Dynamic Coprecipitation Process

When the number of components involved in a multicomponent system is increased, typically the process complexity is increased accordingly. In this subsection, two aspects will be explored: (i) if a similar methodology can be

used to develop the process trajectory and (ii) if the PCA-based process trajectory can reveal critical process events such as nucleation and growth.

9.4.4 3D Map of NIR Absorbance–Wavelength–Process Time (or Process Sample) of the Coprecipitation Process

When in-line NIRS is applied to pharmaceutical unit operations, a lot of process-related information may be captured in real time. For the dynamic coprecipitation process evaluation, prior to the introduction of the nonsolvent (water) to the coprecipitation vessel, there was a ternary system of naproxen–Eudragit L100–alcohol already present in the vessel. When water is gradually introduced into the vessel, according to classical thermodynamics, the overall composition point in the four-component system phase diagram is moved from the liquid phase toward the solid (crystal)–liquid equilibrium line; the zone that is bordered by the solid–liquid equilibrium line (the solubility curve) and the metastable limit constitutes the metastable zone for coprecipitation. The graphical illustration of the movement of equilibrium line and metastable zone is shown in Figure 9.4. Once the solid–liquid equilibrium line is crossed, coprecipitate begins to form and grow. The overall composition point will be located within the solidus region. As discussed later, the three-dimensional (3D) process map can illustrate how the coprecipitation process progress and what process events occur during the process.
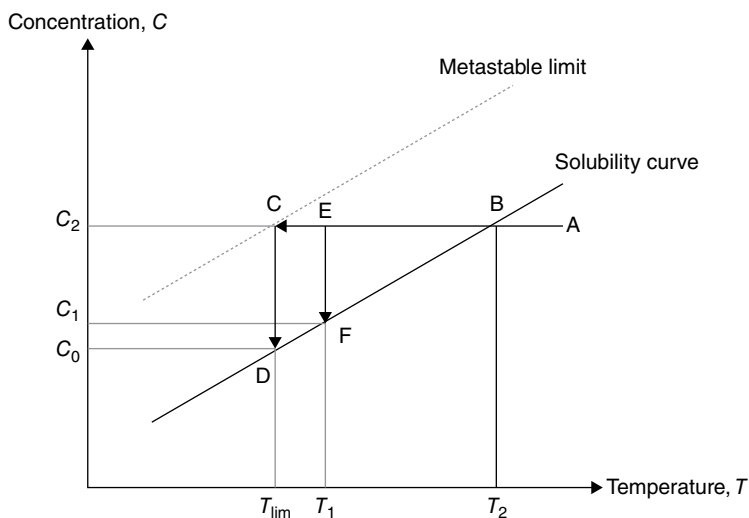

Figure 9.4 Graphical illustration of the movement of equilibrium line and metastable line zone. Concentrations C_0, C_1, and C_2 represent the solubility of solute at temperature T_{lim}, T_1, and T_2, respectively.

To enhance the resolution of the 3D process map for identifying process signatures, the addition rate of the nonsolvent into the system was kept slow, such that the transition of one process state to another state could be observed clearly. The formulation components used, the amount of the nonsolvent added and its adding sequence, and significant visual observations or indications from turbidity measurement are summarized in Table 9.1. As one example, the process 3D (NIR spectral number–wavelength–NIR absorbance) map for batch started with formulation components of 4.7617 g naproxen and 1.205 g Eudragit L100 (batch code (a) in Table 9.1) is shown in Figure 9.5.

9.4.5   Process Signature Identification

The 3D process map is able to provide visual evidence about significant process event for the ongoing process, such as addition of a new component into the existing system, depletion of an existing component, etc. The sensitivity and resolution of this kind of 3D direct and raw process map depend on a number of factors such as (i) the sensitivity and resolution of the instrumentation and the signal/noise ratio of the instrumentation used, (ii) the mass transfer rates in the vicinity of process probe tip and the bulk solution in the vessel, and (iii) the hydrodynamics of the fluid and mixing characteristics in the vicinity of water dispersion area inside the vessel. The physical distance between the process probe tip and the point where the nonsolvent water is introduced and in contact with the bulk solution presents a factor to limit the mass transfer process. In our study, despite the fact that the magnetic stirring of the solution could limit the possible concentration gradient within the vessel eventually, the time lag between the initial moment when the process event take places locally and the moment when the process probe is able to actually detect is always unavoidable due to the existence of physical distance. When a signal associated with a process event is relatively weak, especially when the process event is at its embryo stage, its appearance on the 3D process map will not be that obvious. In this case, other tools may be needed for magnifying or identifying the signature. Furthermore, there are some important questions associated with the 3D process map to be answered. Questions include but are not limited to: (1) What are the implications of the inflection points in the 3D map? (ii) Is there any new phase formation during the process? In this case, other techniques such as chemometrics may play a vital role in terms of identifying process signature associated with process event and provide essential information regarding process progression. In this work, PCA was applied to the process NIR spectra to construct process trajectory and identify SPs such that critical process information and knowledge could be extracted.

Applying PCA to the process NIR spectra of the aforementioned process batch showed that two principal components are sufficient to characterize the variability embedded with NIR spectra. PC1 and PC2 are able to account for

Table 10.1 Risk assessment of the blending process step using FMEA according to Adam *et al*. [1].

Effect	Severity (S)	Failure mode	Detectability (D)	Cause		Probability (P)	RPN
Varying content uniformity of final dosage form	5	Inhomogeneity of final blend	4	Raw material properties	Particle-size ratio	5	100
					Weight ratio	5	100
					Particle shape	5	100
					Cohesivity	5	100
					Material density	4	80
				Process parameters	Fill volume	4	80
					Blending time/total revolutions	5	100
					Blend rpm	1	20
					Order of addition	3	60
				Equipment design	Blender type	4	80
					Blender total volume	3	60
				Environment	Relative humidity	3	60
					Temperature	2	40

Five-point scales for S, D, and P, with 5 as worst-case value and 1 as best-case value, are used. An RPN of 50 is defined as the cutoff limit.
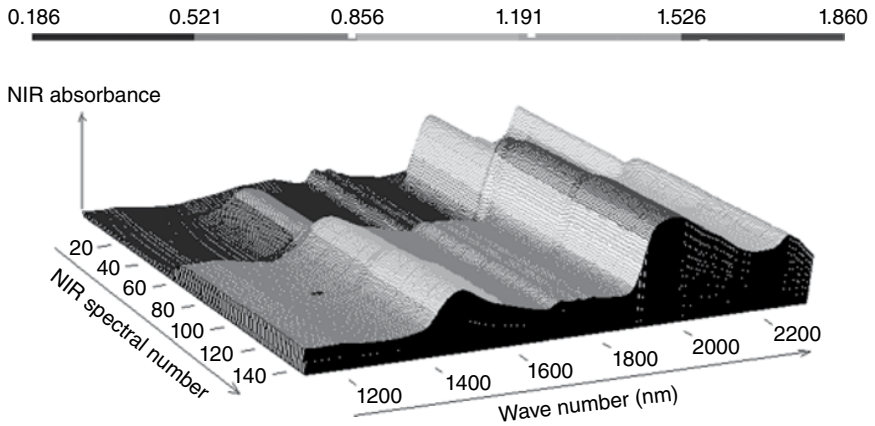
0.186   0.521   0.856   1.191   1.526   1.860

Figure 9.5 Three-dimensional map of the process state over the entire process course of adding water to the ternary system. An inflection point occurs among process spectra window between spectral numbers #60 and #67. New phase is detected between spectral numbers #60 and #67, which maybe an indication of nucleation and growth.
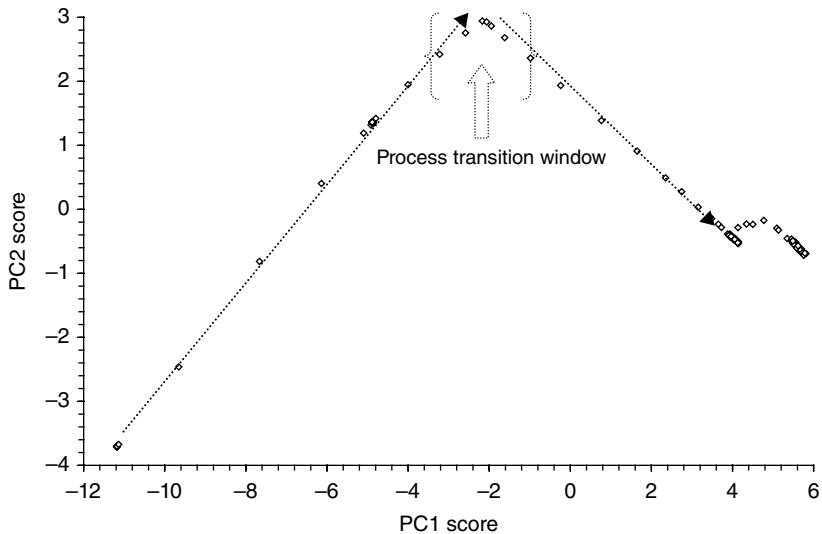

Figure 9.6 PCA score plot for process NIR spectra of the process batch started with formulation components of 4.7617 g naproxen and 1.205 g Eudragit L100.

93 and 7%, respectively, of the total variability embedded with the process NIR spectra for this batch. As shown in Figure 9.6, when PC2 score is plotted against PC1 score, there is a sharp transition on its PCA score plot as reflected by the change in the slope of the straight line of PC2 score versus PC1 sore. The slope

is dramatically changed from positive value to negative value. As a matter of fact, the two segments of the process trajectory consisting of PC2 score versus PC1 score are almost perpendicular to each other. When tracking the sample number or NIR spectral number on this PCA score plot, it was found that (i) the process trajectory that has positive slope value corresponds to the process state of transparent solution during the course of adding water into the ternary system, (ii) the sharp transition of the process trajectory corresponds to the period that solution starts to become cloudy when water is introduced to the ternary system for the third time, and (iii) the process trajectory that has negative slope value corresponds to the process state of cloudy solution during the course of adding water into the ternary system.

The process trajectory depicted by the PCA score plot demonstrated that certain process samples #61–66 can be identified, which cover the process transition window where nucleation and growth take place, as supported by both visual evidences and real-time turbidity profile of the dynamic coprecipitation process. In order to further verify this finding, similar data analysis method was extended to dynamic coprecipitation processes that have different drug/polymer ratios in the starting formulation preparation. In-line NIRS was used to monitor the dynamic coprecipitation process in real time for various batches with different drug/polymer ratios. PCA was applied to each NIR spectra data set associated with a particular drug/polymer ratio. Process trajectory for each coprecipitation process was constructed based on the PCA results. These results together with our previous work [9] collectively demonstrated that process trajectory based on online NIR real-time process monitoring and PCA can differentiate various distinguishable process events and accurately track various process stages such as incubation, nucleation, and crystal growth.

9.4.6 Online Turbidity Monitoring of the Process

For the process runs being conducted in this work, we also used online turbidity measurement to monitor the turbidity change real time during the process. The real-time process slurry turbidity profile of a dynamic coprecipitation process (batch code (b) in Table 9.1) is shown in Figure 9.7. As we can see from Figure 9.7, there are several plateaus observed during the course of introducing water into the ternary system. Before adding any water to the ternary system, the baseline turbidity data of the ternary system were around 0.33–0.45 NTU. The first plateau (1.7–2.6 NTU) and the second plateau (50–70 NTU) correspond to the first and the second addition of 100 ml of water into the ternary system, respectively. The third plateau (140–150 NTU) occurred due to third addition of 100 ml of water into the ternary system. Given that the turbidity value of first plateau of the four-component system is still very close to the baseline turbidity value of the ternary system, most likely it indicates that
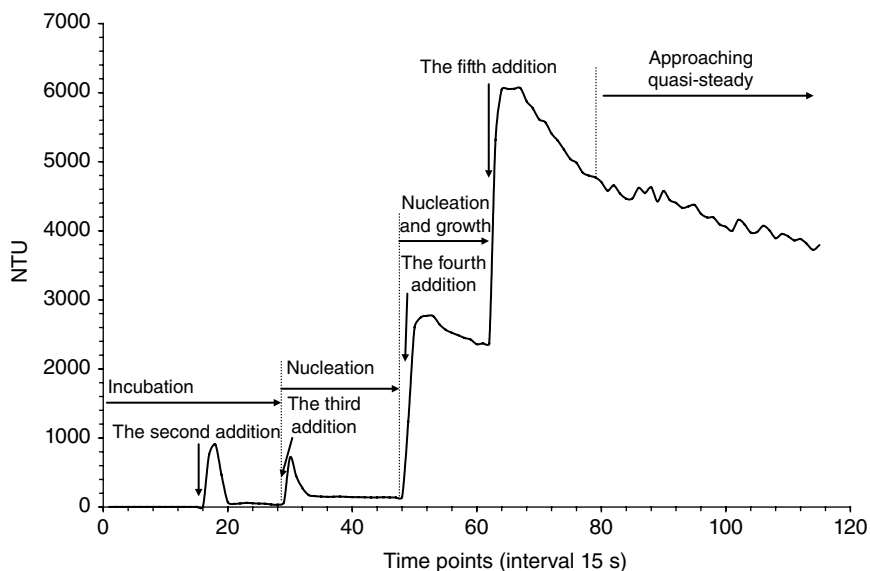
Figure 9.7 Turbidity profile during the process of sequentially adding water to the ternary system of API (0.1320 g), Eudragit L100 (6.0970 g), and alcohol (320 ml) to form the coprecipitate.

the system remains as transparent solution and no new phase is formed (as confirmed by the online process NIR monitoring). The relative low turbidity values for the second turbidity plateau suggest that the process is still at the incubation stage. At this early stage, the composition point of the system is still located within the liquid region. Therefore, no nuclei were formed and the solution was still transparent and thus had a small NTU value. When more water was introduced gradually, the third turbidity plateau (140–150 NTU) occurred, which is probably correlated with the nucleation process stage. At this stage, the system composition point reaches the solid–liquid equilibrium line, and thus nuclei were initiated and a bit higher turbidity values than the second stage were detected. Afterward, when more water was introduced, the system composition point reaches the metastable zone. Consequently, a large supersaturation was created and a crystal growth period was observed. This stage has a turbidity plateau of 2360–2770 NTU. When more and more water was introduced, the system approaches a quasi-steady state where an even higher turbidity plateau of 4000–5180 NTU was observed.

The aforementioned experimental observations could be explained from a multiphase transfer perspective. When water was introduced into the ternary system initially, there were some small nuclei formed initially in the vicinity of the water dispersion area due to local building up of supersaturation; then, the newly formed local nuclei were dissolved quickly due to the

de-supersaturation phenomena via mass transfer (diffusion and convection) enhanced by the magnetic stirring of the slurry solution. However, since the existence of fine nuclei was so short and did not have sufficient time to travel to the area that was directly exposed to the process NIR probe tip, this short shelf-life nuclei was not detected by the Brimrose Process NIRS system. The relative high turbidity values for the third turbidity plateau (140–150 NTU) plus the visible slightly cloudiness of the solution suggest that the process is at the nucleation stage. At this stage, the system composition point reaches the solid–liquid equilibrium line, and thus nuclei were initiated and a bit higher turbidity values than the second stage were detected. When more water was introduced gradually, the fourth turbidity plateau (2360–2770 NTU) occurred, which is probably correlated with a stage that both nucleation and crystal growth are involved since the system composition point reaches the metastable zone. Afterward, when more water was introduced, consequentially, a large supersaturation was created and a crystal growth period was observed.

9.5 Challenges and Opportunities for PCA-Based Data Analysis and Modeling in Pharmaceutical PAT and QbD Development

PCA is a popular and powerful technique for feature extraction and dimensionality reduction and probably one of the most employed techniques of multivariate analysis. Maggio *et al.* reported a new PCA-based approach for testing "similarity" of drug dissolution profiles [25]. Comparison between the area enclosed by the confidence ellipses of the weighted scores plot and the region obtained from the bootstrap-calculated acceptable values of the corresponding f_2 tests suggested that PCA confidence region represents, in general, a more discriminated standard. Otsuka *et al.* used PCA for identifying and predicting the most important variables in the process of granulation and tableting [26]. However, there are some limitations with PCA that we should take into consideration, which include the following: (i) the absence of an associated probability density or generative model; (ii) the subspace itself is restricted to a linear mapping as PCA is a linear method; and (iii) PCA does not reveal any reliable information on time scales that are not actually sampled [27] (e.g., a short molecular dynamics trajectory does not yield an accurate covariance matrix of protein fluctuations). The high-order statistical information is discarded during the linear mapping. Basically PCA involves rotating the ellipsoid in such a way that the direction of the variance of the data comes as the first component. This works fine as long as the X/Y relation is fairly linear. For a situation where the X/Y relation is nonlinear, there is a problem. While PCA still tries to produce components by variance, it fails as the largest variance is not along a

single vector, but along a nonlinear path, with the assumption that the observed data is independent. In real-world process measurement environment, this may not be a valid assumption due to the facts that (i) the measurement data at consecutive time points could be interrelated and (ii) NIR absorbance values at nearby wavelengths could be correlated too. If a system is highly nonlinear or the observed data are not independent, then the limitations of PCA-based process trajectory become a problem. In this challenging case, other methods may provide better option to handle the nonlinearity encountered, as discussed in the following text.

A number of strategies have been proposed to address the aforementioned limitations of PCA. For example, probabilistic principal component analysis (PPCA) [28], kernel principal component analysis (KPCA) [29], and probabilistic kernel principal component analysis (PKPCA) [30–32] have been developed to deal with the first two limitations of PCA. In addition, a hidden Markov model (HMM) [33] was used to obtain an optimized representation of the observed data through time. On the other hand, neural networks [17, 34] are perfectly capable of dealing with nonlinear problems and can on their own do this. Furthermore, they can do scaling directly so that the principal components can be scaled by their importance.

In the pharmaceutical manufacturing setting, it is possible to encounter nonlinear process features due to many variables (such as formulation variables, process variables, environmental variables, etc.) coexisting and possible interactions. Risk analysis and risk assessment [35] may help to rank the relative importance of those variables and thus provide a list of critical variables for further scientific investigation and design space development [17]. On the other hand, depending on the ranges of various variables selected, the impact of interaction among variables could be significant, marginal, or insignificant. This can be quantitatively assessed via design of experiments (DOE) and ANOVA, as illustrated previously [17]. For those critical variables identified via risk analysis and initial DOE study mentioned earlier, another DOE could be conducted to establish the linkage between critical variables and essential response variables (such as key quality attributes) as demonstrated in most DOE-based approaches. Or, as demonstrated in our recent work, an integrated PAT and DOE approach can be developed to establish dynamic linkages between the real-time process behaviors and the essential response variables at both transition state and steady state [17]. Both approaches are important to ensure product quality and help to achieve QbD.

The PCA-based process trajectory is useful to diagnose process healthiness and identify outliers and abnormal situations. However, the applicability of the PCA-based process trajectory for this kind of process QC and process QA really depends on the scope of the design space [36, 37] (formulation and process variables) investigated, for which the process trajectory

was based on. Therefore, it is important to have a clear understanding of where and when PCA-based process trajectory can be applied for process control purpose and of how robust the PCA-based process trajectory is to the process disturbance in real applications, especially for the cases where events and disturbances appear in many different time scales. Issues surrounding PCA-based process trajectory, design space development, PAT process monitoring and control strategies, and QbD methodologies, as discussed briefly in this section, deserve much attention in future research and development during the implementation of PAT [7, 37] and QbD [7, 36–39] in the pharmaceutical sector.

9.6 Conclusions

This work provides a QbD case study that focuses on process trajectory development for a dynamic pharmaceutical coprecipitation process based on an integrated real-time PAT process monitoring strategy. The dynamic coprecipitation process is visualized via three-dimensional map of NIR absorbance–wavelength–process time. The process trajectory based on the results of applying PCA to real-time process NIR spectra data clearly demonstrated that physical meanings can be assigned to various SPs that occurred. Furthermore, those SPs in the process trajectory are directly linked to various distinguishable process events and process signatures such that incubation, nucleation, and crystal growth could be accurately tracked and differentiated. This information and knowledge are essential for developing a suitable design space and operational process space for a pharmaceutical unit operation. The challenges and opportunities of PCA-based process trajectory under the PAT framework and ICH Q8(R2) have been discussed.

Acknowledgments

References

1 Blagden N, Matas M, and Gavan PT, York P. Crystal engineering of active pharmaceutical ingredients to improve solubility and dissolution rates. Advanced Drug Delivery Reviews. 2007;59:617–630.

2 Mcnamara DP, Childs SL, Giordano J, Iarriccio A, Cassidy J, Shet MS, Mannion R, O'Donnell E, and Park A. Use of a glutaric acid cocrystal to improve oral bioavailability of a low solubility API. Pharmaceutical Research. 2006;23:1888–1896.

3 Trask AV, Motherwell WDS, and Jones W. Pharmaceutical cocrystallization: engineering a remedy for caffeine hydration. Crystal Growth & Design. 2005;5:1013–1021.

4 Khan MA, Bolton S, and Kislalioglu MS. Optimization of process variables for the preparation of ibuprofen coprecipitates with Eudragit S100. International Journal of Pharmaceutics. 1994;102:185–192.

5 Zaghloul AA, Faltinek J, Vaithiyalingam SR, Reddy IK, and Khan MA. Naproxen-Eudragit microspheres: screening of process and formulation variables for the preparation of extended release tablets. Pharmazie. 2001;56(4):321–324.

6 Kislalioglu MS, Khan MA, Blount C, Goettsch RW, and Bolton S. Physical characterization and dissolution properties of ibuprofen: Eudragit coprecipitates. Journal of Pharmaceutical Sciences. 1991;80:799–804.

7 FDA. Guidance for Industry. PAT—A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance. 2004. Available at: http://www.fda.gov/downloads/Drugs/ GuidanceComplianceRegulatoryInformation/Guidances/UCM070305.pdf (accessed on April 20, 2017).

8 Yu L, Lionberger RA, Raw AS, D'Costa R, Wu H, and Hussain AS. Applications of process analytical technology to crystallization processes. Advanced Drug Delivery Reviews. 2004;56:349–369.

9 Wu H and Khan MA. 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. Journal of Pharmaceutical Sciences. 2010;99:1516–1534.

10 Wu H, Lyon RC, Hussain AS, Drennen JD III, DasGupta D, Voytilla RJ, and Bejic L. Application of principal component analysis in assessing pharmaceutical formulation design: exploring the casual links between the tablet processing conditions and drug dissolution rate. In: Proceedings of Symposium on Innovations in Pharmaceutical and Biotechnology Development and Manufacturing, 2003 AIChE Annual Meeting [TE001], 2003, pp. 635–642. San Francisco, CA, November 17–21, 2003.

11 Wu H, Heilweil EJ, Hussain AS, and Khan MA. Process analytical technology (PAT): quantification approaches in terahertz spectroscopy for pharmaceutical application. Journal of Pharmaceutical Sciences. 2007;97(2):970–984.

12 Sulub Y, Wabuyele B, Gargiulo P, Pazdan J, Cheney J, Berry J, Gupta A, Shah R, Wu H, and Khan M. Real-time on-line blend uniformity monitoring using near-infrared reflectance spectroscopy: a noninvasive off-line calibration approach. Journal of Pharmaceutical and Biomedical Analysis. 2009;49:48–54.

13 Ciurczak EW and Drennen JD III. Near-infrared spectroscopy in pharmaceutical applications. In: Handbook of Near-Infrared Analysis, 2nd edition, edited by Burns DA and Ciurczak EW, Marcel Dekker, New York, 2001.

14 Xie L, Wu H, Shen M, Augsburger L, Lyon RC, Khan MA, Hussain AS, and Hoag SW. 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. Journal of Pharmaceutical Sciences. 2008;97(10):4485–4497.

15 Wu H and Khan MA. Quality-by-design (QbD): an integrated approach for evaluation of powder blending process kinetics and determination of powder blending end-point. Journal of Pharmaceutical Sciences. 2009;98(8):2784–2798.

16 Wu H, Tawakkul M, White M, and Khan MA. Quality-by-design (QbD): an integrated multivariate approach for the component quantification in powder blends. International Journal of Pharmaceutics. 2009;372(1–2):39–48.

17 Wu H, White M, and Khan MA. Quality-by-design (QbD): an integrated process analytical technology (PAT) approach for a dynamic pharmaceutical co-precipitation process characterization and process design space development. International Journal of Pharmaceutics. 2011;405:63–78.

18 Ündey C, Tatara E, and Çinar A. Real-time batch process supervision by integrated knowledge-based systems and multivariate statistical methods. Engineering Applications of Artificial Intelligence. 2003;16:555–566.

19 Kourti T and MacGregor JF. Process analysis, monitoring and diagnosis, using multivariate projection methods. Chemometrics and Intelligent Laboratory Systems. 1995;28:3–21.

20 Wu H, Hussain AS, and Khan MA. Process control perspective for process analytical technology: integration of chemical engineering practice into semiconductor and pharmaceutical industries. Chemical Engineering Communications. 2007;194(6):760–779.

21 Srinivasan R and Qian MS. Off-line temporal signal comparison using singular points augmented time warping. Industrial and Engineering Chemistry Research. 2005;44:4697–4716.

22 Doan X-T, Srinivasan R, Bapat PM, and Wangikar PP. Detection of phase shifts in batch fermentation via statistical analysis of the online measurements: a case study with rifamycin B fermentation. Journal of Biotechnology. 2007;132:156–166.

23 Volpe G, Volpe G, and Petrov D. Singular-point characterization in microscopic flows. Physical Review. 2008;E77:037301–037304.

24 Jackson DA and Chen Y. Robust principal component analysis and outlier detection with ecological data. Environmetrics. 2004;15:129–139.

25 Maggio RM, Castellano PM, and Kaufman TS. A new principal component analysis-based approach for testing "similarity" of drug dissolution profiles. European Journal of Pharmaceutical Sciences. 2008;34:66–77.

26 Otsuka T, Iwao Y, Miyagishima A, and Itai S. Application of principal component analysis to effectively find important physical variables for optimization of fluid bed granulation conditions. International Journal of Pharmaceutics. 2011;409:81–88.

27 Balsera MA, Wriggers W, Oono Y, and Schulten K. Principal component analysis and long time protein dynamics. Journal of Physical Chemistry. 1996;100:2567–2572.

28 Tipping M and Bishop C. Probabilistic principal component analysis. Journal of the Royal Statistical Society, Series B. 1999;21(3):611–622.

29 Schölkopf B, Smola A, and Müller K. Nonlinear component analysis as a kernel eigenvalue problem. Neural Computation. 1998;10(5): 1299–1319.

30 Tipping M. Sparse kernel principal component analysis. In: Advances in Neural Information Processing Systems, NIPS'00, vol. 13, edited by Leen TK, Dietterich TG; Tresp V, pp. 633–639, 2001.

31 Zhou C. Probabilistic analysis of kernel principal components: mixture modeling and classification. CFAR technical report, car-tr-993, University of Maryland, Department of Electrical and Computer Engineering, College Park, MD, 2003.

32 Zhang Z, Wang G, Yeung D, and Kwok J. Probabilistic kernel principal component analysis. Technical report hkust-cs04-03, The Hong Kong University of Science and Technology, Department of Computer Science, Hong Kong, 2004.

33 Alvarez M and Henao R. Probabilistic kernel principal component analysis through time. In: The 13th International Conference on Neural Information Processing, Lecture Notes in Computer Science, vol. 4232, edited by King I, Wang J, Chan LW, and Wang D, pp. 747–754, Springer-Verlag, Hong Kong, China, October 3–6, 2006.

34 Despagne F and Massart DL. Neural networks in multivariate calibration. The Analyst. 1998;123:157R–178R.

35 FDA/ICH. Guidance for Industry. Q9 Quality Risk Management. 2006. Available at: http://www.fda.gov/downloads/Drugs/ GuidanceComplianceRegulatoryInformation/Guidances/UCM073511.pdf (accessed on April 20, 2017).

36 FDA/ICH. Guidance for Industry. Q8 Pharmaceutical Development. 2006. Available at: http://www.fda.gov/downloads/Drugs/ GuidanceComplianceRegulatoryInformation/Guidances/UCM073507.pdf (accessed on June 20, 2012).

37 FDA/ICH. Guidance for Industry. Q8(R2) Pharmaceutical Development. 2009. Draft available at: http://www.fda.gov/downloads/Drugs/ GuidanceComplianceRegulatoryInformation/Guidances/ucm073507.pdf (accessed on June 20, 2012).

38 FDA/ICH. Q10. Pharmaceutical Quality System. Draft Consensus Guideline. 2007. Available at: http://www.fda.gov/downloads/Drugs/ GuidanceComplianceRegulatoryInformation/Guidances/UCM073517.pdf (accessed on April 20, 2017).

39 Wu H and Khan MA. 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. Journal of Pharmaceutical Sciences. 2011;100(5):1969–1986.