Electronic Supplementary Information (ESI)

Polymorph control in batch seeded crystallizers. A case study with paracetamol.

Lucrèce Nicoud, Filippo Licordari, and Allan S. Myerson

1 Seeds characterization

1.1 Polymorphism

The seeds of form II were characterized by PXRD, Raman spectroscopy and FTIR spectroscopy. Experimental details about offline Raman spectroscopy are given in the Materials and Methods section of the article, whereas experimental details about PXRD and offline FTIR spectroscopy are given here.

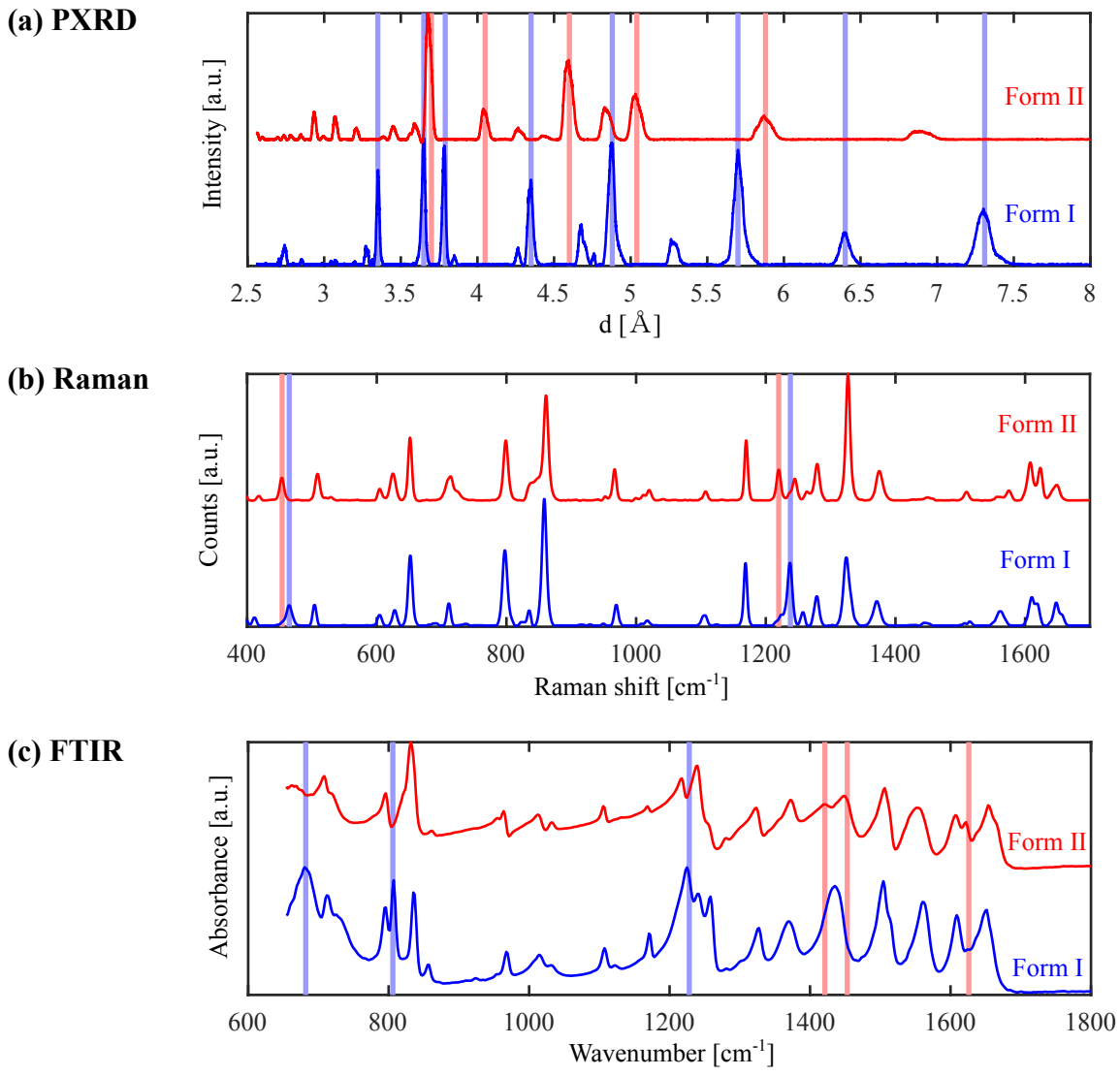

Figure 1: Characterization of both forms of paracetamol with (a) PXRD, (b) Raman spectroscopy, (c) Infrared spectroscopy.

Table 1: Some characteristic features of the PXRD, Raman and IR spectra of both paracetamol forms.

XRD [Å]		**Raman** [cm^{-1}]		**IR** [cm^{-1}]	
Form I	Form II	Form I	Form II	Form I	Form II
3.35	3.70	465	454	802	1421
3.65	4.05	1238	1220	682	1453
3.79	4.60		peak split at 1620	1228	1626
4.35	5.04				
4.88	5.88				
3.35					
5.70					
6.40					
3.35					
7.31					

PRXD measurements were performed with a PANalytical X'pert PRO instrument equipped with a Cu $K_{\alpha 1}$ source ($\lambda = 0.154056$ nm). Data were collected over the range $2\theta = 3 - 35^{\circ}$ at 45 mV with an anode current of 40 mA. Sample spinning was used to reduce preferential orientation effects. The interplanar distance d was computed with Bragg's equation:

$$d = \frac{m\lambda}{2\sin(\theta)} \tag{1}$$

where m is an integer, which was set to 1 here.

Offline FTIR spectroscopy was performed on solid samples with a IdentifyIR instrument from Smith Detection. The instrument uses a diamond attenuated total reflection interface. For each sample, 128 spectra were collected.

The results of the characterization of the seeds with PXRD, Raman and FTIR spectroscopy are shown with red solid lines in panels (a), (b) and (c), respectively, of Figure 1. As a comparison, the results of commercial paracetamol form I are shown with blue solid lines. Some characteristic features of each form are highlighted in the figure and summarized in Table 1 for each analytical technique. Those results are in agreement with literature data [1]. According to the spectra shown in Figure 1, no trace of crystals of form I can be detected in the seeds.

The detection limit of PXRD was assessed by preparing samples of form II containing a small amount of form I, namely 2 and 5% in mass. The corresponding spectra are shown in Figure 2. These results suggest that the detection limit of PXRD is around 5%.

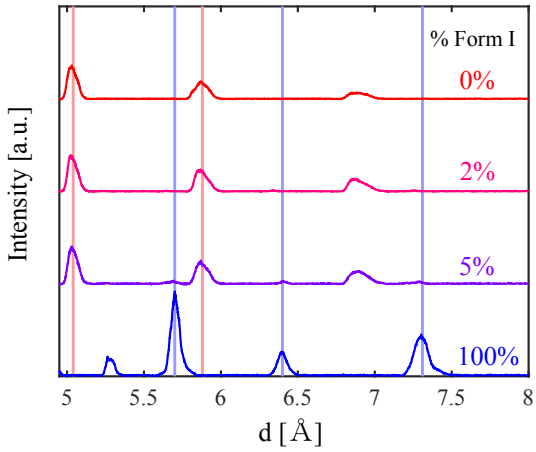

Figure 2: Estimation of the detection limit of PXRD. Samples of paracetamol form II were prepared with various amounts of form I, as indicated in the legend. As a reference, the spectrum of the pure forms are also reported.

1.2 Presence of metacetamol

The potential presence of metacetamol in the seeds was assessed by PXRD and Raman spectroscopy. The spectra of metacetamol are shown in Figure 3 together with those of the seeds. The absence of an intense peak at 5.25 Å in the PXRD spectrum and at 1000 cm^{-1} in the Raman spectrum indicates that the amount of metacetamol in the seeds is so low that it cannot be detected with PXRD nor with Raman spectroscopy.

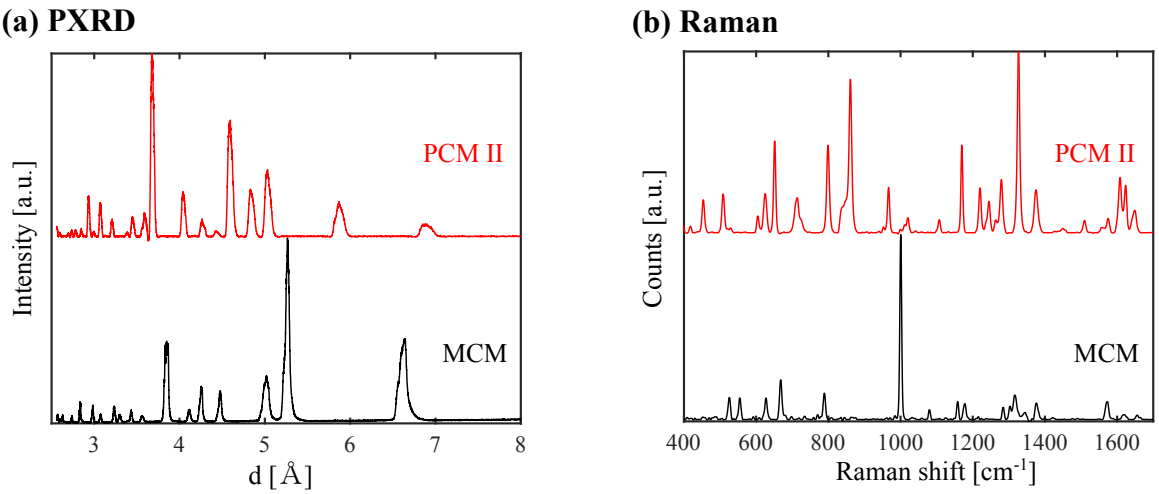

Figure 3: Estimation of the presence of metacetamol (MCM) in the seeds of paracetamol form II (PCM II) with (a) PXRD and (b) Raman spectroscopy.

1.3 Size

Finally, the seeds were characterized in terms of size by analyzing the pictures acquired with the online imaging probe at the beginning of crystallization experiments. The seeds were detected with a Matlab algorithm, which performs the following operations: (1) contrast enhancement, (2) binarization of the image by thresholding, (3) removal of objects smaller than 2 pixels, (4) morphological closing, (5) removal of objects smaller than 100 pixels, (6) object labeling, (7) object sizing. The size of the objects was assessed as the diameter of the sphere having the same area. Approximately 150 seeds were analyzed. The so-obtained distribution is represented as a bar chart in Figure 4. The experimental distribution was then fitted with a Gamma distribution given by:

$$n^{II}_{seeds}(L) = a(aL)^{b-1}\exp(-aL)/\Gamma(b) \tag{2}$$

where L is the size in m, while a and b are positive real quantities that were estimated here to 2.45×10^5 and 2.94, respectively. The average size is given by b/a (here 120 µm), whereas the standard deviation is given by $\sqrt{b/a^2}$ (here 70 µm).

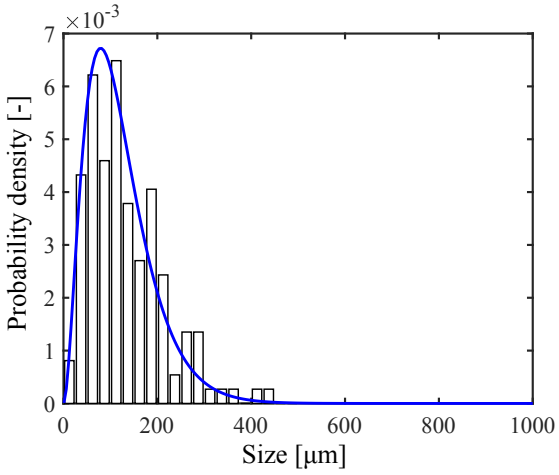

Figure 4: Size distribution of the seeds determined from the pictures acquired with the online imaging probe (bars) together with the fitting with a Gamma distribution (blue line). The area is normalized to 1.

2 Solubility form II

The solubility of form II was assessed from the dissolution behaviour of seeds in solutions of increasing concentrations. As an example, Figure 5 shows pictures acquired at +10 °C, where the solubility of form I is 147 mg/g$^{\mathrm{EtOH}}$. It is seen that at 160 mg/g$^{\mathrm{EtOH}}$, the seeds dissolve within 2 h, which indicates that the concentration under investigation is below

the solubility of form II. On the other hand, at 180 mg/g$^{\mathrm{EtOH}}$ and 200 mg/g$^{\mathrm{EtOH}}$, the seeds do not dissolve during the time frame of the experiment.

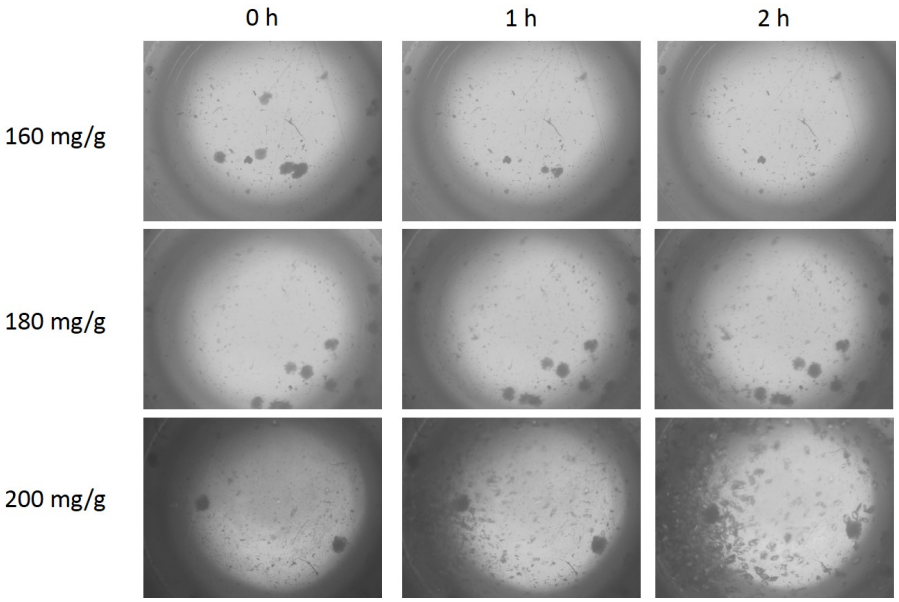

Figure 5: Examples of images recorded to measure the solubility of form II. The experiment was performed at $+10$ °C, where the solubility of form I is 147 mg/g$^{\mathrm{EtOH}}$.

3 Solution density

The density of paracetamol solutions was measured with a density meter from Anton Paar (DMA 4500) at various temperatures and paracetamol concentrations. Experimental conditions were selected below the solubility of form I to avoid precipitation in the instrument. Experimental results are shown with black dots in Figure 6.

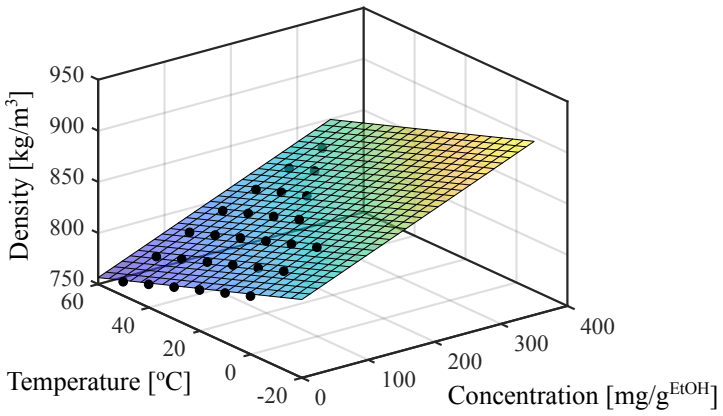

Figure 6: Density of ethanolic paracetamol solutions as a function of temperature and paracetamol concentration.

These data were then fitted with a linear model. The density (in kg/m^3) can be computed as a function of the parac-

etamol concentration (in mg/g$^{\mathrm{EtOH}}$) and temperature (in °C) with the following equation:

$$\rho = 809 + 0.2664\, c^{liq} - 0.877\, T \tag{3}$$

4 Calibration of process analytical tools

4.1 ATR-FTIR

ATR-FTIR was calibrated using paracetamol solutions free of crystals in the concentration range from 100 to 300 mg/g$^{\mathrm{EtOH}}$ and in the temperature range from -10 to $+10$ °C. As an example, the spectrum of a paracetamol solution at 300 mg/g$^{\mathrm{EtOH}}$ acquired at $+10$ °C is shown in Figure 7(a) with a black line. For comparison purposes, the spectrum of pure ethanol is shown with a grey line.

IR spectra were recorded over the course of 10 min and average absorbance values were considered for calibration. The stirring speed was set to 75 rpm to avoid crystal formation during the time frame of the experiment and it was verified with online microscopy that no crystals were formed in the investigated conditions. In addition, it was verified that the stirring speed does not affect the IR spectra at 100 mg/g$^{\mathrm{EtOH}}$, where it is possible to use a high stirring speed without inducing crystal formation. The calibration was performed based on the peak height at 1514 cm^{-1} after substracting the intensity at 1801 cm^{-1}, which corresponds to the intensity of the baseline. The results of the calibration are shown in Figure 7(b). The equations of the linear regressions at each temperature are given in the main text of the article.

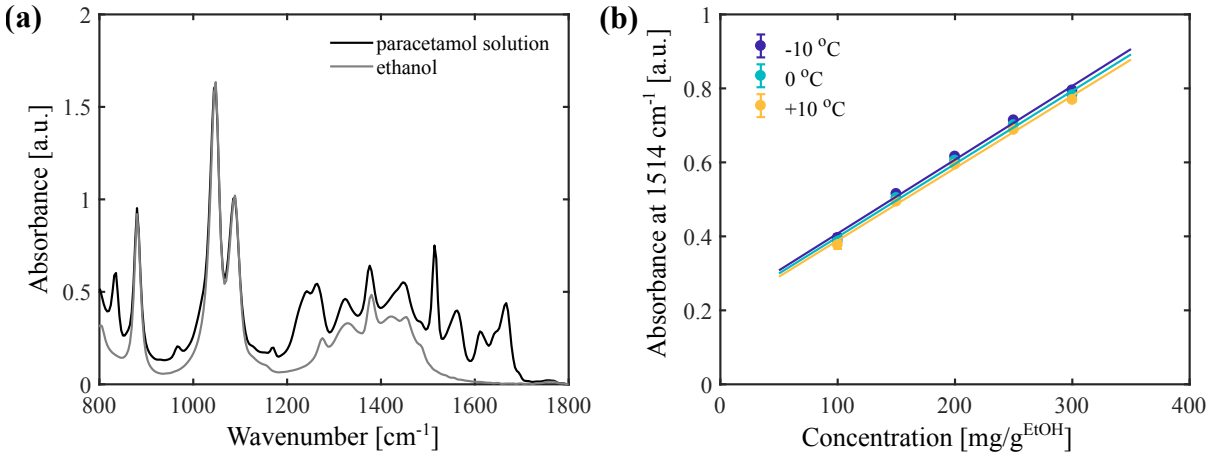

Figure 7: (a) ATR-FTIR spectrum of an ethanolic solution of paracetamol at the concentration of 300 mg/g$^{\mathrm{EtOH}}$ and temperature of $+10$ °C. As a comparison, the spectrum of pure ethanol is also shown. (b) Absorbance at 1514 cm^{-1} as a function of the concentration of dissolved paracetamol for three temperatures. The intensity at 1801 cm^{-1}, which corresponds to the intensity of the baseline, was substracted.

4.2 Offline Raman

The offline Raman calibration was performed with powder samples containing various proportions of form I and form II paracetamol. Fine powders of each form were used to allow a good mixing. The mixture was gently mixed without grinding to avoid any potential solid-state transformation. For each acquired spectrum, the areas of the peaks at 454 cm^{-1} and 465 cm^{-1} were computed by fitting the corresponding region of the spectrum with the sum of two Gaussian functions after baseline substraction. An example of peak deconvolution is shown in Figure 8(a).

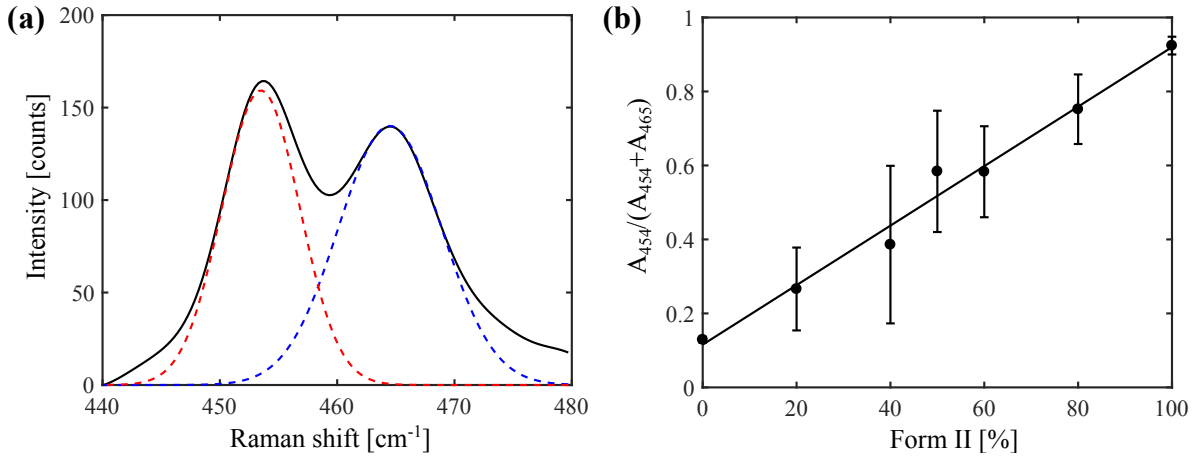

Figure 8: (a) Example of peak deconvolution of an offline Raman spectrum (50:50 mixture). (b) Ratio of the area at 454 cm^{-1} divided by the sum of the areas at 454 cm^{-1} and 465 cm^{-1} as a function of the mass percentage of form II.

The ratio of the area at 454 cm^{-1} divided by the sum of the areas at 454 cm^{-1} and 465 cm^{-1} is plotted as a function of the mass percentage of form II in Figure 8(b). The points and error bars correspond to the averages and standard deviations obtained by collecting ten spectra at different locations of the sample. It is seen in Figure 8(a) that the area at 454 cm^{-1} divided by the sum of the areas at 454 cm^{-1} and 465 cm^{-1} increases linearly with the mass percentage of form II. The regression parameters are given in the main text of the article.

4.3 Online Raman

Figure 9(a) shows the online Raman spectra acquired during an experiment with the following operating conditions: $T = 0$ °C, $c^0_{liq} = 300$ mg/g$^{\mathrm{EtOH}}$, $N = 400$ rpm. The solid black line corresponds to the spectra collected at the the moment of seeds addition, while the dotted black line corresponds to the spectrum collected at equilibrium (approximately 6.5 h). As a reference, the spectrum of pure ethanol is shown with a gray solid line.

The impact of temperature on the Raman signal was assessed using a paracetamol solution at 300 mg/g$^{\mathrm{EtOH}}$. The solution was kept for about 10 min at each temperature at a stirring speed of 75 rpm. It was verified with the online

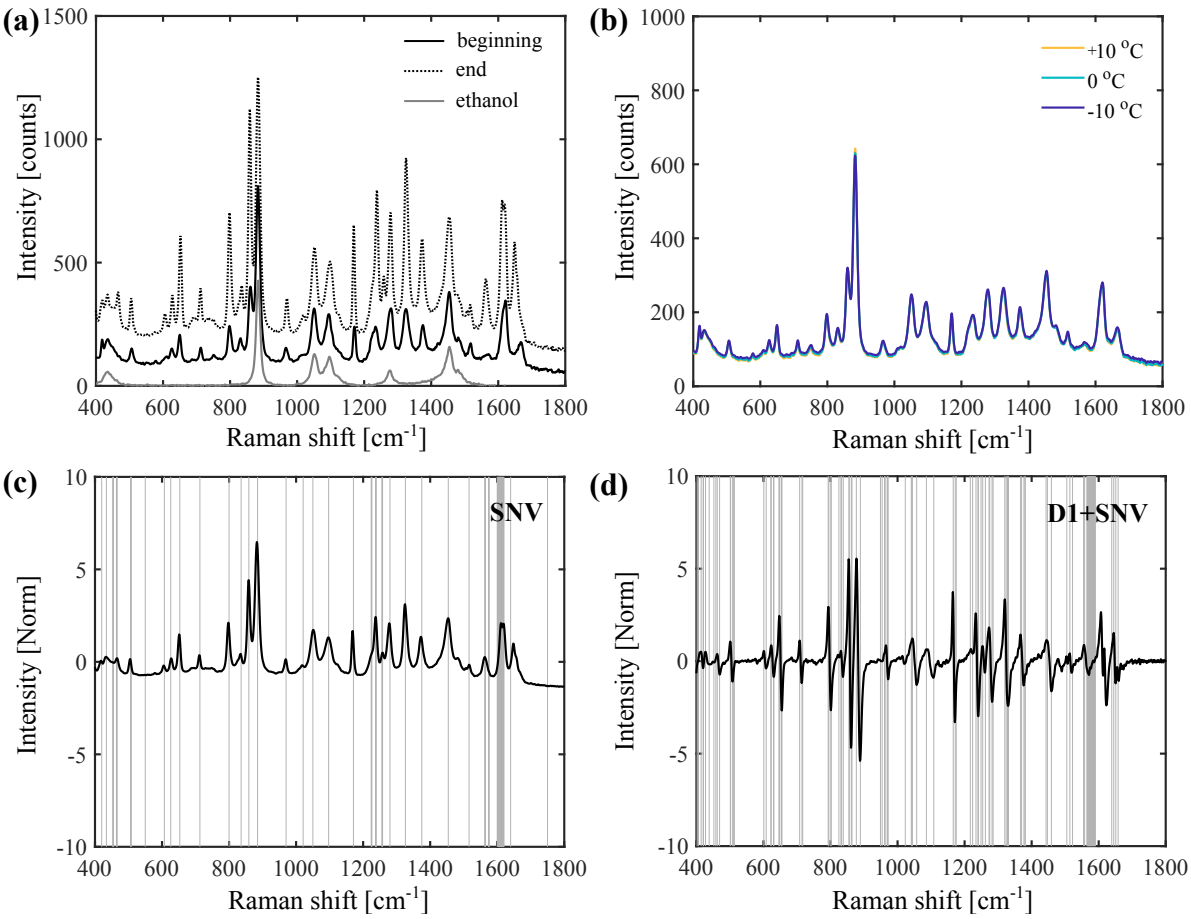

Figure 9: (a) Raman spectra obtained at the beginning and end of a crystallization experiment. The following operating conditions were used: $T = 0$ °C, $c^0_{liq} = 300$ mg/g$^{\mathrm{EtOH}}$, $N = 400$ rpm. The spectrum obtained with pure ethanol at the same temperature is also shown. (b) Raman spectra of a paracetamol solution at 300 mg/g$^{\mathrm{EtOH}}$ at three temperatures. (c) Spectrum 'end' of panel (a) after standard normal variate treatment together with selected Raman shifts for PCA. (d) Spectrum 'end' of panel (a) after first derivative followed by standard normal variate treatment together with selected Raman shifts for PCA.

imaging probe that no crystals formed during the experiment. The results are shown in Figure 9(b). It is seen that the Raman signal is insensitive to temperature changes, at least in the investigated range.

To apply PCA, the number of Raman shifts should be smaller than the number of measured samples (i.e., here time points). To satisfy this constraint, a reduced number of wavelength was selected. These wavelength primarily consist of peak maxima and are highlighted in Figure 9(c and d) for the two types of pre-processing investigated, namely standard normal variate and first derivative followed by standard normal variate, respectively.

5 Shape factor of crystals of form II

Figure 10 shows the morphology selected to describe form II crystals, which is a simplification of the morphology reported in the literature [5]. The symbols L_l, L_w, and L_t denote the length, width, and thickness of crystals, respectively. The symbols H and θ denote the angle and height of the pyramids, while h is defined as shown on the right-hand-side of Figure 10.

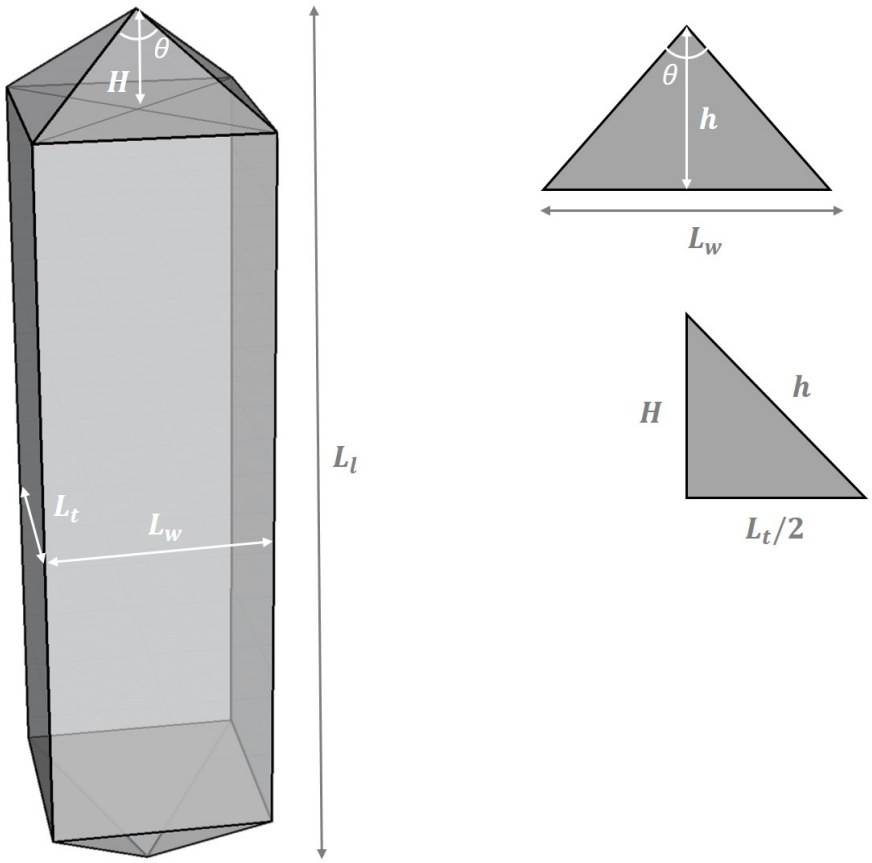

Figure 10: Morphology selected to estimate the shape factor of form II crystals.

The volume of a crystal is computed as:

$$V = (L_l - 2H)L_w L_t + 2\frac{L_w L_t H}{3} \tag{4}$$

Here, L_w is chosen as the characteristic length so that the shape factor is defined as:

$$V = k_v L_w^3 \tag{5}$$

so the goal is to express V as a function of L_w only.

From simple trigonometric relations we can write:

$$h^2 = H^2 + \left(\frac{L_t}{2}\right)^2 \tag{6}$$

$$\frac{\cos(\theta/2)}{\sin(\theta/2)} = \frac{h}{L_w/2} \tag{7}$$

so that

$$H = \frac{1}{2}\sqrt{\frac{\cos^2(\theta/2)}{\sin^2(\theta/2)}L_w^2 - L_t^2} \tag{8}$$

To compute the shape factor, it is necessary to assume that the length and thickness are proportional to the width. Accordingly, we define ω_l and ω_t as:

$$\omega_l = \frac{L_l}{L_w} \quad \text{and} \quad \omega_t = \frac{L_t}{L_w} \tag{9}$$

It follows that:

$$V = \omega_t \left(\omega_l - \frac{2}{3}\sqrt{\frac{\cos^2(\theta/2)}{\sin^2(\theta/2)} - \omega_t^2} \right) \times L_w^3 \tag{10}$$

Therefore:

$$k_v = \omega_t \left(\omega_l - \frac{2}{3}\sqrt{\frac{\cos^2(\theta/2)}{\sin^2(\theta/2)} - \omega_t^2} \right) \tag{11}$$

By analyzing approximately 30 crystals, the following parameter values were estimated: $\theta \approx 100°$, $\omega_l \approx 8.8$, $\omega_t \approx 0.23$. This leads to $k_v \approx 2$.

6 Additional kinetic experiments

Figure 11 shows two experiments that were used to determine kinetic parameters in addition to those presented in the main article.

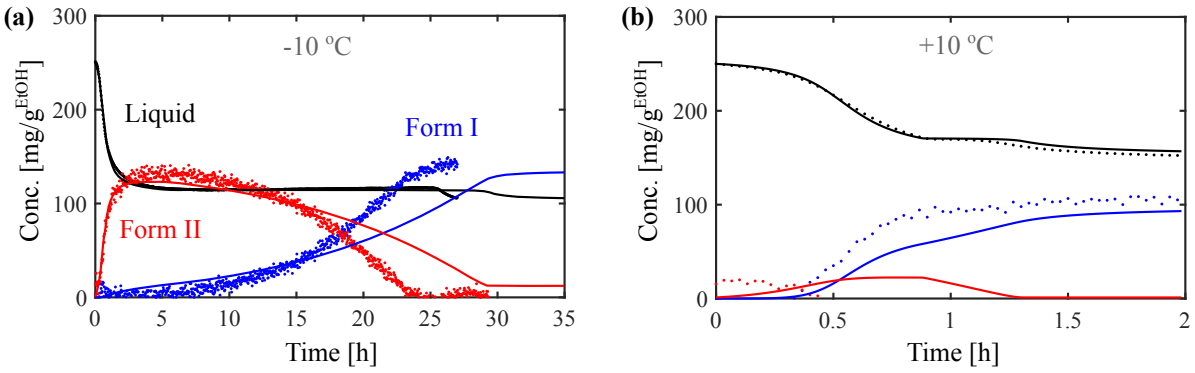

Figure 11: Kinetic experiments performed at (a) -10 °C, and (b) $+10$ °C. In both cases, $c_0^{liq} = 250$ mg/g$^{\text{EtOH}}$, and $N = 400$ rpm .

7 Comparison of nucleation and growth rates with the literature

The secondary nucleation and growth rates of form I estimated in this work are compared to those obtained by Frawley [2], Li [3], Mitchell [4], and Worlitshek [6]. In all these studies, paracetamol form I was obtained by cooling crystallization

from an ethanolic solution. The results are plotted as a function of c^{liq}/c_I^* in Figure 12.

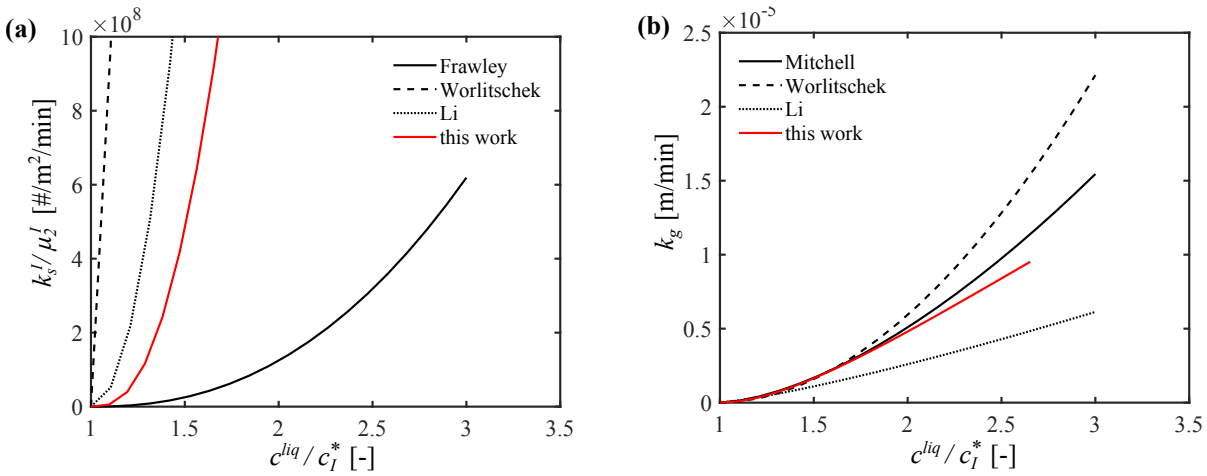

Figure 12: Comparison of the (a) secondary nucleation rates and (b) growth rates of form II as a function of c^{liq}/c_I^* reported in different studies.

8 Process performance parameters

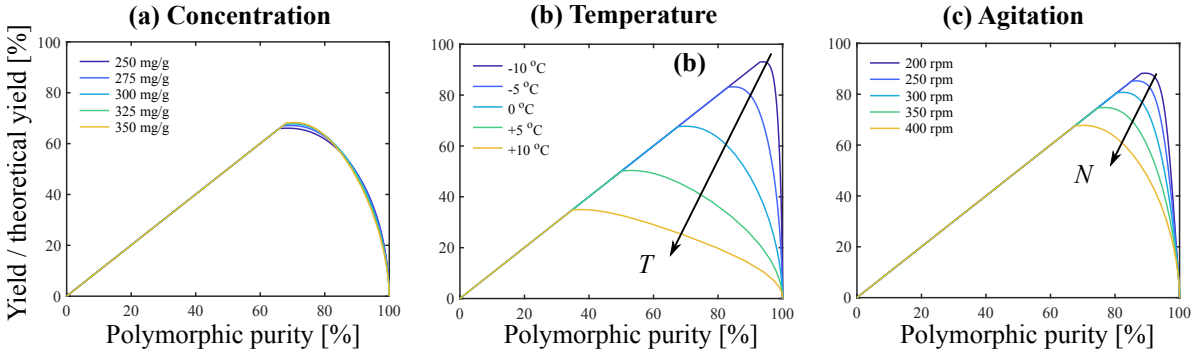

Figure 13: Impact of (a) concentration, (b) temperature, and (c) agitation on the trade-off between yield and polymorphic purity. The yield is normalized by the theoretical yield defined as $\vartheta^* = (c_0^{liq} - c_{II}^*)/c_0^{liq}$. In panel (a) $T = 0$ °C and $N = 400$ rpm. In panel (b) $c_0^{liq} = 300$ mg/g$^{\text{EtOH}}$ and $N = 400$ rpm. In panel (c) $T = 0$ °C and $c_0^{liq} = 300$ mg/g$^{\text{EtOH}}$.

References

[1] N Al-Zoubi, JE Koundourellis, and S Malamataris. FT-IR and raman spectroscopic methods for identification and quantitation of orthorhombic and monoclinic paracetamol in powder mixes. *J Pharm Biomed Anal*, 29(3):459–467, 2002.

[2] PJ Frawley, NA Mitchell, Clifford T Ó'Ciardha, and KW Hutton. The effects of supersaturation, temperature, agitation and seed surface area on secondary nucleation. *Chem Eng Sci*, 75:183–197, 2012.

[3] H Li, Y Kawajiri, MA Grover, and RW Rousseau. Modeling of nucleation and growth kinetics for unseeded batch cooling crystallization. *Indus Eng Chem Res*, 56(14):4060–4073, 2017.

[4] NA Mitchell, CT Ó'Ciardhá, and Patrick J Frawley. Estimation of the growth kinetics for the cooling crystallisation of paracetamol and ethanol solutions. *J Cryst Growth*, 328(1):39–49, 2011.

[5] G Nichols and CS Frampton. Physicochemical characterization of the orthorhombic polymorph of paracetamol crystallized from solution. *J Pharm Sci*, 87(6):684–693, 1998.

[6] J Worlitschek and M Mazzotti. Model-based optimization of particle size distribution in batch-cooling crystallization of paracetamol. *Cryst Growth Des*, 4(5):891–903, 2004.