Electronic Supplementary Material (ESI) for CrystEngComm.
This journal is © The Royal Society of Chemistry 2019

SUPPLEMENTARY INFORMATION

Methods for Estimating Supersaturation in Antisolvent Crystallization Systems

*Jennifer M. Schall, Gerard Capellades and Allan S. Myerson**

Department of Chemical Engineering, Massachusetts Institute of Technology, E19-502D, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139-4307, United States

***Corresponding author e-mail:** myerson@mit.edu.

Consolidated gPROMS models

gPROMS model summary for MFAD supersaturation estimation

Calculated variable	Equation
Solute supersaturation	$$\sigma = \ln\left(\frac{\gamma x}{\gamma_{sat} x_{sat}}\right)$$
Mother liquor concentration (molar)	$$x = \frac{C_{ml} / MW_{sol}}{\sum w_i / MW_i}$$
Effective temperature	$$\ln(x) = \left(\alpha_{11} + \alpha_{12} v_S + \frac{\alpha_{13}}{v_S} + \alpha_{14} \ln(v_S)\right) + \frac{\alpha_{21} + \alpha_{22} v_S + \alpha_{23} / v_S + \alpha_{24} \ln(v_S)}{T_e} + \left(\alpha_{31} + \alpha_{32} v_S + \frac{\alpha_{33}}{v_S} + \alpha_{34} \ln(v_S)\right) \ln(T_e)$$
Solute activity	$$\gamma = \frac{1}{x} exp\left[-\frac{\Delta H_m}{R T_m} ln \frac{T_m}{T_e}\right]$$
Solute solubility	$$\ln(x_{sat}) = \left(\alpha_{11} + \alpha_{12} v_S + \frac{\alpha_{13}}{v_S} + \alpha_{14} \ln(v_S)\right) + \frac{\alpha_{21} + \alpha_{22} v_S + \alpha_{23} / v_S + \alpha_{24} \ln(v_S)}{T} + \left(\alpha_{31} + \alpha_{32} v_S + \frac{\alpha_{33}}{v_S} + \alpha_{34} \ln(v_S)\right) \ln(T)$$
Saturation activity	$$\gamma^{sat} = \frac{1}{x^{sat}} exp\left[-\frac{\Delta H_m}{R T_m} ln \frac{T_m}{T}\right]$$

gPROMS variables*

Symbol	Description	Units	gPROMS identifier	Size	Type
C_{ml}	Solute mother liquor concentration (mass)	g/g	ml_concentration_mass	-	ML_conc_mass
T	Crystallizer temperature	K	temperature	-	temperature_gFP
T_e	Effective temperature	K	temperature_effective	-	temperature_gFP
v_S	Solvent volume fraction in the magma's solvent system	mL/mL	vol_fraction_solvent	-	volume_fraction
w_i	Mass fraction of component i in the MSMPR magma	g/g	mass_fraction_tot	components	mass_fraction_gFP
x	Solute mother liquor concentration (molar)	mol/mol	ml_concentration_mf	-	ML_conc_mf
x_{sat}	Solute solubility (molar)	mol/mol	solubility	-	ML_conc_mf
γ	Solute activity coefficient	-	activityc_solute_ml	-	activityc
γ_{sat}	Saturation activity coefficient	-	activityc_saturation	-	activityc
σ	Solute supersaturation	-	supersaturation	-	supersaturation_ratio

*Gray shading indicates variables specified as model inputs.

Variable types ending in gFP are default from gPROMS Formulated Products. Other are user created.

gPROMS parameter summary for LAM

Parameter inputs for gPROMS model for LAM						
Symbol	**Description**	**Units**	**gPROMS identifier**	**Size**	**Type**	**Value**
NC	System components	-	components	-	ORDERED_SET	3 (API, Solvent, antisolvent}
MW_1	Component 1 molar weight (API)	kg/mol	molar_weight	components	REAL	0.150134
MW_2	Component 2 molar weight (Solvent)	kg/mol	molar_weight	components	REAL	0.01801528
MW_3	Component 3 molar weight (Antisolvent)	kg/mol	molar_weight	components	REAL	0.060096
R	Ideal gas constant	kJ/mol/K	ideal_gas_constant	-	REAL	0.00831
T_m	Solute melting point	K	melting_point_sol	-	REAL	511.75
α_{11}	Solubility expression term	various	(real number)	12	REAL	3539.593045
α_{12}	Solubility expression term	various	(real number)	12	REAL	-4110.478551
α_{13}	Solubility expression term	various	(real number)	12	REAL	660.4109913
α_{14}	Solubility expression term	various	(real number)	12	REAL	3645.735023
α_{21}	Solubility expression term	various	(real number)	12	REAL	-174415.0757
α_{22}	Solubility expression term	various	(real number)	12	REAL	197949.6118
α_{23}	Solubility expression term	various	(real number)	12	REAL	-31558.52507
α_{24}	Solubility expression term	various	(real number)	12	REAL	-174717.3736
α_{31}	Solubility expression term	various	(real number)	12	REAL	-520.4480379
α_{32}	Solubility expression term	various	(real number)	12	REAL	605.9438293
α_{33}	Solubility expression term	various	(real number)	12	REAL	-97.48256172
α_{34}	Solubility expression term	various	(real number)	12	REAL	-537.4671788
ΔH_m	Solute enthalpy of fusion	kJ/mol	enthalpy_fusion_sol	-	REAL	135.03

Conversion of volume based concentrations to mass fractions was conducted assuming additive volumes and using the following densities:

Component	**Density (kg/m^3)**
LAM	1543
Water (solvent)	1000
Isopropanol (antisolvent)	786

gPROMS parameter summary for API

Parameter inputs for gPROMS model for API						
Symbol	**Description**	**Units**	**gPROMS identifier**	**Size**	**Type**	**Value**
NC	System components	-	components	-	ORDERED_SET	3 (API, Solvent, antisolvent)
MW_1	Component 1 molar weight (API)	kg/mol	molar_weight	components	REAL	(REDACTED)
MW_2	Component 2 molar weight (Solvent)	kg/mol	molar_weight	components	REAL	0.047603989
MW_3	Component 3 molar weight (Antisolvent)	kg/mol	molar_weight	components	REAL	0.01801488
R	Ideal gas constant	kJ/mol/K	ideal_gas_constant	-	REAL	0.00831
T_m	Solute melting point	K	melting_point_sol	-	REAL	(REDACTED)
α_{11}	Solubility expression term	various	(real number)	12	REAL	-7889.3347
α_{12}	Solubility expression term	various	(real number)	12	REAL	12288.8427
α_{13}	Solubility expression term	various	(real number)	12	REAL	-4942.5091
α_{14}	Solubility expression term	various	(real number)	12	REAL	-16439.7744
α_{21}	Solubility expression term	various	(real number)	12	REAL	331574.5041
α_{22}	Solubility expression term	various	(real number)	12	REAL	-520921.6338
α_{23}	Solubility expression term	various	(real number)	12	REAL	211152.3762
α_{24}	Solubility expression term	various	(real number)	12	REAL	700687.5715
α_{31}	Solubility expression term	various	(real number)	12	REAL	1190.5329
α_{32}	Solubility expression term	various	(real number)	12	REAL	-1853.9568
α_{33}	Solubility expression term	various	(real number)	12	REAL	744.8459
α_{34}	Solubility expression term	various	(real number)	12	REAL	2479.0816
ΔH_m	Solute enthalpy of fusion	kJ/mol	enthalpy_fusion_sol	-	REAL	(REDACTED)

Error analysis

Because of the experimental difficulty in measuring differential heat capacities, ΔC_p, the supersaturation calculation methodology presented in this work approximates this term as the solute's entropy of fusion. This approximation was considered, instead of the van't Hoff approach, as it is expected that common organic crystallization solutes will present high melting points and high differential heat capacities.

The generalized solubility expression has the form of:

$$x^{sat} = \frac{1}{\gamma^{sat}} exp\left[\frac{\Delta H_{tp}}{R}\left(\frac{1}{T_{tp}} - \frac{1}{T}\right) - \frac{\Delta C_p}{R}\left(ln\frac{T_{tp}}{T} - \frac{T_{tp}}{T} + 1\right)\right] \quad \text{(S1)}$$

When the solution is ideal, the activity coefficient has a value of unity and equation S1 gives the ideal solubility. This value is only dependent on the solute's physical properties and the operating temperature, and it can be used to assess potential errors associated with different approximations.[1] First, we want to demonstrate that, while the van't Hoff approximation is common practice,[2] it can lead to significant errors in the prediction of ideal solubilities for common pharmaceuticals. This can be done through the use of equation S1, and assuming that the triple point and the enthalpy of change can be approximated as the melting point and the enthalpy of fusion, respectively. For an ideal system, equation S1 reduces to:

$$x^{sat} = exp\left[\frac{\Delta H_m}{R}\left(\frac{1}{T_m} - \frac{1}{T}\right) - \frac{\Delta C_p}{R}\left(ln\frac{T_m}{T} - \frac{T_m}{T} + 1\right)\right] \quad \text{(S2)}$$

For this study, we chose four relevant pharmaceuticals, with the relevant physical properties summarized in Table S1. These properties can be used to calculate the ideal solubilities both when the heat capacity term is neglected (van't Hoff approximation), and when it is approximated as the entropy of fusion.

Table S1. Properties of the studied pharmaceuticals.

Compound	Data source	ΔH_m (kJ/mol)	T_m (K)	ΔS_m (J/molK)
Acetaminophen	Neau *et al.*[3]	27.0	441.7	61.1
Ibuprofen	Pappa *et al.*[4]	25.5	347.2	73.6
Mannitol	Neau *et al.*[3]	50.6	438.7	115.3
Naproxen	Neau *et al.*[3]	31.5	428.5	73.5

Figure S1 shows the ideal solubilities as a function of temperature for the four studied pharmaceuticals. In every case, the two methods give similar results when the temperature is near the melting point. This behavior has been used to justify the use of the van't Hoff approximation as a general rule. While this approach may be applicable for compounds with low melting points like ibuprofen, an organic molecular pharmaceutical that melts at 347 K is an exception rather than the rule. Most pharmaceuticals will operate in the ranges shown for acetaminophen, naproxen or mannitol, and give much higher errors in the estimated solubilities when neglecting the heat capacity term.

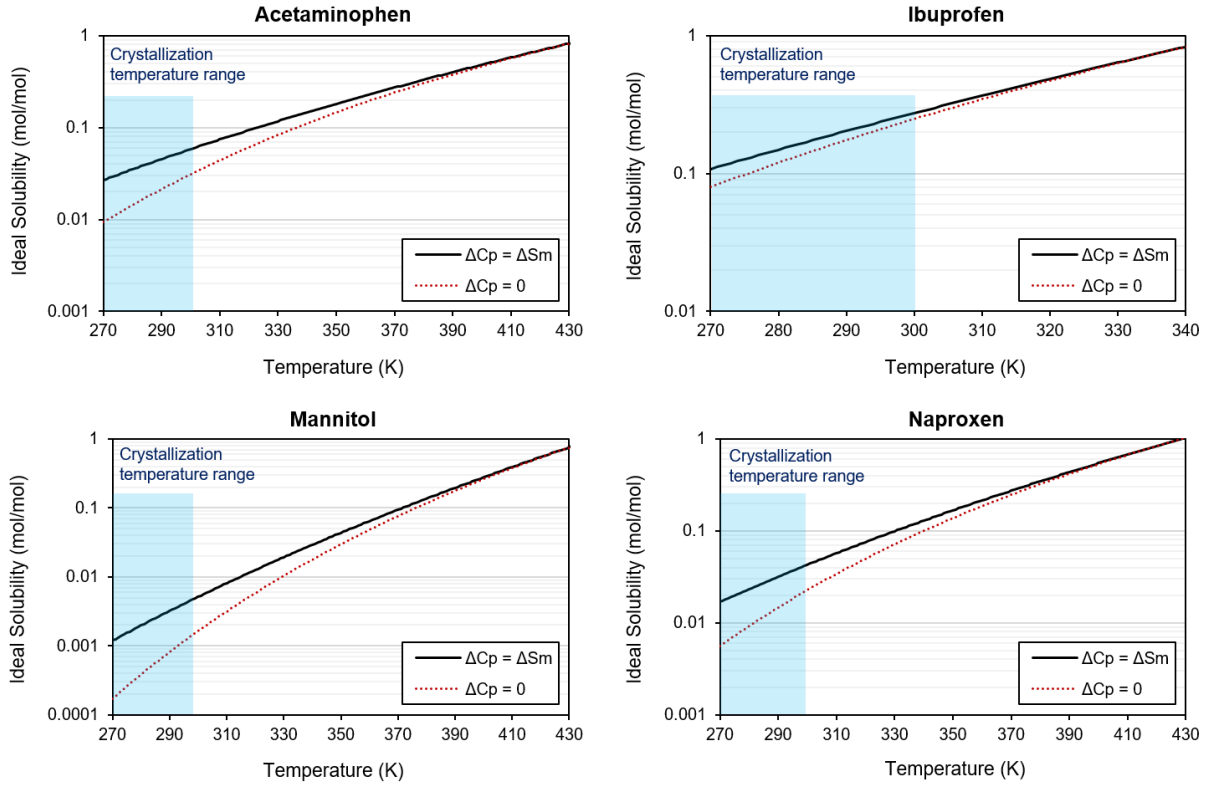

Figure S1. Calculated ideal solubility curves for acetaminophen, ibuprofen, mannitol, and naproxen.

The methods presented in the manuscript use the generalized solubility expression as means to calculate the activity coefficients for a given solute concentration. Consequently, they will be sensitive to errors in the estimated differential heat capacity. However, because the calculated supersaturation is a function of a ratio of activity coefficients, this parameter is not as sensitive to variations in ΔC_p as the ideal solubility.

The mole fraction and activity dependent (MFAD) supersaturation expression takes the form of:

$$\sigma = \ln\left(\frac{x\gamma}{x^{sat}\gamma^{sat}}\right) \tag{S3}$$

The product $x^{sat}\gamma^{sat}$ can be obtained from the generalized solubility expression (equation S1), while the product $x\gamma$ will express the solute's activity at supersaturated conditions. As explained in the manuscript, by assuming that the activity coefficient is a strong function of concentration and a weak function of temperature, γ can be approximated from the solubility curve using an effective temperature, T_e:

$$x = \frac{1}{\gamma} exp\left[\frac{\Delta H_{tp}}{R}\left(\frac{1}{T_{tp}} - \frac{1}{T_e}\right) - \frac{\Delta C_p}{R}\left(ln\frac{T_{tp}}{T_e} - \frac{T_{tp}}{T_e} + 1\right)\right] \tag{S4}$$

Incorporating equations S1 and S4 into equation S3, and approximating the triple point temperature and enthalpy of change as the melting point and the enthalpy of fusion, the MFAD supersaturation expression becomes a function of the crystallization temperature and the solute's effective temperature. Indirectly, these values express a concentration difference between the supersaturated and the saturated states, as both temperatures are tied to a given concentration through the solubility curve. The MFAD expression reduces to:

$$\sigma = \frac{\Delta H_m}{R}\left(\frac{1}{T} - \frac{1}{T_e}\right) + \frac{\Delta C_p}{R}\left(ln\frac{T_e}{T} + \frac{T_m}{T_e} - \frac{T_m}{T}\right) \tag{S5}$$

This expression can now be used to estimate the error propagations coming from a poor estimation of the heat capacity term. For the four compounds in Table S1, and assuming that $T_e = T + 10K$, the differences between the van't Hoff approach and the entropy of fusion approximation are reported in Figure S2.

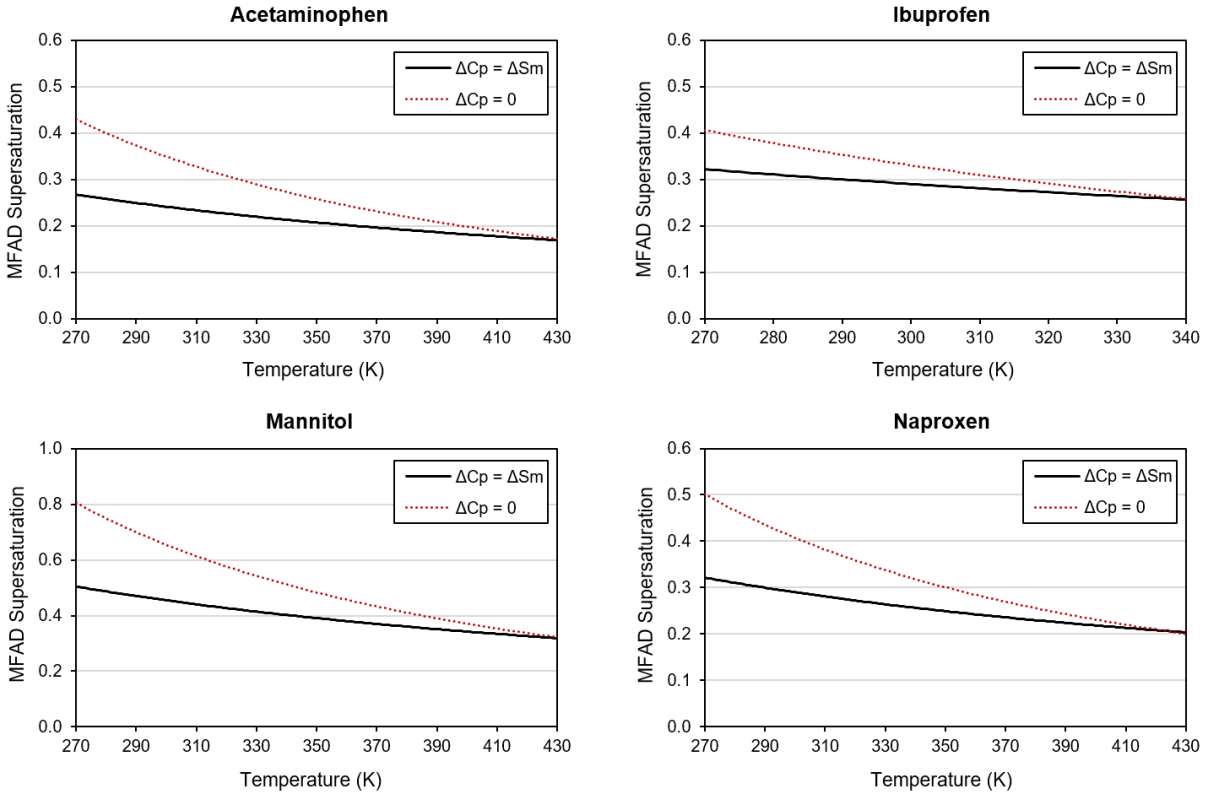

Figure S2. Calculated supersaturation trends for acetaminophen, ibuprofen, mannitol, and naproxen, when $T_e = T + 10K$.

For the four pharmaceuticals, it is apparent from Figures S1 and S2 that the errors in calculated supersaturation are smaller than those in the calculated ideal solubilities. Using equation S6 to estimate the errors in supersaturation, the largest differences between the two methods are 60.7%, 26.3% 59.6%, and 55.9% for acetaminophen, ibuprofen, mannitol, and naproxen, respectively. These differences are consistent regardless of the chosen effective temperature, with the estimated errors decreasing by only 1 - 3% when the same study is conducted for $T_e = T + 20K$. If these compounds were crystallized at 298 K, neglecting the heat capacity term would lead to supersaturation estimation errors between 18.5% and 50.8%.

$$\varepsilon = \frac{\left|\sigma_{\Delta C_p=0} - \sigma_{\Delta C_p=\Delta S_m}\right|}{\sigma_{\Delta C_p=\Delta S_m}} \quad \text{(S6)}$$

Note that these results assume that the entropy of fusion is a perfect estimation of the differential heat capacity. Unfortunately, this is difficult to prove experimentally. However, the alternative estimation methods available in the literature frequently give differential heat capacity values within 30% of the entropy of fusion for small molecule pharmaceuticals.[3,4] Assuming that the real differential heat capacities will fall within $\Delta S_m \pm 30\%$, we can estimate the expected confidence intervals for our supersaturation estimations. For the four pharmaceuticals in Table S1, these results are given in Figure S3.

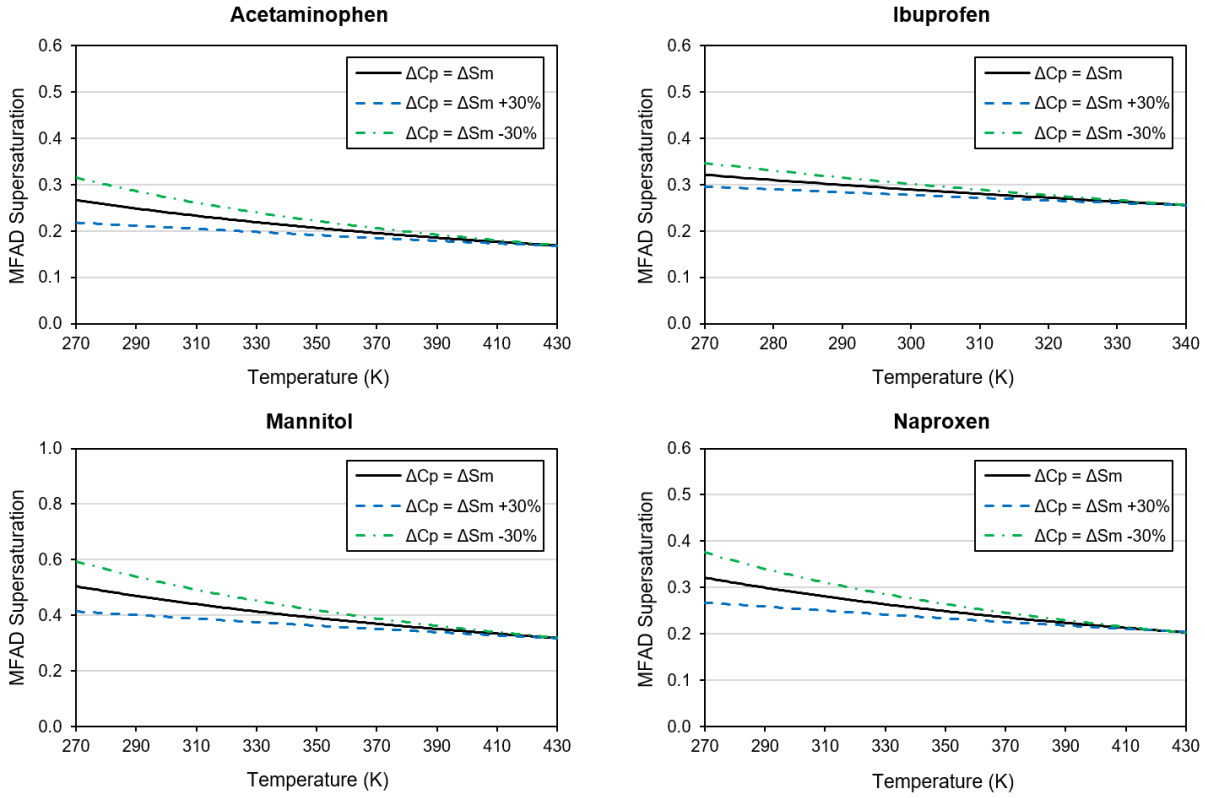

Figure S3. Calculated confidence intervals for estimating MFAD supersaturations of four pharmaceuticals when $T_e = T + 10K$.

For the crystallization range of interest, $270\,K < T < 300K$, the confidence intervals for the estimated MFAD supersaturations varied between 14.9% and 18.2% for acetaminophen, between 5.3% and 7.9% for ibuprofen, between 14.6% and 17.9% for mannitol, and between 13.6% and 16.8% for naproxen.

For the two compounds investigated in the manuscript, LAM gave confidence intervals between 22.1% to 25.8%, and API between 15.6% and 18.9%. As discussed in the manuscript, LAM presents a worst-case scenario because the enthalpy of fusion and, thus, the calculated entropy of fusion, are likely overpredicted because of solute decomposition near the melting point. Despite the expected large errors in estimating the MFAD supersaturation for this compound, the conclusions of this work are unaffected. This can be seen in Figure S4, where the manuscript's Figure 1 is reproduced within 30% confidence for the estimated differential heat capacity.

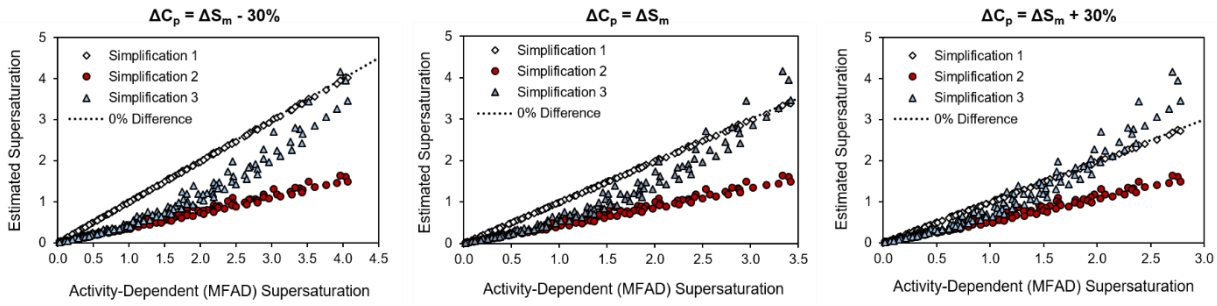

Figure S4. Comparison of simplified supersaturation with MFAD supersaturation calculations for LAM system, including error propagations from the estimation of differential heat capacities within a 30% confidence.

Note that, because the activity coefficients at supersaturation will always exceed those at equilibrium, simplification 2 (assuming the activity coefficient fraction is unity) will always underpredict supersaturation. Simplification 3 is consistently unreliable, as whether it gives a good prediction or not depends highly on the investigated system. Similar trends are found for API, as seen in Figure S5.

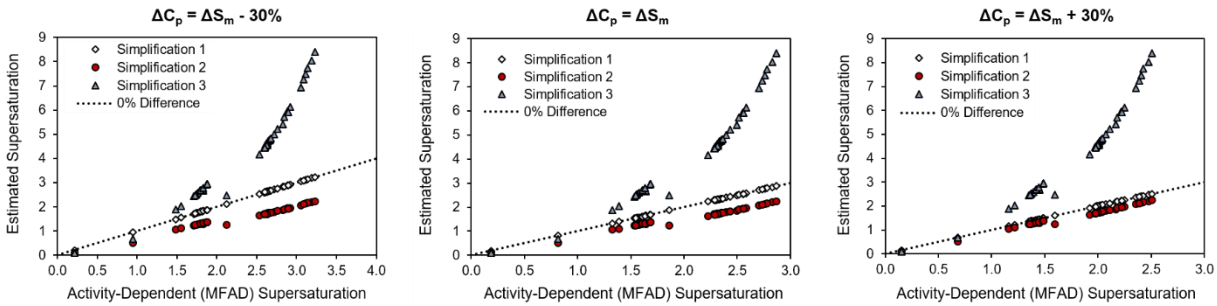

Figure S5. Comparison of simplified supersaturation with MFAD supersaturation calculations for API system, including error propagations from the estimation of differential heat capacities within a 30% confidence.

In this case, while the trends are consistent within the confidence intervals, the errors associated with simplification 2 are highly sensitive to error propagation from a poor estimation of differential heat capacities. Overall, the conclusions from this work are consistent regardless of error propagation from the chosen approximations: the MFAD supersaturation expression is the most accurate approach for the estimation of supersaturations, as it accounts for the solute activities instead of assuming an ideal system. For systems where the activity coefficient ratio is assumed to be unity, supersaturation should be expressed as a logarithmic function of the system concentrations, preferrably expressed as mole fractions. Other approaches are inconsistent and can lead to significant errors in the estimation of crystallization kinetics.

References

1 S. H. Yalkowsky and M. Wu, *J. Pharm. Sci.*, 2010, **99**, 1100–1106.

2 A. S. Myerson, *Handbook of Industrial Crystallization*, Butterworth-Heinemann, 2nd ed., 2002.

3 H. Neau, Steven, V. Bhandarkar, Satej and W. Hellmuth, Eckhard, *Pharm. Res.*, 1997, 14, 601–605.

4 G. D. Pappa, E. C. Voutsas, K. Magoulas and D. P. Tassios, *Ind. Eng. Chem. Res.*, 2005, **44**, 3799–3806.