

In silico modelling in an industrial crystallisation process

George Taylor – APMF 2017

ADDOPt - Advanced Digital Design of Pharmaceutical Therapeutics



Funded by **AMSCI** (Advanced Manufacturing Supply Chain Initiative)

- The Medicines Manufacturing Industry Partnership (**MMIP**) brought together key industry stakeholders to develop and shape the bid ultimately leading to successful funding from UK Government via **AMSCI**
- The project aim to **integrate** leading scientists, university research groups, Industry and SMEs with the UK's advanced **E-infrastructure** research facilities to **optimise pharmaceutical manufacturing processes**.

Explore & exploit 'in-silico' methodologies

- Pre-competitive industry data sharing
- Advanced materials and process modelling
- Advanced Predictive Analytics & Informatics
- Super computing
- Integrated software & control systems

£20.5 million budget

£12.2 million government
£8.3 million industry

2015

2016

2017

2018

2019

PARTNERS:

Research organisations



SMEs: Process Systems Enterprise (lead)



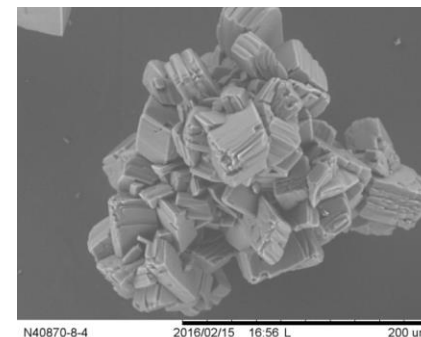
Primes:



GSK will not receive funding
and only providing 'in-kind'
contribution to the consortium

Can we model a crystallisation process to:

- Predict scale up effects?
- Construct design space?
- Optimise conditions to meet a specification?



Available Software – gPROMS (gCRYSTAL)

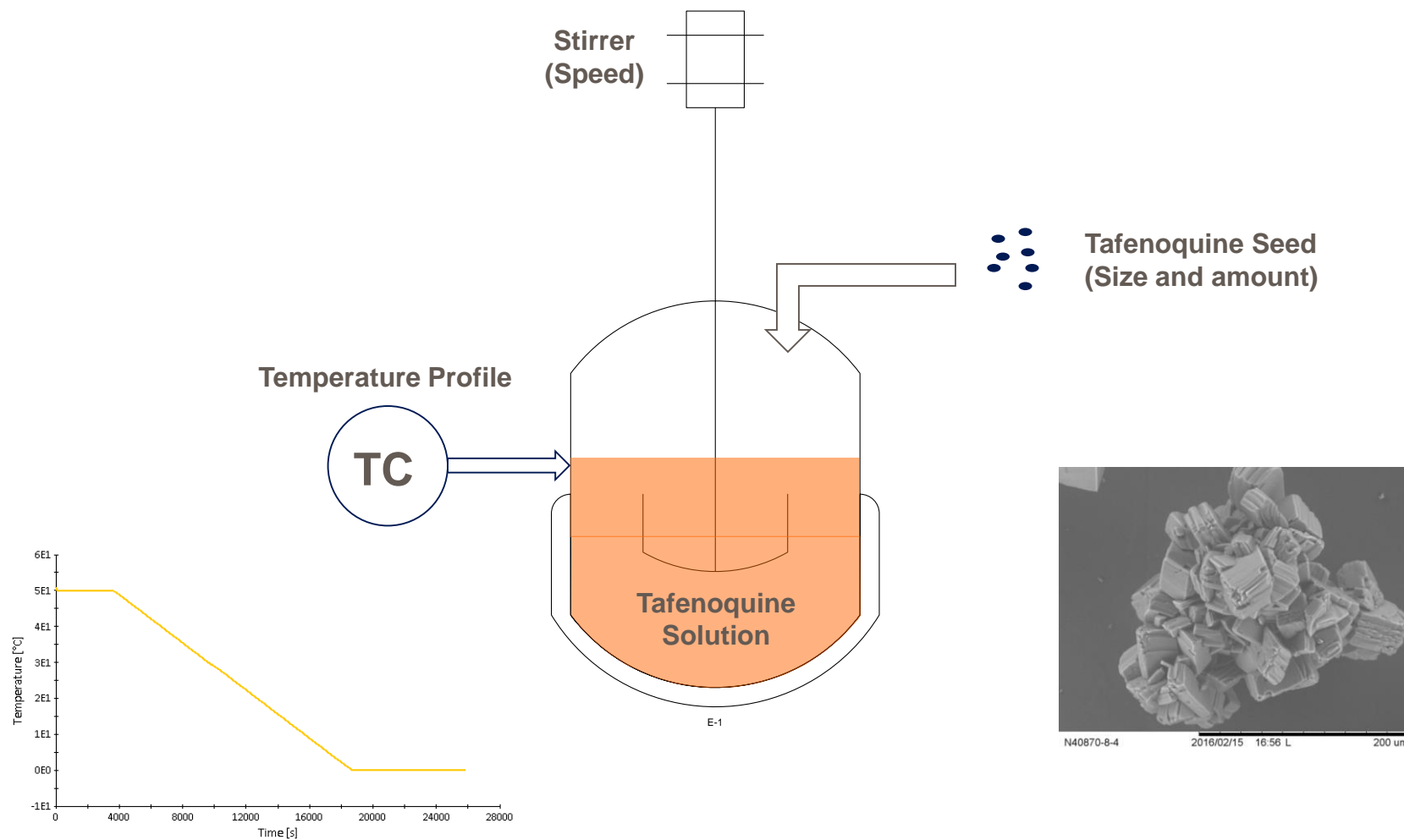
– Crystallisation modelling software from first principals. A mechanistic model.

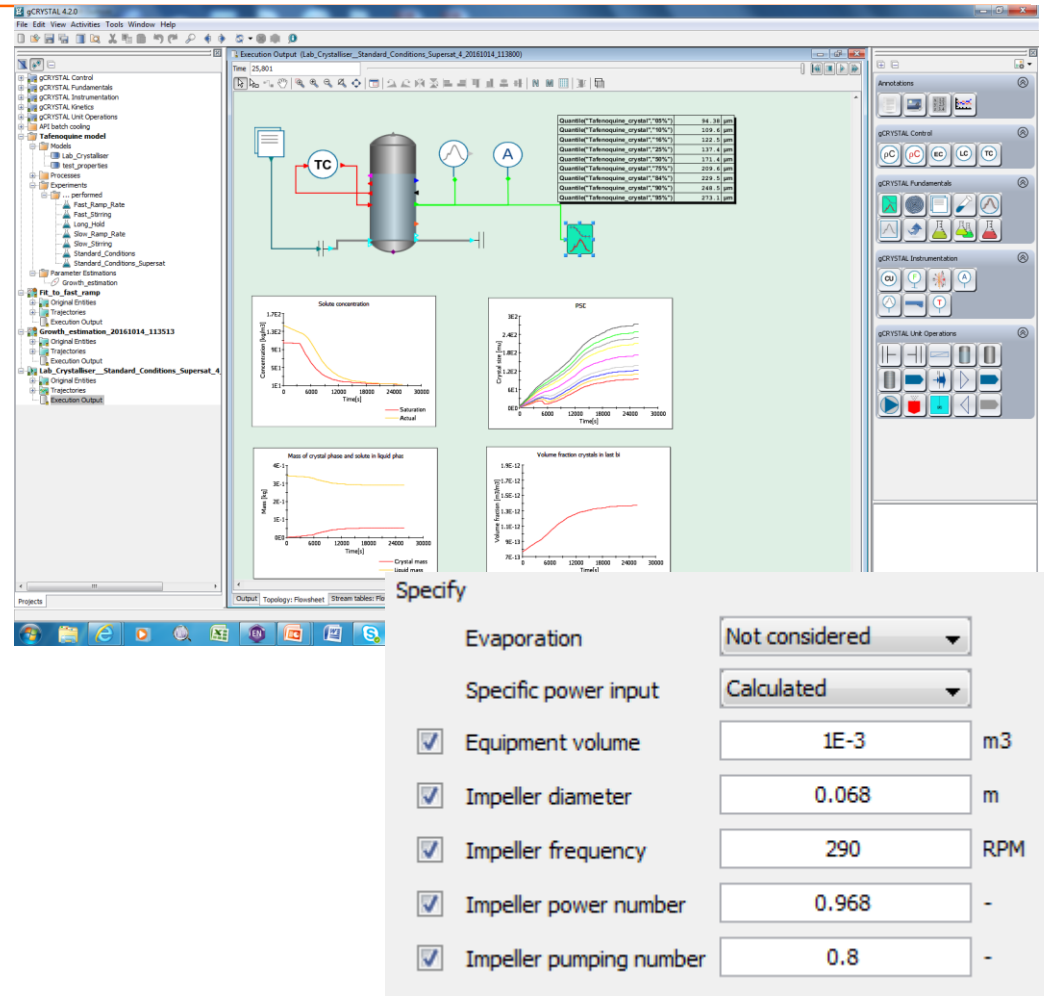
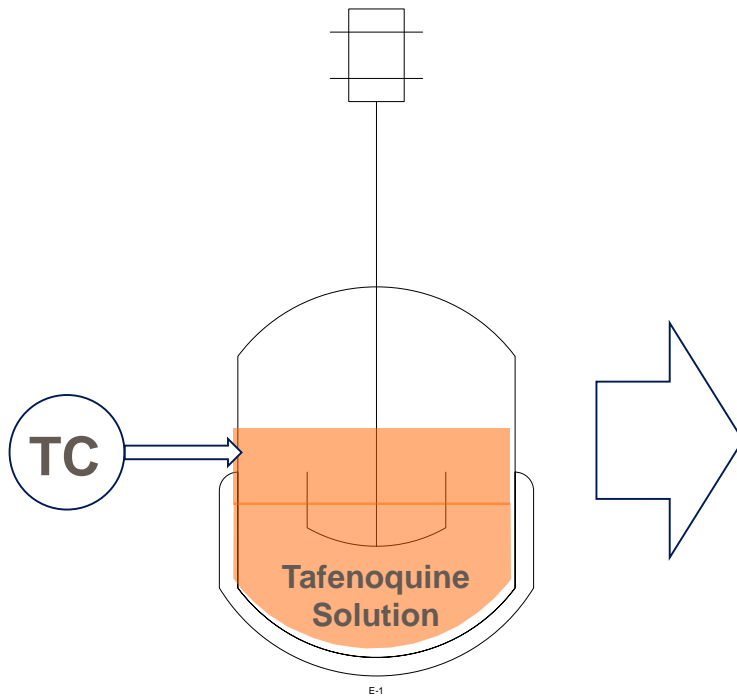
Proposal – Tafenoquine Case Study

- To use Tafenoquine crystallisation to construct a mechanistic model
- Understand the data requirement to construct the model
- Evaluation of the model based on how it predicts PSD
- Evaluate how well it predicted the scale up of the process

The Process

Batch Reactor





Gathering Data

A Thermodynamic Knowledge of the System:

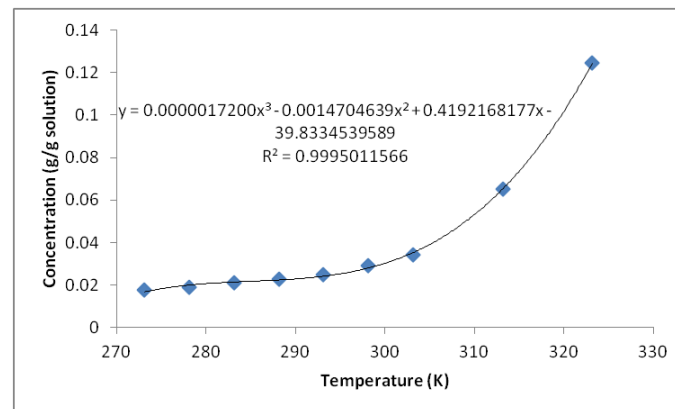
- Solubility Curve
- Equipment Specifications (Agitator dimensions, vessel size etc.)
- Seed PSD
- Crystal density (from Single crystal structure)

A Kinetic knowledge of the System:

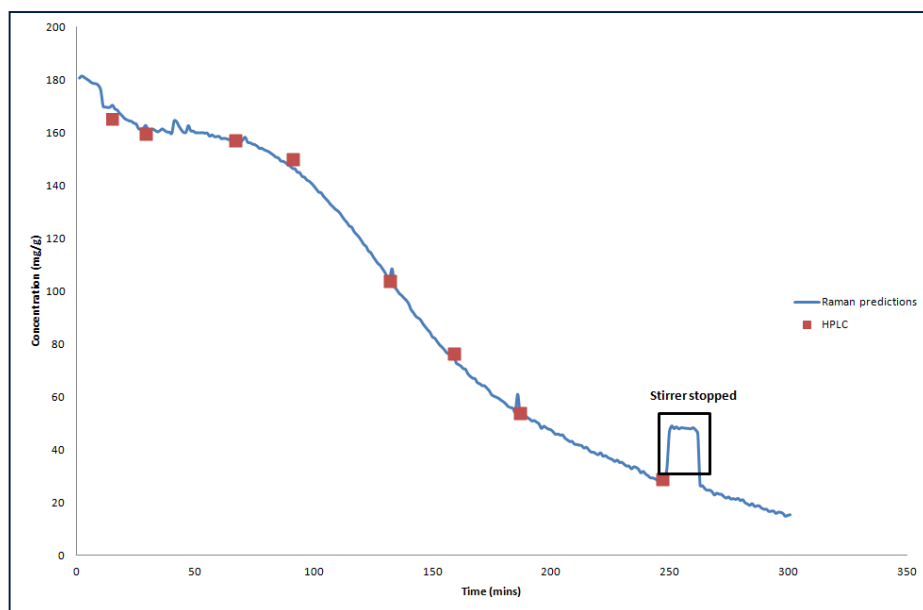
- Desupersaturation data
 - Obtained using insitu Raman or IR
- Idea of the important crystallisation mechanisms
 - PVM and FBRM data can give an insight into this

Specify

Evaporation	<input type="text" value="Not considered"/>	
Specific power input	<input type="text" value="Calculated"/>	
<input checked="" type="checkbox"/> Equipment volume	<input type="text" value="1E-3"/>	m3
<input checked="" type="checkbox"/> Impeller diameter	<input type="text" value="0.068"/>	m
<input checked="" type="checkbox"/> Impeller frequency	<input type="text" value="290"/>	RPM
<input checked="" type="checkbox"/> Impeller power number	<input type="text" value="0.968"/>	-
<input checked="" type="checkbox"/> Impeller pumping number	<input type="text" value="0.8"/>	-



- An insitu Raman probe was used to monitor concentration changes in the liquid phase
- A quantitative PLS model was calibrated to offline HPLC measurements
 - Gives an insight into desupersaturation of the system

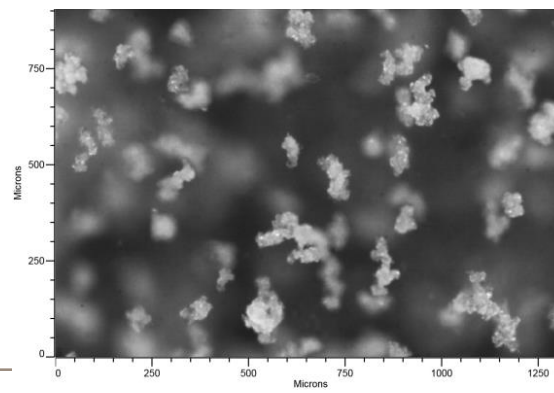
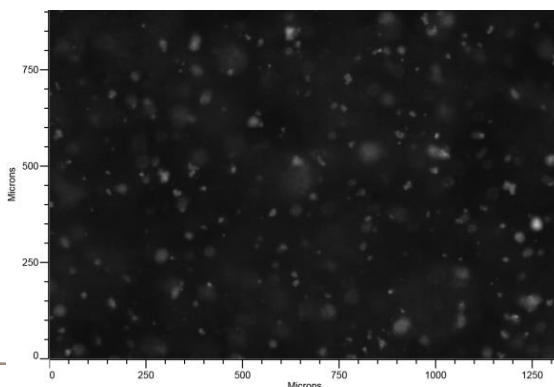
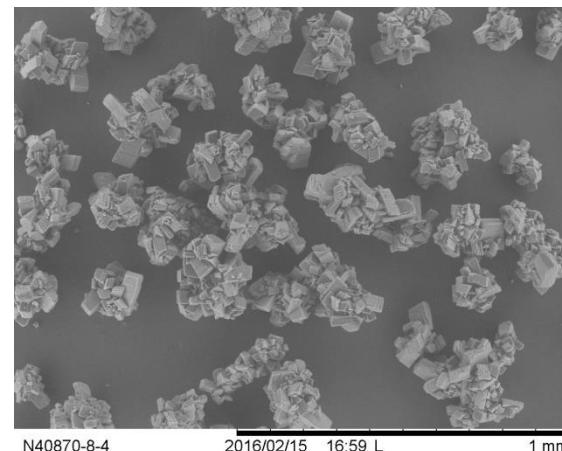


Conducted 4 experiments to be used in the model:

- Standard Conditions
- Long Hold
- Fast Stirrer
- Fast Ramp

Experiments performed in an CLR with FBRM and PVM probes to investigate crystallisation mechanisms:

- 1) Primary Nucleation
- 2) Secondary Nucleation
- 3) Growth
- 4) Agglomeration
- 5) Attrition



1) Primary Nucleation



2) Secondary Nucleation



3) Growth



4) Agglomeration



5) Attrition



1) Primary Nucleation

$$J_{prim} = \exp(\ln A_0) \exp\left(\frac{-16\pi(\alpha\sigma)^3 v_0^2}{3k^3 T^3 (\ln S)^2}\right)$$

2) Secondary Nucleation

$$J_{sec} = A_t \left[\exp(\ln k_s) \frac{D_{AB}}{d_m^4} \exp\left(-\pi \left(K \ln\left(\frac{c_c}{c^*}\right)\right)^2 \frac{1}{v \ln S_a}\right) \right]$$

3) Growth

$$G(L) = k_g \exp\left(\frac{-E_{A,g}}{RT}\right) \left[\frac{C_{int}(L) - C_{sat}}{\rho_{orgs}} \right]^g \quad \text{if } C_{bulk} - C_{sat} > 0$$

4) Agglomeration

$$\beta_{agg} = \left(\sqrt{\frac{8\pi\epsilon}{15v}} \cdot \bar{d}_{3,0}^3 \right) \cdot \frac{(A_{50}G)/(\epsilon\rho\bar{d}_{3,0}^2)}{1 + (A_{50}G)/(\epsilon\rho\bar{d}_{3,0}^2)}$$

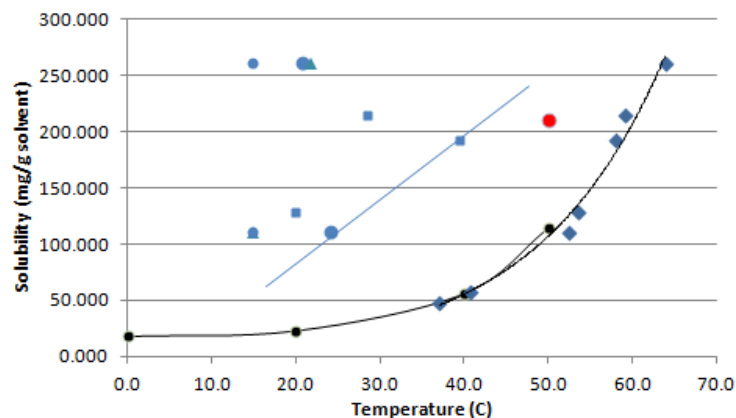
5) Attrition

$$J_{sec\,ci} = \exp(\ln k_{n-ci}) \sigma^{n_{ci}} \frac{N_Q}{N_P} k_v \rho_c \epsilon \int_{L_{min-ci}}^{\infty} n L^3 dL$$

Which Mechanisms to consider?

1) Primary Nucleation

Spontaneous crystallisation



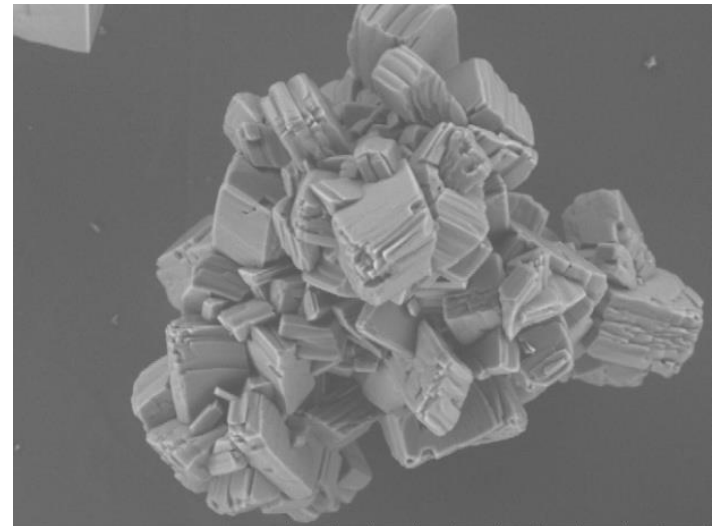
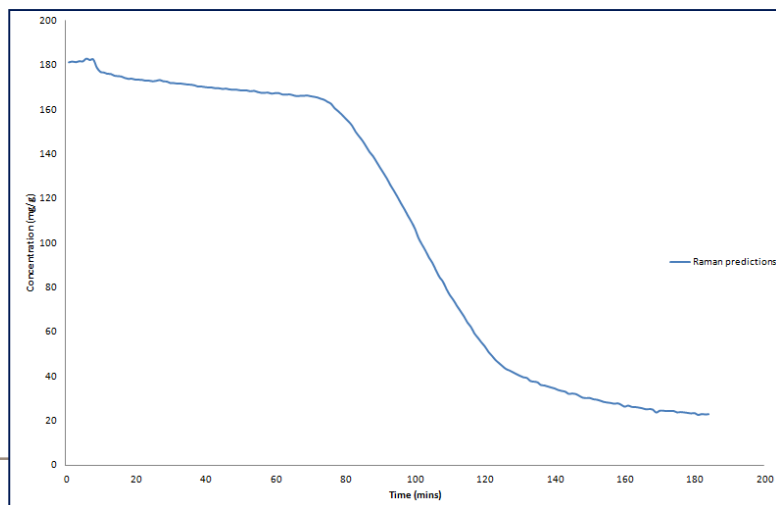
Not likely to occur at 50degC during a seeded crystallisation – can ignore!

Which Mechanisms to consider?

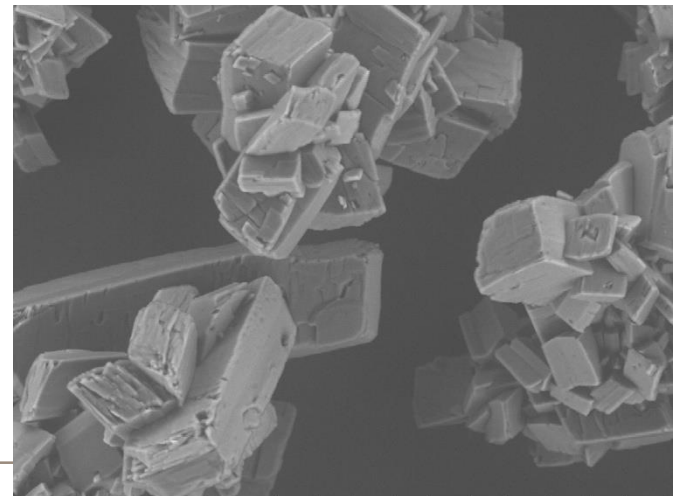
2) Secondary Nucleation

3) Growth

4) Agglomeration



N40870-8-4 2016/02/15 16:56 L 200 um

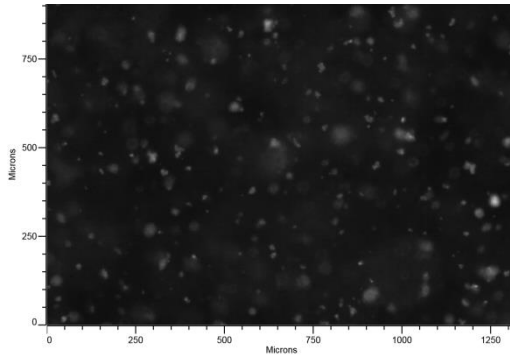


N40870-8-6 2016/02/15 17:10 L 200 um

5) Attrition

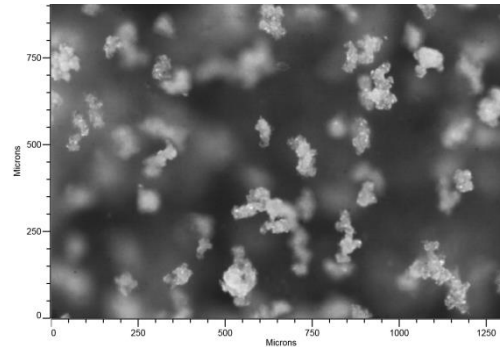
Extended stirring, FBRM and PVM data suggests that attrition is not occurring in the system

Can Ignore!



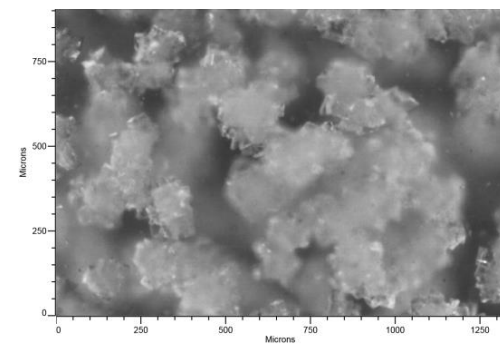
METTLER TOLEDO

T = 0



METTLER TOLEDO

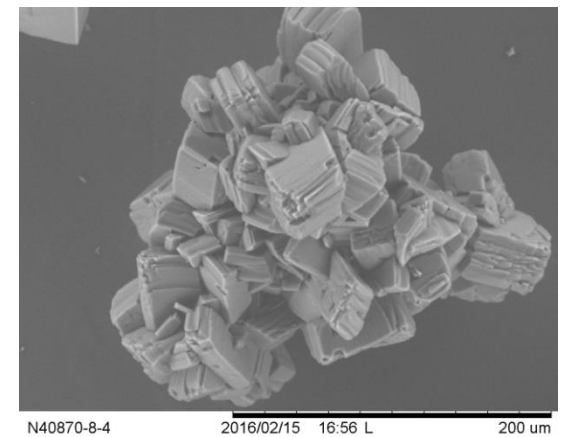
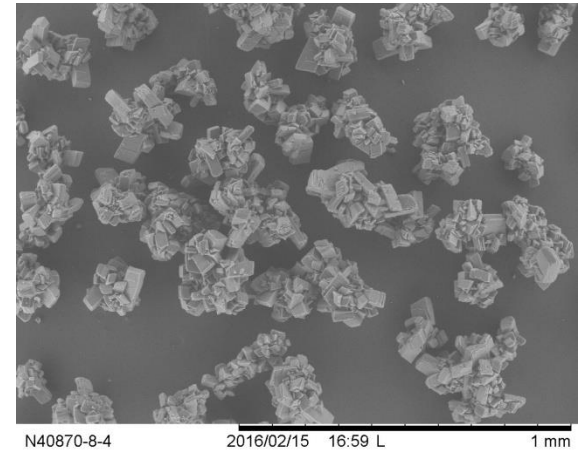
T = 1h



METTLER TOLEDO

T = 16h

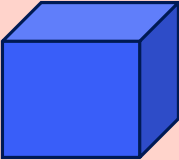
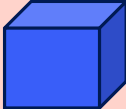

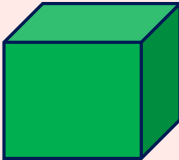
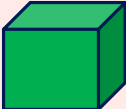
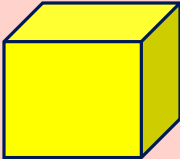
- 1) ~~Primary Nucleation~~
- 2) Secondary Nucleation
- 3) Growth
- 4) Agglomeration
- 5) ~~Attrition~~



Modelling

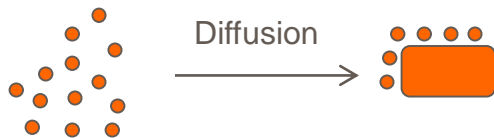
Incremental Parameter estimation

– Progressive addition of crystallisation mechanisms

Kinetic Parameter	Stage 1	Stage 2	Stage 3
Growth			
Agglomeration	Inactive		
Secondary Nucleation	Inactive	Inactive	

Mersmann two-step kinetics

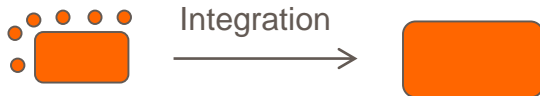
1) Mass transfer to the crystal surface



$$G(L) = k_d(L) \left[\frac{C_{bulk} - C_{int}(L)}{\rho_{orys}} \right]$$

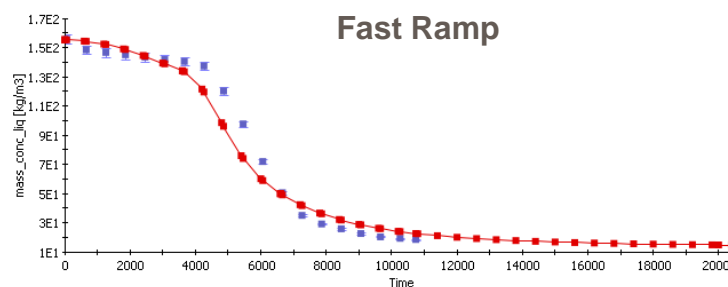
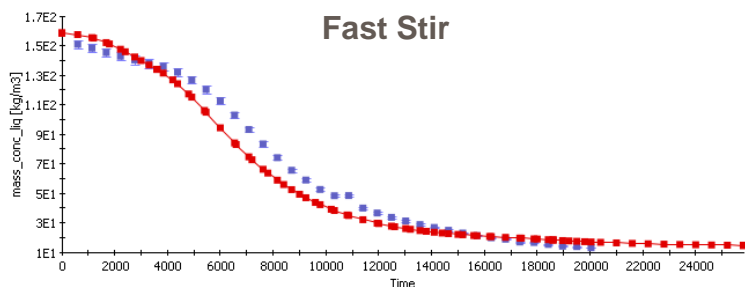
Calculated from the mass transfer in the system.
Does not require fitting

2) Integration into the crystal surface

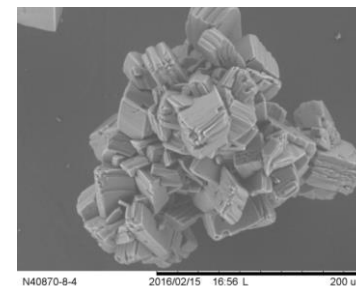
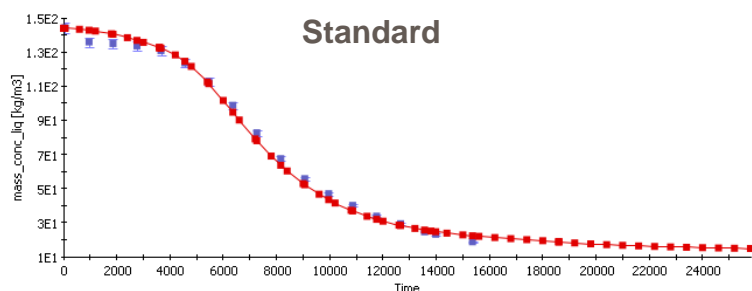


$$G(L) = k_g \exp\left(\frac{-E_{A,g}}{RT}\right) \left[\frac{C_{int}(L) - C_{sat}}{\rho_{orys}} \right]^g \quad \text{if } C_{bulk} - C_{sat} > 0$$

This model requires fitting of three parameters: Rate constant (k_g), Activation Energy (E_a) and order relative to supersaturation (g).



Good fit could NOT be obtained with Growth only model
Requires further mechanisms to achieve desired PSD



Exp	D50 (pred)	D50 (exp)
Std	11.4	170
Fast Ramp	11.7	140
Fast Stir	11.8	131

Growth and Agglomeration

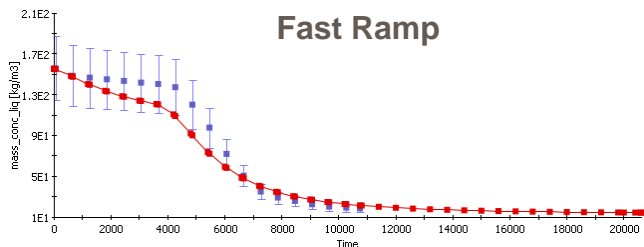
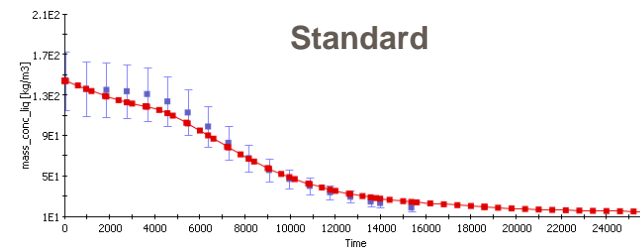
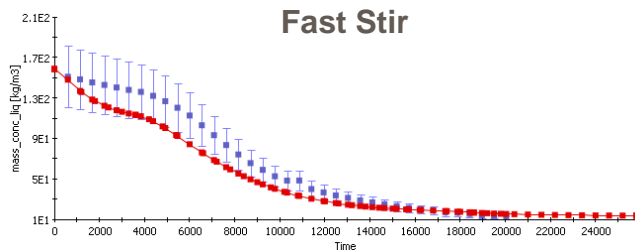
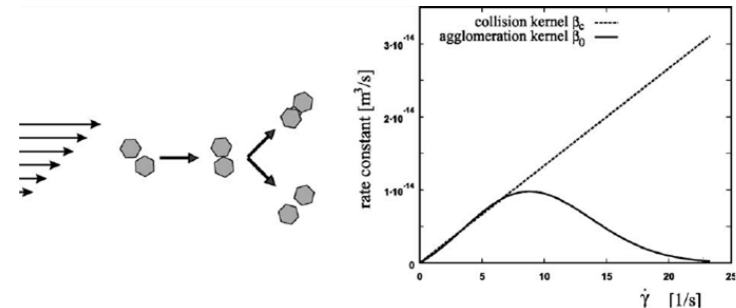


Need to include an agglomeration term to describe the system

- Modelled using Mumtaz kinetics which is based on collisions and growth rate

$$A_{50} = \frac{L \cdot \sigma^*}{M_{50}} \quad \beta_{agg} = \left(\sqrt{\frac{8\pi\epsilon}{15\nu}} \cdot \bar{d}_{3.0}^3 \right) \cdot \frac{(A_{50}G)/(\epsilon\rho\bar{d}_{3.0}^2)}{1 + (A_{50}G)/(\epsilon\rho\bar{d}_{3.0}^2)}$$

$$M = \frac{\text{strength of the neck}}{\text{hydrodynamic force/area of the neck}} = \frac{L \cdot \sigma^* \cdot G}{\epsilon \cdot \rho \cdot \bar{d}_{3.0}^2}$$



Exp	D50 (pred)	D50 (exp)
Std	160	170
Fast Ramp	173	140
Fast Stirr	97	131

Growth, Agglomeration and Secondary Nucleation



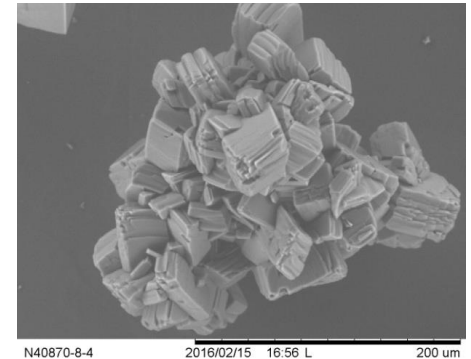
Activated secondary nucleation

May be possible to describe the system using a combination of growth, agglomeration and secondary nucleation.

This gives the system many degrees of freedom (7 variables)

- Likely to be able to describe most systems
- Risk of become empirical model rather than mechanistic

$$G(L) = k_g \exp\left(\frac{-E_{A,g}}{RT}\right) \left[\frac{C_{int}(L) - C_{sat}}{\rho_{orgs}} \right]^g \quad \text{if } C_{bulk} - C_{sat} > 0$$

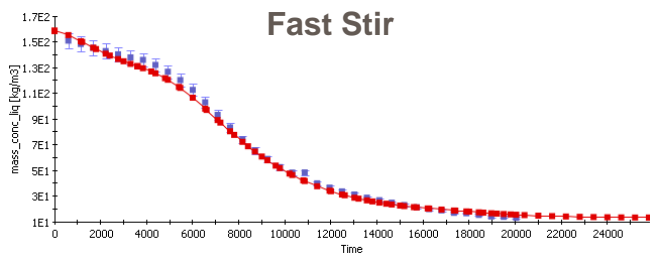
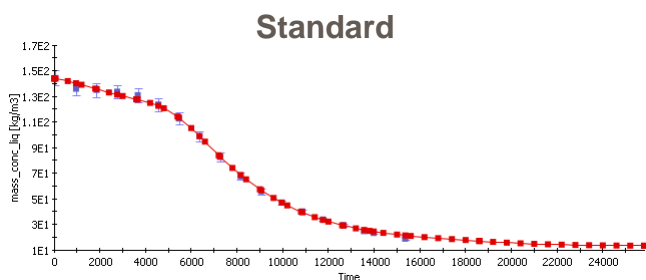
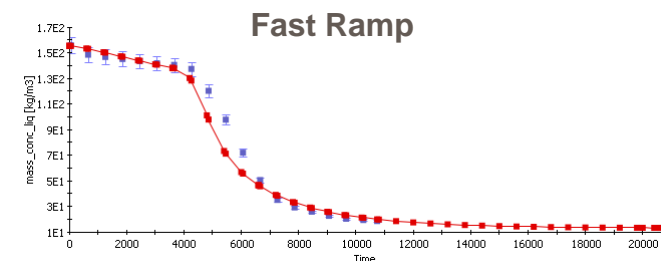


$$\beta_{agg} = \left(\sqrt{\frac{8\pi\epsilon}{15\nu}} \cdot \bar{d}_{3,0}^3 \right) \cdot \frac{(A_{50}G)/(\epsilon\rho\bar{d}_{3,0}^2)}{1 + (A_{50}G)/(\epsilon\rho\bar{d}_{3,0}^2)} \quad J_{sec} = A_t \left[\exp(\ln k_s) \frac{D_{AB}}{d_m^4} \exp\left(-\pi \left(K \ln\left(\frac{c_c}{c^*}\right) \right)^2 \frac{1}{\nu \ln S_a} \right) \right]$$

Growth, Agglomeration and Secondary Nucleation



Good fitting obtained



Constant	Value
A_{50}	0.03571
K_g	2.786E-7
g	0.793
E_a	5663
$\text{Ln}k_s$	-27.5128
K	0.409
v	1

Exp	D50 (pred)	D50 (exp)
Std	173	170
Fast Ramp	130	140
Fast Stirr	133	131

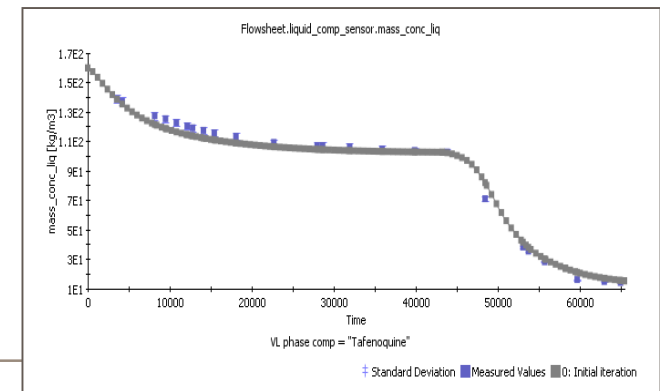
How well did it predict CPPs



Parameter	Predicted D50	Experimental D50
Std Conditions	173	170
Seed size	175	185
Seed Loading	173	210
Stirrer speed	133	131
Seed hold	234	210
Seed temp (low)	182	187
Seed temp (high)	169	170
Scale up (630L)	164	158-170

Model predicts certain parameters well. It predicts the effects of seed size and mass poorly.

Model is a good fit to plant scale data. Further validation would be to simulate more scale up variables (eg power number, vessel geometry etc.)



Pros

Predicted certain parameters very well

Did not require any data we wouldn't normally obtain

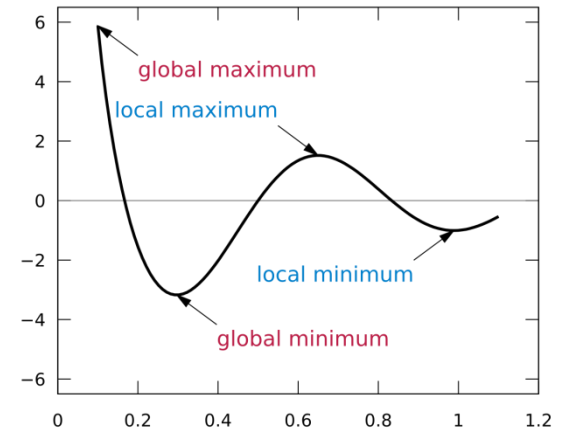
Valid across scales

Cons

Parameter estimation is a derivative based solver

- Easily gets stuck in local minima
- Requires initial parameter estimation to be close to the true values.
- Multiple parameter estimations can give similar outputs

Can be poor at predicting parameters the model isn't fitted to



Constant	Value
A_{50}	0.03571
K_g	2.786E-7
g	0.793
E_a	5663
$\text{Ln}k_s$	-27.5128
K	0.409
v	1

A useful tool for crystallisation prediction, although it will require multiple iterations to truly model the system.

Potential applications:

- DoE design/*in silico* DoE
- Identifying CPPs
- Predictive tool to supply material for early formulation development
- Scale up prediction
- Failure mode prediction

Has the potential to save many laboratory hours if implemented during early development!

Acknowledgements



I'd like to thank:

- Laura Palmer
- Tafenoquine project team
- Chrismono Himawan
- PSE, in particular Niall Mitchell

Thank you for Listening