



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Process modelling and simulation for continuous pharmaceutical manufacturing of ibuprofen

### Citation for published version:

Jolliffe, HG & Gerogiorgis, D 2015, 'Process modelling and simulation for continuous pharmaceutical manufacturing of ibuprofen', *Chemical Engineering Research and Design*, vol. 97, pp. 175-191.  
<https://doi.org/10.1016/j.cherd.2014.12.005>

### Digital Object Identifier (DOI):

[10.1016/j.cherd.2014.12.005](https://doi.org/10.1016/j.cherd.2014.12.005)

### Link:

[Link to publication record in Edinburgh Research Explorer](#)

### Document Version:

Peer reviewed version

### Published In:

Chemical Engineering Research and Design

### General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Process Modelling and Simulation for Continuous Pharmaceutical Manufacturing of Ibuprofen

Hikaru G. Jolliffe and Dimitrios I. Gerogiorgis\*

*Institute for Materials and Processes (IMP), School of Engineering, University of Edinburgh,  
The King's Buildings, Edinburgh, EH9 3JL, United Kingdom*

*\*Corresponding author: D.Gerogiorgis@ed.ac.uk (+44 131 6517072)*

## ABSTRACT

Pharmaceutical corporations face rapidly rising process research and development (R&D) as well as production costs due to globalised competition. Batch production processes are dominant in the pharmaceutical industry and have multiple advantages, including equipment flexibility, high-fidelity quality control and the ability to recall specific batches; they however suffer disadvantages such as limited heat transfer and mixing scalability and low operational asset efficiency. Continuous Pharmaceutical Manufacturing (CPM) has a documented potential to reduce cost because continuous production techniques can be easier to scale up and can be designed to be more efficient in terms of both solvent and energy use: therefore, it is both timely and important to explore the expanding feasibility limits of this emerging technology. The literature has been extensively surveyed in order to identify a series of candidate Active Pharmaceutical Ingredients (API) for flowsheet synthesis, process modelling and mass balance simulation toward rapid assessment of CPM potential. Ibuprofen [2-(4-isobutylphenyl)propanoic acid], the widely used non-steroidal anti-inflammatory drug (NSAID), has emerged as an ideal CPM candidate because it is in high global demand and can generate significant profit margins. The flowsheet is based on a published organic synthesis pathway and produces 50 kg of ibuprofen annually using three plug flow reactors (PFRs) in series, followed by a final separation for purification. Kinetic and thermodynamic parameter estimation modelling has been employed in order to compute essential data for design, and all PFR reactors have been designed based on reported conversions of feed and intermediate organic molecules in the respective pharmaceutical synthesis reactions. Theoretically computed reactor designs are in good agreement with experimental prototypes constructed for the same organic synthesis, as well as with previously reported CPM systems. The developed continuous final separation performs very well in accordance with green chemistry principles, and with relatively low environmental impact (an E-factor of 25.4).

## 1 INTRODUCTION

Batch manufacturing processes feature advantages including equipment usage versatility, flexible production planning and scheduling, and a wide range of attainable products, and have dominated the pharmaceutical industry for a number of decades; they still are considerably more preferable due to regulatory and licensing considerations and the option to quickly recall specific batches of products (Plumb, 2005). Nevertheless, they have several disadvantages: new processes are often difficult to scale up to production level due to poor heat transfer and mixing (potentially resulting in unacceptable product quality) and efficiency can be very low with high volumes of unrecovered solvent (Anderson, 2012; Gernaey et al., 2012). Batch production plants also require significant intermediate storage capacity between process stages, resulting in large inventories of feedstock organic chemicals and sensitive intermediates.

In recent years, pharmaceutical firms have faced mounting pressure to reduce ever-increasing R&D as well as production costs (Fig. 1). The commercialisation of new drugs can require up to \$1.8 billion in total per product, a high cost which has been steadily increasing (Fig. 2). Competition from generic drug manufacturers also increases the pressure on pharmaceutical firms (Behr et al., 2004). Taking into account clinical approval and patent duration, by the time the product reaches the market half of its patent life may have expired; once this occurs, the original developer may lose up to 90% of market share to generic manufacturers (Plumb, 2005).

Continuous Pharmaceutical Manufacturing (CPM) offers many advantages: lower costs, reduced waste, decreased time-to-market for new drugs, continuous flow reactors can deliver significantly higher yields, and solvent and energy waste can be decreased by up to 90% by steady-state operation (Kockmann et al., 2008; Roberge et al., 2008; Anderson, 2012). For identical production rates, continuous systems require considerably smaller equipment volumes than batch processes (Ashe and Lazowski, 2012). Lower solvent requirement and more efficient heat transfer during exothermic reactions both lead to safer operation (LaPorte and Wang, 2007; Kockmann et al., 2008; Roberge et al., 2008), therefore CPM can offer cost savings to pharmaceutical corporations, as it does in commodity and fine chemicals industries (Schaber et al., 2011). Labour cost can also be reduced, due to the limited need for human intervention in material transport and quality control (Plumb, 2005).

Experimental demonstration and process validation at pilot plant scale are key to showcasing CPM benefits: as this is laborious, costly and time-consuming, flowsheet synthesis and process modelling can be used to investigate the feasibility and viability of candidate CPM processes (Douglas, 1988; Gerogiorgis and Barton, 2009). Minimal investment is required, and numerous process options and configurations can be rapidly evaluated by means of suitable mathematical models and simulations, especially when pilot plants are expensive and know-how is limited (Gerogiorgis and Ydstie, 2005).

This work presents an adapted flowsheet and a steady-state process model which has been developed for the continuous synthesis and purification of an Active Pharmaceutical Ingredient (API) identified as a promising candidate for this new production paradigm. Reactor design has been performed based on published kinetics of the continuous organic synthesis pathway, and a final separation stage for product purification has been designed, with multiple solvents evaluated with respect to API recovery and green chemistry performance. Process modelling and simulation have thus been successfully applied to illustrate the benefits and feasibility of CPM for ibuprofen.

The paper is organized as follows: the evaluation of candidate APIs is presented, and following this the static model flowsheet and mass balances are shown. Published data is analysed to determine reaction orders and kinetic constants, which are then used to design continuous plug flow reactors. A summary of material properties is given next, followed by heat transfer requirements and final separation design. Finally, mass balance and unit operation design results are discussed in relation to the technical feasibility of the process and the emerging field of CPM process R&D. Detailed kinetic, thermodynamic and physicochemical property prediction models are presented in the Appendix.

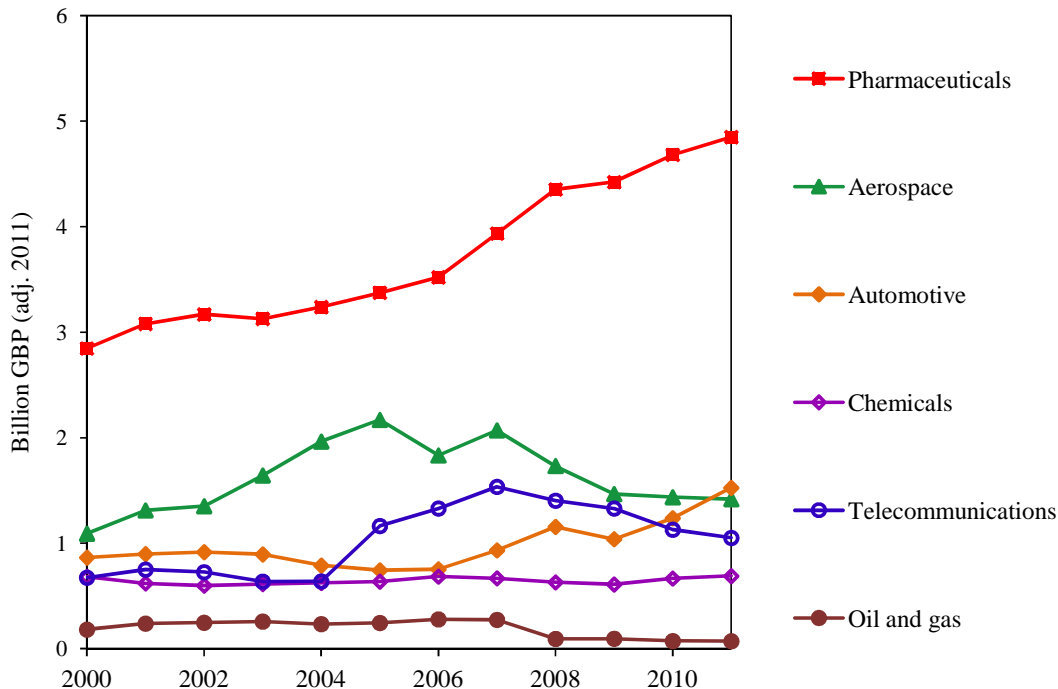


Fig. 1. R&D expenditure of the pharmaceutical industry vs. other UK sectors (Office for National Statistics, 2011).

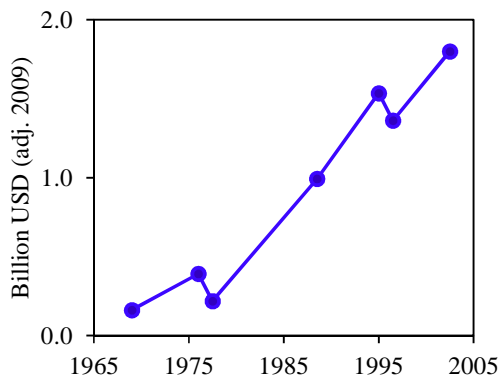


Fig. 2. Cost of new pharmaceutical product R&D and commercialisation (Morgan et al., 2011).

## 2 API SELECTION

### 2.1 Identification of candidate APIs

An API is the biologically active chemical species of a pharmaceutical drug (the remainder being composed of inert material, the excipient). Although there is a multitude of APIs in commercially available formulations today, not all are amenable to continuous manufacturing. To identify suitable CPM candidates, the literature has been extensively surveyed in order to identify published organic flow synthesis pathways.

The following candidates have been identified: artemisinin, casein-kinase 1 (CK1) inhibitor, vitamin D<sub>3</sub>, diphenhydramine,  $\delta$ -opioid receptor (DOR) agonists, (R)-flurbiprofen, ibuprofen, imatinib, taurine- and glycine-conjugated ursodeoxycholic acid (TUDCA and GUDCA), and reflux inhibitor AZD6906.

*Artemisinin* has well-known anti-malarial properties and is extensively used in traditional Chinese medicine. First extracted from the sweet wormwood plant *Artemisia annua* in 1972, artemisinin and its derivatives are the fastest-acting antimalarial drugs available today (Tu, 2011). The current

production method relies on plant cultivation and subsequent extraction (White, 2008); the continuous production of artemisinin from the waste product of this extraction process has been recently demonstrated (Kopetzki et al., 2013). This method converts dihydroartemisinic acid (DHAA) into artemisinin, and construction of a scaled-up system with an annual capacity of one metric tonne is currently under way (Kopetzki et al., 2013).

**Casein-kinase I (CKI) inhibitor** is potent in the regulation of mammal circadian rhythms, and research indicates it is promising for the regulation of sleep disorders and mood disorders (Lee et al., 2009). Recently, Sanofi-Aventis developed a class of CKI inhibitors, and certain ones have been synthesised continuously (Almario et al., 2009).

**Vitamin D<sub>3</sub>** has extreme physiological importance, as this molecule and its analogues are used to treat osteoporosis and renal failure (Posner and Kahraman, 2003). A continuous route for the synthesis of activated vitamin D<sub>3</sub> and some of its analogues has been reported (Fuse et al., 2012).

**Diphenhydramine**, discovered in 1942, is a first-generation antihistamine first marketed as Benadryl<sup>®</sup> (Ravina and Kubinyi, 2011). More effective than many newer antihistamines at treating allergic symptoms, it also has sedative properties and is used in sleep aids (Richardson et al., 2002). Global demand of diphenhydramine is high (at over 100 metric tonnes a year), and recent work has demonstrated its continuous flow synthesis using molten ionic salts (Snead and Jamison, 2013).

**DOR agonists** have been shown to have analgaesic properties without the side-effects associated with opioids; recently, researchers have presented a continuous synthesis method for the production a DOR (developed by Astra Zeneca) and some of its analogues (Wei et al., 2000; Qian et al., 2010).

**Flurbiprofen** is a non-steroidal anti-inflammatory drug (NSAID) used to treat several pathological conditions, including osteoarthritis. The R-enantiomer of flurbiprofen, R-flurbiprofen, has undergone clinical trials for a variety of conditions (Wechter et al., 2000; Quann et al., 2007), and a continuous synthesis pathway and isolation method of R-flurbiprofen has been reported (Tambornini et al., 2012).

**Ibuprofen** [2-(4-isobutylphenyl)propanoic acid] is a widely used NSAID which treats pain, fever, and inflammation, and is listed as essential by the World Health Organisation. The original 6-step synthesis devised in the 1960s has been replaced by a more efficient 3-step method (Adams, 1992). A continuous synthesis pathway for ibuprofen has been published recently (Bogdan et al., 2009).

**Imatinib** is effective in treatment of multiple cancers, especially chronic myeloid leukaemia (Goldman and Melo, 2003). Marketed as Glivec<sup>®</sup>, it has been highly successful commercially (with revenues of \$4.7 billion in 2012) and commands one of the highest prices for an anti-cancer treatment at \$92,000 a year (Abboud et al., 2013). A continuous production method has been developed as well in order to handle poorly soluble intermediates (Hopkin et al., 2010).

**Ursodeoxycholic acid** (UDCA) is a secondary bile acid which can be conjugated with amino-acids (e.g. taurine and glycine) to form TUDCA and GUDCA, with potential therapeutic uses in stroke, Parkinson's, and Huntington's diseases (Rivard et al., 2007; Boatright et al., 2009). A continuous flow synthesis pathway has also been reported for the conjugation of UDCA with taurine and glycine at the multi-gram scale (Venturoni et al., 2012).

**Reflux inhibitor AZD6906** is a promising API without central nervous system side-effects. However, its batch synthesis is difficult due to highly exothermic reactions, toxic reagents, and unstable products; consequently, researchers have developed a continuous process which adapts a previous batch synthesis route (Gustafsson et al., 2012).

## 2.2 Evaluation of candidate APIs

The identified candidate APIs require systematic evaluation: technical as well as economic considerations have been taken into account and ten (10) broad criteria have been defined in order to comparatively rank the candidates and determine which API would be the most attractive to pursue for flowsheet synthesis, process modelling and simulation.

**A. Is the process fully continuous?** One of the driving forces for CPM is the need to overcome drawbacks such as batch process intermediate storage and scheduling, which result in low operational asset efficiency. A fully continuous process is more desirable than one with even one batch unit operation (typically a separation or purification).

**B. Is the final product a currently marketed API?** Economic evaluation is of high importance for CPM development. The business case for a continuous flow process is stronger if the API is commercially available, because this implies a clear scope for revenue and cost reduction.

**Table 1.** Systematic evaluation of candidate APIs for CPM.

API Candidate	Selection Criteria										Score	Rank
	A	B	C	D	E	F	G	H	I	J		
Ibuprofen	5	5	4	2	4	4	3	5	2	5	39	1
Diphenhydramine	5	5	5	1	4	4	2	5	3	4	38	2
Artemisinin	3	5	3	5	2	4	4	3	4	4	37	3
TUDCA/GUDCA	3	1	4	2	3	3	3	2	1	1	33	4
Activated vitamin D <sub>3</sub>	5	4	3	0	1	2	4	5	3	3	30	5
Imatinib	2	5	1	0	2	4	2	0	5	5	26	6
R-flurbiprofen	5	1	4	0	2	4	4	2	0	0	22	7
AZD6906	0	0	2	0	4	4	4	4	0	0	18	8
CK1 inhibitor	4	0	2	0	3	3	3	1	0	0	16	9
DOR agonist	2	0	2	0	3	3	3	0	0	0	13	10

**C. Is the process of low to medium complexity?** The feasibility and viability of complex CPM pathways are subject to higher uncertainties, so processes which require fewer and/or conventional unit operations and configurations are more preferable.

**D. Has the process been implemented?** Evidence of actual or planned CPM implementation (either at pilot-plant or production level) strengthens the likelihood of success.

**E. Are feedstock and reagent materials readily available?** The availability of feedstocks and reagents is key to feasibility: processes which rely on chemicals which are prohibitively expensive (or whose supply cannot be guaranteed) are less favourable for implementation.

**F. Does the continuous synthesis pathway offer improved performance over the batch process?** A promising API synthesis method must demonstrate clear advantages over current batch production processes.

**G. Can the process be scaled up easily and safely?** A CPM route must be feasible at a pilot plant as well as production level in order to realise the expected benefits over batch production.

**H. Are intermediate separations eliminated?** The most attractive process has a minimal number of intermediate separations (excluding final purification), which imply lower solvent and energy use.

**I. How high is the price of the commercially available product?** The business case (and profitability) of a CPM process is stronger if the final API is sold at a high price.

**J. How high is the global sales volume?** If the product API is in high global demand, the business case for a CPM process will be stronger.

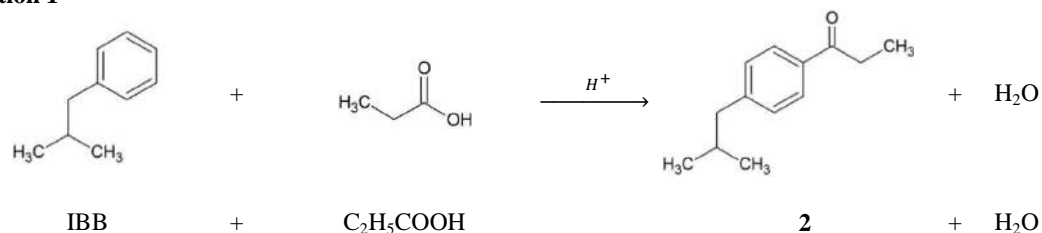
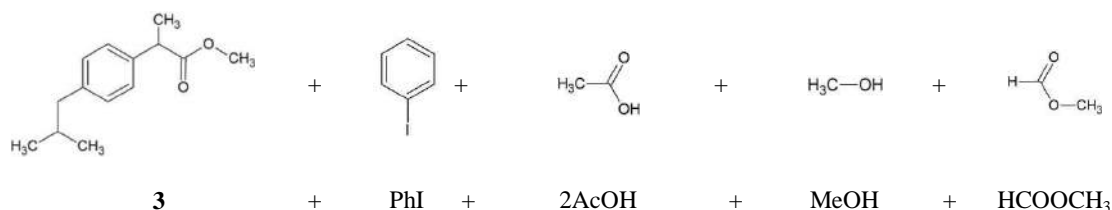
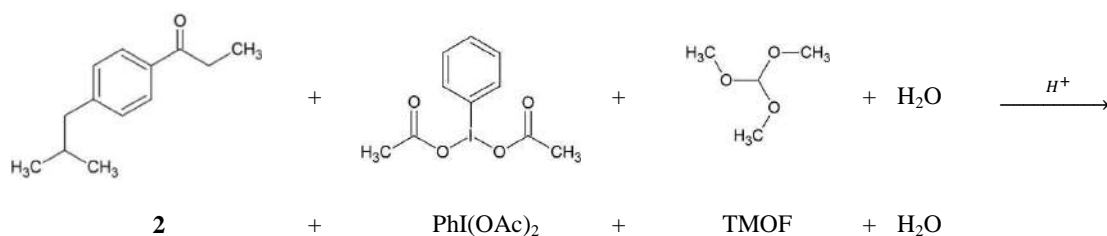
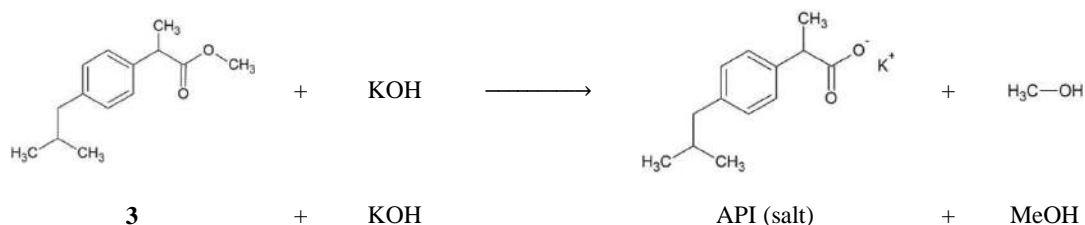
The identified candidate APIs have been graded using a scale of 0 to 5: here, 0 indicates that the criterion is not at all being satisfied and 5 indicates that it is satisfied completely. The complete qualitative analysis and the final API ranking are presented in Table 1.

Ibuprofen emerges as the most promising candidate API. Both artemisinin and diphenhydramine are both widely marketed APIs and thus strong candidates; however, there is limited data on the raw material for artemisinin, while CPM of diphenhydramine relies on the use of molten ionic salts (Snead and Jamison, 2013). Imatinib commands a high price, but its continuous synthesis is complex and requires batch separations (Hopkin et al., 2013). The remaining candidates are less commercially significant, require batch processes, or do not satisfy one or more of the other foregoing criteria.

### 3 FLOWSHEET AND PROCESS MODELLING

The flowsheet considered here is an adaptation of a published continuous flow organic synthesis which relies on a series of three plug flow reactors (PFRs) in series for the production of ibuprofen (Bogdan et al., 2009). The complete reaction scheme is illustrated in Fig. 3.

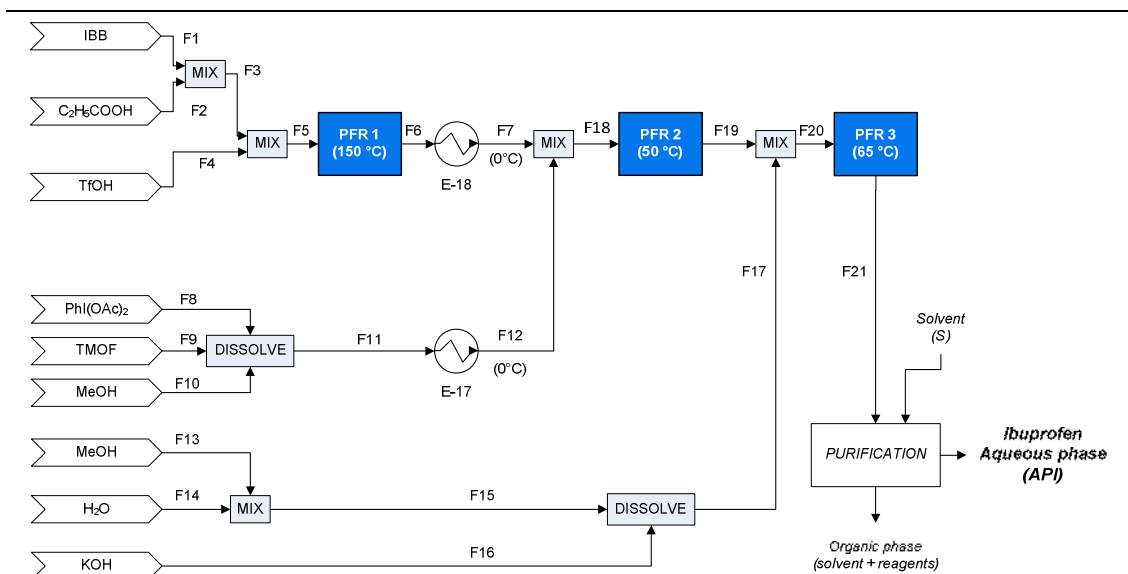
Isobutylbenzene (IBB) is mixed with propanoic acid. The resulting stream is then mixed with neat triflic acid (TfOH, an acid catalyst) and the mixture enters the first reactor, operated at 150 °C: IBB undergoes Friedel-Crafts acylation to produce 1-(4-isobutylphenyl)propan-1-one (intermediate **2**), and the reactor outlet stream is cooled to 0 °C.

**Reaction 1****Reaction 2****Reaction 3**

**Fig. 3.** Reaction stoichiometries and by-products. Reaction 1: Friedel-Crafts acylation in PFR 1. Reaction 2: PhI(OAc)<sub>2</sub>-mediated 1,2-aryl migration in PFR 2. Reaction 3: base hydrolysis in PFR 3 (Bogdan et al., 2009).

Diacetoxyiodobenzene, PhI(OAc)<sub>2</sub>, is dissolved with trimethyl orthoformate (TMOF) in methanol (MeOH). The resulting stream is cooled to 0 °C before mixing with the cooled first reactor product. The combined stream is fed into the second reactor, where intermediate **2** undergoes PhI(OAc)<sub>2</sub>-mediated 1,2-aryl migration to produce methyl 2-(4-isobutylphenyl)propanoate (intermediate **3**), with TfOH again acting as acid catalyst. The second reactor is operated at 50 °C: its product is mixed with potassium hydroxide (KOH) dissolved in a methanol-water mixture and fed into the third reactor.

In the third reactor, intermediate **3** is saponified (base-hydrolysed) to the salt form of ibuprofen, and a series of batch operations has been previously used to separate and purify ibuprofen: the 15-step batch process consists of quenching, extraction, crystallisation and filtration operations, and must be replaced by a continuous final separation. An important mass balance assumption is that the product of the third reactor is entirely acidified to obtain a potassium-free acid form of ibuprofen, as the salt obtained from the third reaction is unsuitable for pharmaceutical formulations (Bogdan et al., 2009). The continuous organic synthesis and purification flowsheet considered here is illustrated in Fig. 4.



**Fig. 4.** Conceptual flowsheet for continuous production of ibuprofen (adapted from Bogdan et al., 2009).

### 3.1 Mass balance calculations

Performing PFR reactor design requires the calculation of the process mass balances (Table 2-3). The starting point is a recent synthesis study (Bogdan et al., 2009), with the following assumptions:

- Reactions occur only inside the PFR reactors and not in any of the lines connecting them.
- The conversion of intermediate **3** to K-ibuprofen in the third reactor can reach 99%.
- Isothermal operation is ensured in all reactors via suitable heating media and thermostats.
- Temperature changes cause no phase transformation or precipitation affecting the flow.
- Esterification reactions between methanol and all organic acids encountered are negligible.
- Reactions between unreacted reagents/by-products and subsequent chemicals are negligible.
- The API salt (K-ibuprofen) is entirely acidified with a suitable acid prior to the final separation.

## 4 KINETIC PARAMETER ESTIMATION

The determination of the order of three reactions and the calculation of reaction rate constants is essential for efficient CPM design. Bogdan et al. (2009) have published kinetic data for the first two continuous organic synthesis reactions. Reaction orders and rate constants have been computed here by evaluating several candidate kinetic expressions, plotting functions of concentrations against time and calculating the corresponding coefficients of determination ( $R^2$ ) for each case considered (Fig. 5). For the first reaction (Friedel-Crafts acylation), experimental data is given for varying residence time (5, 10, 20 minutes), reaction temperature (50, 100, 150 °C), and equivalents of propanoic (1.0, 1.1 and 1.5) and triflic acid (4.0 and 5.0). A second-order reaction is the most plausible case ( $R^2 = 0.836$ , compared to  $R^2 = 0.721$  for zero order and  $R^2 = 0.776$  for first order) and the reaction rate constant has been calculated as  $k_f = 31.41 \text{ L mol}^{-1} \text{ hr}^{-1}$  (at  $T = 150 \text{ °C}$ ), with 1.1 and 5.0 equivalents of propanoic acid and TfOH, respectively. Complex kinetic models would imply multi-stage reaction mechanisms.



Table 2. Process mass balance (g hr<sup>-1</sup>).

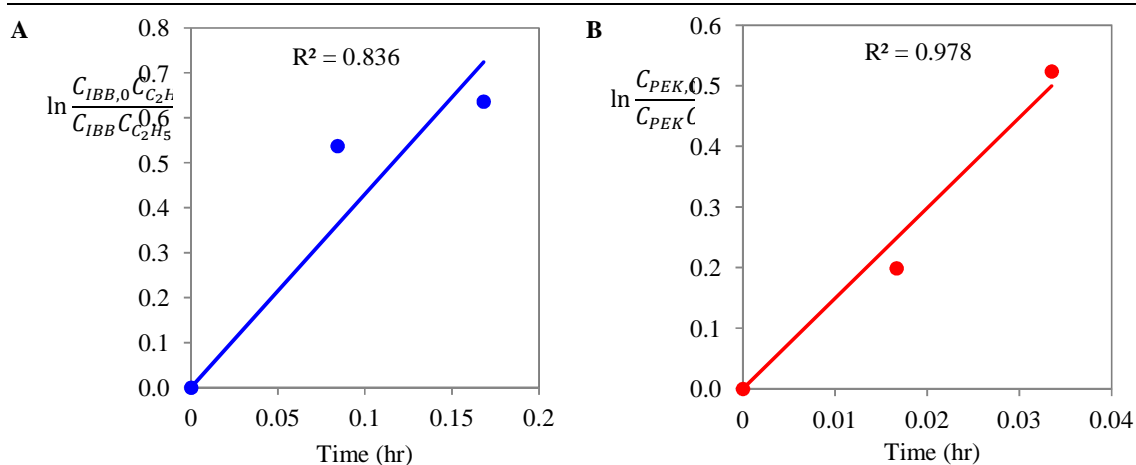
Component	Stream																				
	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21
IBB	5.84	0	5.84	0	5.84	0.53	0.53	0	0	0	0	0	0	0	0	0	0	0.53	0.53	0.53	0.53
2	0	0	0	0	0	7.54	7.54	0	0	0	0	0	0	0	0	0	0	7.54	0.15	0.15	0.15
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8.55	8.55	0.09
API (salt)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9.39
C <sub>2</sub> H <sub>5</sub> COOH	0	3.22	3.22	0	3.22	0.29	0.29	0	0	0	0	0	0	0	0	0	0	0.29	0.29	0	0
C <sub>2</sub> H <sub>5</sub> COOK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.44
PhI(OAc) <sub>2</sub>	0	0	0	0	0	0	14.24	0	0	14.24	14.24	0	0	0	0	0	0	14.24	1.73	1.73	1.73
TMOF	0	0	0	0	0	0	0	18.49	0	18.49	18.49	0	0	0	0	0	0	18.49	14.38	14.38	14.38
HCOOCH <sub>3</sub>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.33	2.33	2.33
H <sub>2</sub> O	0	0	0	0	0	0.71	0.71	0	0	0	0	0	0	30.10	30.10	0	30.10	0.71	0.01	35.49	35.49
TIOH	0	0	0	32.51	32.51	32.51	32.51	0	0	0	0	0	0	0	0	0	0	32.51	32.51	0	0
TIOK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40.76	40.76
AcOH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4.66	0	0
AcOK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7.62	7.62
PhI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7.92	7.92	7.92
KOH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	48.97	46.81
MeOH	0	0	0	0	0	0	0	0	0	41.46	41.46	41.46	95.33	0	95.33	0	95.33	41.46	42.71	138.04	139.27
Total	5.84	3.22	9.07	32.51	41.58	41.58	41.58	14.24	18.49	41.46	74.19	74.19	95.33	30.10	125.43	65.70	191.12	115.77	115.77	306.90	306.90

**Table 3.** Component molar and mass flow changes due to all three continuous flow reactions.

Component	Stoichiometric coefficient	Flow in		Flow change		Flow out	
		(mmol hr <sup>-1</sup> )	(g hr <sup>-1</sup> )	(mmol hr <sup>-1</sup> )	(g hr <sup>-1</sup> )	(mmol hr <sup>-1</sup> )	(g hr <sup>-1</sup> )
<b>PFR 1</b>							
IBB	1	43.52	5.84	39.60	5.31	3.91	0.52
<b>2</b>	+1	0	0	+39.60	+7.53	39.60	7.53
C <sub>2</sub> H <sub>5</sub> COOH	1	43.52	3.22	39.60	2.93	3.91	0.29
H <sub>2</sub> O	+1	0	0	+39.60	+0.71	39.60	0.71
TfOH		216.62	32.51			216.62	32.51
<b>PFR 2</b>							
IBB		3.91	0.52			3.91	0.52
<b>2</b>	1	39.60	7.53	38.82	7.38	0.78	0.15
<b>3</b>	+1	0	0	+38.82	+8.55	38.82	8.55
C <sub>2</sub> H <sub>5</sub> COOH		3.91	0.29			3.91	0.29
H <sub>2</sub> O	1	39.60	0.71	38.82	0.69	0.78	0.01
TfOH		216.62	32.51			216.62	32.51
PhI(OAc) <sub>2</sub>	1	44.19	14.23	38.82	12.50	5.37	1.73
TMOF	1	174.32	18.49	38.82	4.11	135.49	14.37
MeOH	+1	1294.12	41.46	+38.82	+1.24	1332.95	42.70
HCOOCH <sub>3</sub>	+1	0	0	+38.82	+2.33	38.82	2.33
AcOH	+2	0	0	+77.65	+4.66	77.65	4.66
PhI	+1	0	0	+38.82	+7.91	38.82	7.91
<b>PFR 3</b>							
IBB		3.91	0.52			3.91	0.52
<b>2</b>		0.78	0.15			0.78	0.15
<b>3</b>	1	38.82	8.55	38.37	8.46	0.44	0.08
API (salt)	+1	0	0	+38.37	+9.39	38.37	9.39
C <sub>2</sub> H <sub>5</sub> COOK		3.91	0.43			3.91	0.43
H <sub>2</sub> O		1969.27	35.48			1969.27	35.48
TfOK		216.62	40.76			216.62	40.76
PhI(OAc) <sub>2</sub>		5.37	1.73			5.37	1.73
TMOF		135.49	14.37			135.49	14.37
MeOH	+1	4308.24	138.03	+38.37	+1.23	4346.62	139.26
HCOOCH <sub>3</sub>		38.82	2.33			38.82	2.33
AcOK		77.65	7.61			77.65	7.61
PhI		38.82	7.91			38.82	7.91
KOH	1	872.74	48.96	38.37	2.15	834.37	46.81

For the second reaction [PhI(OAc)<sub>2</sub>-mediated 1,2-aryl migration], experimental data has also been analysed for varying residence time (1 and 2 minutes), temperature (35 °C, 50 °C, and 60 °C), and excesses of TfOH (3.0 and 5.0 equivalents) and TMOF (2.0, 4.0 and 9.0 equivalents). In their evaluation of optimum reaction conditions they use propiophenone (PEK) in place of intermediate **2**, with the stated assumption that the results are applicable to **2** on the basis of molecular similarity. From data analysis (Fig. 5), the most plausible case is again a second-order reaction ( $R^2 = 0.978$ , compared to  $R^2 = 0.733$  for zero order and  $R^2 = 0.926$  for first order), and the reaction rate constant has been calculated as  $k_2 = 2732 \text{ L mol}^{-1} \text{ hr}^{-1}$  (at  $T = 50 \text{ °C}$ ), with 5.0 and 4.0 equivalents of TfOH and TMOF, respectively. Kinetic parameter estimation equations and details are given in the Appendix.

For the third reaction (API salt production), empirical data is not available; therefore, a predictive tool for physical properties and chemical reactivity of organic molecules (SPARC) has been used to estimate  $k_3$ , the reaction rate constant for the base hydrolysis reaction (Hilal and Karickhoff, 2003). The feed stream of the third reactor (F20) contains mainly water and methanol (ca. 80% vol.), a fact which justifies our assumption that (for the purpose of reaction rate estimation only) these are the only two solvents considered present, in their actual volume ratios (16.57% and 83.43%, respectively). For the operating temperature of the third reactor (65 °C), the reaction rate constant  $k_3$  has been estimated as  $15.57 \text{ L mol}^{-1} \text{ hr}^{-1}$ . Guidelines on SPARC model implementation are available by its developers.



**Fig. 5.** Reaction order and kinetic parameter determination. A: reaction 1 (feed concentrations:  $C_{IBB,0} = 1.42$  M,  $C_{C_2H_5COOH,0} = 1.56$  M). B: reaction 2 (feed concentrations:  $C_{PEK,0} = 0.38$  M,  $C_{PhI(OAc)_2,0} = 0.39$  M).

## 5 PLUG FLOW REACTOR (PFR) DESIGN

Reactor design is performed assuming ideal solutions, homogeneous mixtures, no radial concentration gradients, and isothermal operation without axial or radial temperature gradients: specific heat capacities and reaction enthalpies remain constant within each PFR, and phase changes do not occur. All reactors are assumed to be submerged in heating media with adequate flow circulation; due to the small radial size considered, heat transfer suffices so as to ensure constant and uniform temperature.

The first reaction is assumed to be second-order (first in IBB and first in propanoic acid), the second reaction is also assumed to be second-order (first order in intermediate 2 and first order in  $PhI(OAc)_2$ ), and the third reaction is assumed to be pseudo-first-order (first-order in intermediate 3 due to a considerable excess of potassium hydroxide used for the base hydrolysis). These assumptions are justified by the kinetic parameter estimation and model reactions with the highest possible fidelity.

The standard plug flow reactor performance equation has been used:

$$\tau_i = C_{A,0} \int_0^{X_{A,f}} \frac{dX_A}{-r_A} \quad (1)$$

Published data indicate that 91% conversion is achievable in the first reactor, and that over 98% is attainable in the second one (Bogdan et al., 2009). No data has been published for the conversion of the base hydrolysis in the third reactor, for which a value of 99% has been assumed. The assumptions that have been made and the parameters estimated toward PFR design are summarised in Table 4.

**Table 4.** Plug Flow Reactor (PFR) design basis and assumptions.

	PFR 1	PFR 2	PFR 3
Reaction	1	2	3
Reaction order	2	2	2
Reaction type	Friedel-Crafts acylation	1,2-aryl migration	Base hydrolysis
Rate law	$-r_{IBB} = k_1 C_{IBB} C_{C_2H_5COOH}$	$-r_2 = k_2 C_2 C_{PhI(OAc)_2}$	$-r_3 = k'_3 C_3$
Reactor temperature ( $^{\circ}C$ )	150	50	65
Conversion (%)	91	98	99
Rate constant $k_i$ ( $L \text{ mol}^{-1} \text{ hr}^{-1}$ )	31.41	2732.3	15.57
Coefficient of determination ( $R^2$ )	0.836	0.978	–
Method for $k_i$	Data analysis	Data analysis	SPARC model

## 6 MATERIAL PROPERTIES

Thermodynamic and physical properties of certain organic compounds in the CPM flowsheet are not available in the literature and thus require estimation by reliable property modelling methods. The compilation of published data and property modelling estimates conducted is summarised in Table 5.

Reactor heating duties (for feed heating as well as maintaining prescribed operating temperatures) require accurate estimation of reaction enthalpies, which in turn can be computed using the standard formation enthalpies of chemical compounds, and specific heat capacities under constant pressure.

For intermediates **2** and **3**, ibuprofen, PhI(OAc)<sub>2</sub>, methyl formate (HCOOCH<sub>3</sub>), and acetic acid (AcOH), standard formation enthalpies have been calculated using a group contribution method (Constantinou and Gani, 1994):

$$\Delta^{\circ}H_f - h_{f0} = \sum_i N_i h_{1i} + \sum_j M_j h_{2j} \quad (2)$$

Here,  $h_{1i}$  is the contribution of first-order group  $i$ , and  $N_i$  is the number of occurrences of  $h_{1i}$ . Similarly,  $h_{2j}$  is the contribution of second-order group  $j$ , and  $M_j$  is the number of occurrences of  $h_{2j}$ . The constant value  $h_{f0} = 10.835 \text{ kJ mol}^{-1}$  is universal, and tabulated  $h_{1i}$  and  $h_{2j}$  values are available.

The specific heat capacity under constant pressure ( $C_p$ ) has been computed via the Shomate equation for water, IBB, propanoic acid, and potassium hydroxide, using the polynomial coefficients  $A$ ,  $B$ ,  $C$ ,  $D$  and  $E$  which are available in the NIST-JANAF Thermochemical Tables (Chase, 1998):

$$C_p = A + BT + CT^2 + DT^3 + \frac{E}{T^2} \quad (3)$$

The specific heat capacity under constant pressure of all other molecules (for which data is not available) has been computed via a group contribution method (Rihani and Doraiswamy, 1965):

$$C_p = \sum_i n_i a_i + \sum_i n_i b_i T + \sum_i n_i c_i T^2 + \sum_i n_i d_i T^3 \quad (4)$$

Here  $n_i$  is the number of occurrences of functional group  $i$  within the molecule,  $a_i$ ,  $b_i$ ,  $c_i$  and  $d_i$  are tabulated literature values associated with group  $i$ , and  $T$  is the absolute temperature (K).

**Table 5.** Physical properties of CPM flowsheet stream components.

*a: Chase, 1998; b: Rihani and Doraiswamy, 1965; c: Constantinou and Gani, 1994; d: published data.*

Component	MW (g mol <sup>-1</sup> )	Density (g cm <sup>-3</sup> )	Melting point (°C)	Boiling point (°C)	$C_p$ (T = 25 °C) (kJ mol <sup>-1</sup> K <sup>-1</sup> )	$\Delta^{\circ}H_f$ (kJ mol <sup>-1</sup> )
IBB	134.22	0.844	51.7	170.9	242.79 <sup>a</sup>	69.9 <sup>d</sup>
<b>2</b>	190.29	0.960	18.4	248.8	242.76 <sup>b</sup>	139.9 <sup>c</sup>
<b>3</b>	220.31	1.100	38.3	287.0	264.95 <sup>b</sup>	313.6 <sup>c</sup>
Ibuprofen (API)	206.29	1.030	74.5	157.0	278.98 <sup>b</sup>	350.3 <sup>c</sup>
C <sub>2</sub> H <sub>5</sub> COOH	74.08	0.990	21.2	140.9	158.59 <sup>b</sup>	510.4 <sup>d</sup>
H <sub>2</sub> O	18.02	1.000	0.0	100.0	75.37 <sup>a</sup>	285.8 <sup>d</sup>
TfOH	150.08	1.708	40.0	161.9	110.53 <sup>b</sup>	n.r.
PhI(OAc) <sub>2</sub>	322.10	1.814	162.5	456.8	200.16 <sup>b</sup>	628.1 <sup>c</sup>
TMOF	106.12	0.970	53.0	102.8	136.50 <sup>b</sup>	586.9 <sup>d</sup>
MeOH	32.04	0.792	97.2	64.9	81.42 <sup>d</sup>	239.0 <sup>d</sup>
HCOOCH <sub>3</sub>	60.05	0.980	100.0	32.0	93.65 <sup>b</sup>	367.9 <sup>c</sup>
AcOH	60.05	1.049	16.5	118.0	66.58 <sup>b</sup>	442.2 <sup>c</sup>
PhI	204.01	1.850	31.4	188.0	95.05 <sup>b</sup>	234.0 <sup>d</sup>
KOH	56.11	1.050	406.0	1327.0	64.89 <sup>a</sup>	482.4 <sup>d</sup>

Chemical solubilities have been computed using the UNIFAC group contribution method (Fredenslund et al., 1975), via activity coefficients which are sums of a combinatorial and a residual component. The prerequisite enthalpies of fusion have also been calculated via a group contribution method (Chickos and Acree, 1999). All property estimation equations are presented in the Appendix.

## 7 HEAT TRANSFER DESIGN

Energy balance calculation is essential for process R&D and CPM evaluation: therein, heat transfer duty constitutes the dominant energy requirement, since process streams require heating or cooling, while reactor operating temperatures must be maintained constant for maximum process efficiency.

All three plug flow reactors are assumed isothermal (PFR 1: 150 °C, PFR 2: 50 °C, PFR 3: 65 °C). Reactor and process lines are considered submerged in heat transfer media with adequate circulation, and heat transfer dominates over ancillary (e.g. pumping) requirements in this conceptual flowsheet. Reactor heating duty calculations require reaction enthalpy values at corresponding temperatures; these have been computed using the reaction enthalpies at standard state, which in turn emerge from standard enthalpies of formation of the reagents and products for the respective chemical reactions:

$$\Delta H_{rxn}^{\circ} = \sum \Delta H_{f,products}^{\circ} - \sum \Delta H_{f,reagents}^{\circ} \quad (5)$$

Reaction enthalpies at the corresponding reactor temperatures are calculated using Hess's Law, considering the sum of three terms: the enthalpy change in cooling the reactants from the constant reactor temperature  $T$  to standard temperature ( $T^{\circ} = 25$  °C), the reaction enthalpy at standard state, and the enthalpy change in heating the products from 25 °C to the reactor operating temperature:

$$\Delta H_{rxn} = \sum C_{P,reagents}(T^{\circ} - T) + \Delta H_{rxn}^{\circ} + \sum C_{P,products}(T - T^{\circ}) \quad (6)$$

The heating duty for each reactor is thus computed as a sum of feed heating and reaction enthalpy.

For temperatures above the boiling points of certain components (e.g. methanol, methyl formate), empirical observations indicate that partial evaporation (off-gassing) is possible (Bogdan et al., 2009). Nevertheless, this process simulation study assumes that all components remain in the liquid phase throughout the flowsheet and off-gassing is negligible by suitable reactor design (Mascia et al., 2013).

## 8 FINAL SEPARATION DESIGN

Development of continuous separation and purification unit operations is critical in order to attain the full potential of CPM. Product streams frequently contain large quantities of excess reagent to be recycled, and by-products which must be removed; moreover, an aqueous product (API) stream (e.g. via continuous solvent exchange) is essential in downstream processing and final dosage formation.

A Pfizer solvent selection guide classifies solvents as preferred, usable or undesirable, based on their environmental impact (Alfonsi et al., 2008): accordingly, a selection of solvents from all three categories has been evaluated, including ethanol and ethyl acetate (preferred), toluene and acetonitrile (usable), and hexane (undesirable) and methyl acetate (unclassified). Water and methanol (preferred) are both present in the process, as they are essential according to the organic synthesis considered.

The API potassium salt form dissolved in stream F21 cannot be used in commercial formulations, due to its extreme hygroscopicity which compromises stability (Gruber, 2009). Before purification, stream F21 is assumed entirely acidified so that all ibuprofen is in its formal (carboxylic acid) form. The composition of the remainder of stream F21 has been assumed unchanged after this acidification step (the conversion of API salt to carboxylic acid form does not affect other effluent components).

The product stream of the third reactor must undergo final API purification. An efficient choice for this continuous API separation from an organic stream is liquid-liquid extraction (LLE): therein, the introduction of a new solvent (in which the valuable API solute exhibits preferentially high solubility) induces a multiphase mixture, as long as it is partially or entirely immiscible with the original solvent. For LLE design, stream F21 is assumed a binary mixture of water and methanol with several solutes. Ternary phase diagrams have been used to calculate compositions upon solvent addition and phase separation (Skrzecz et al., 1999) at ambient and effluent operating temperatures (25 °C and 65 °C).

Thermodynamic equilibrium between the aqueous and the organic phase is rapidly established in the LLE unit, and no crystallisation or precipitation is considered to occur under these flow conditions. The API distribution in each of the two (organic/O and water/W) phases relies on the assumption that the partition coefficient of ibuprofen is equal to the ratio of corresponding solubilities in each phase; these are computed for each of the resulting (O, W) phases via the UNIFAC method (cf. Appendix).

## 9 RESULTS AND DISCUSSION

### 9.1 Mass balance

The continuous system presented here produces  $9.39 \text{ g hr}^{-1}$  ( $72.6 \text{ kg}$  annually for 46 weeks of operation) of potassium ibuprofen, which corresponds to  $61.3 \text{ kg}$  of ibuprofen annually after acidification and prior to recovery. The computed plantwide mass balance relies on experimental conversions (Bogdan et al., 2009), and a graphic summary of four key streams is illustrated in Fig. 6.

The main organic solvents are methanol (MeOH), and triflic acid (TfOH): the former is the primary carrier solvent of reagents (but also a minor by-product) of the second and third reactors (Figs. 3-4) and forms 45.4% ( $139.3 \text{ g hr}^{-1}$ ) of the reactor effluent stream (F21) which is fed to final purification. The latter occurs in most streams (Table 2) and serves as acid catalyst in the first two reactors (Fig. 3). The majority of streams are hence predominantly organic, but water is also a critical component of the solvent mixture (stream F15) which is essential for the base hydrolysis in the third reactor (Fig. 4).

Apart from key organic molecules (IBB, intermediates **2** and **3**, and ibuprofen in salt form), significant quantities of unreacted reagents and by-products remain in solution, often due to the need

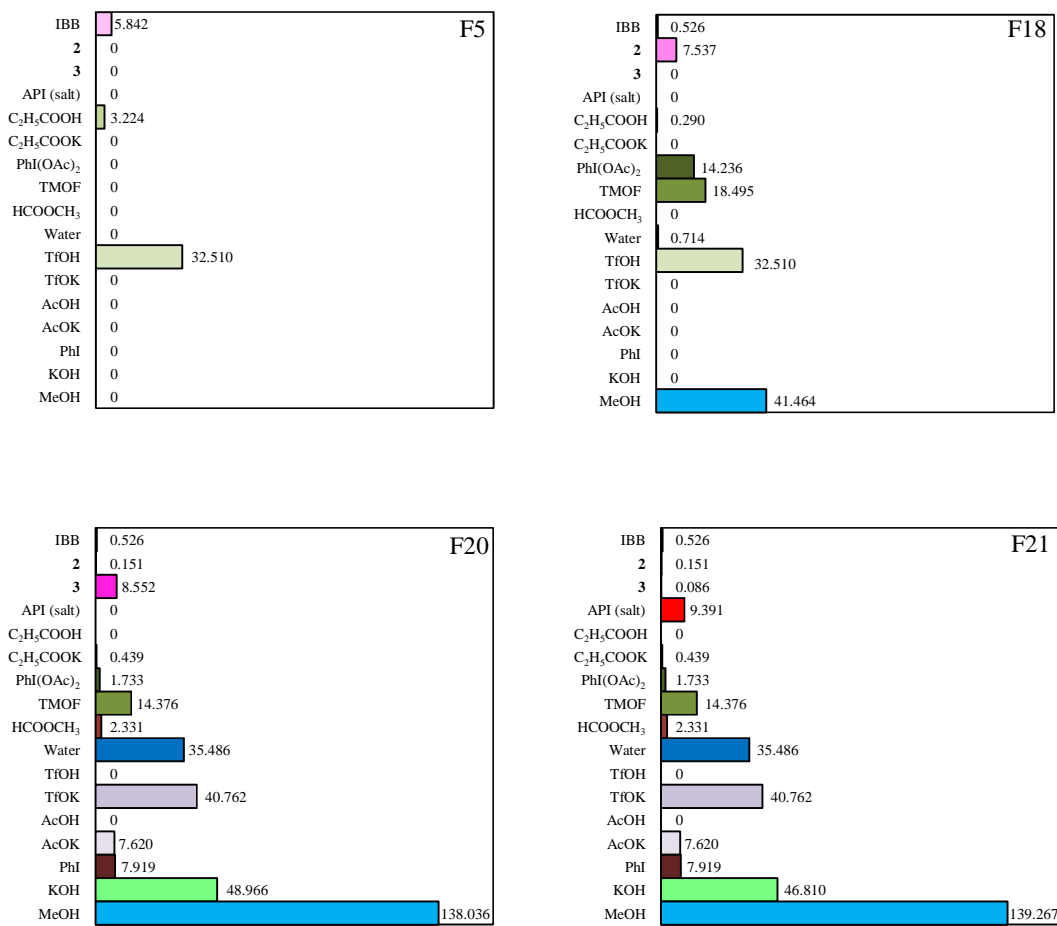


Fig. 6. Mass balance of key flowsheet streams ( $\text{g hr}^{-1}$ ).

**Table 6.** PFR design results.

a: Bogdan et al. (2009); b: assumption.

PFR #	Temperature (°C)	Flowrate (mg hr <sup>-1</sup> )	Conversion (%)	Internal diameter, ID (mm)	Volume (mL)	Length (mm)
PFR 1	150	41.58	91 <sup>a</sup>	5.0	6.392	326
PFR 2	50	115.77	98 <sup>a</sup>	5.0	1.997	102
PFR 3	65	306.90	99 <sup>b</sup>	5.0	28.912	1472

to use reagents in excess (e.g. TMOF, HCOOCH<sub>3</sub>, KOH). This is particularly important in stream F21, where unreacted reagents and by-products form 92.9% of the dissolved chemical substances, implying a clear need for a final separation ensuring API purification before downstream processing.

## 9.2 Reactor design

The reactor sizes computed (PFR 1: 6.392 mL, PFR 2: 1.997 mL, PFR 3: = 28.912 mL) (Table 7) are reasonably small for the target capacity, illustrating the CPM benefit of space-efficient equipment. The reactors have been designed with a relatively small internal diameter (ID = 5.0 mm) to ensure efficient heat transfer; CPM pilot plant systems have been successfully demonstrated at this scale, with heat transfer ID = 1.6 mm and reactor ID = 11.7 to 40 mm, respectively (Mascia et al., 2013). Process simulations have also been performed at the reported capacity and compared to the published reactor dimensions implemented in the laboratory-scale experimental system (Bogdan et al., 2009). Computed reactor volumes are similar; the third reactor estimate is larger than that reported, implying a SPARC underestimation of the base hydrolysis rate constant ( $k_3$ ) due to the lack of kinetic data.

## 9.3 Heat transfer design

Continuous synthesis reactions are endothermic: enthalpies at PFR temperatures are given in Table 7. Inlet streams to all three reactors require energy to reach and be kept at the operating temperature. The computed reactor heating requirements are of the same order of magnitude, but the wide variation in reactor lengths induces a considerably lower specific heating duty for the third reactor (Table 7). Energy duties have been computed solely for preheating feeds and maintaining isothermal operation: pumping, cooling, losses and power for Process Analytical Technology (PAT) are not considered, but heat integration opportunities must be taken into account in the detailed design phase for CPM plants.

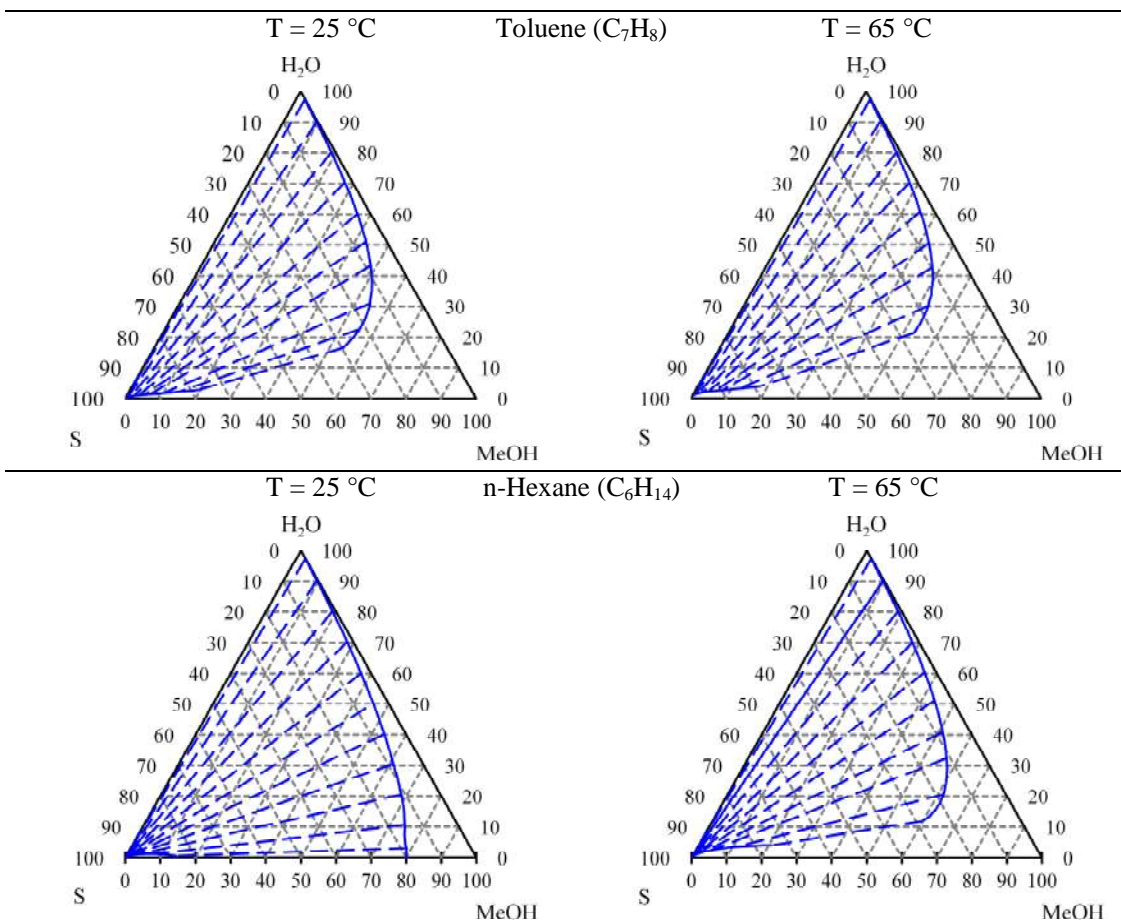
## 9.4 Final separation design

Six organic solvents (ethanol, acetonitrile, toluene, n-hexane, ethyl acetate, methyl acetate) have been evaluated as LLE candidates; the first five are discussed in a Pfizer publication (Alfonsi et al., 2008). Ethanol and acetonitrile are fully miscible with stream F21, yielding no manageable phase separation. System compositions obtained for LLE using methyl/ethyl acetate are outside the multiphase region for all levels of solvent use in both temperatures considered, so miscibility renders them unsuitable.

Ternary equilibrium phase diagrams for likely solvents (toluene and n-hexane) are given in Fig. 7. The solubility of ibuprofen in the resulting immiscible phases has been calculated for each solvent mixture and each operating temperature via the UNIFAC method; results are summarised in Table 8. The solubilities reported here are considerably higher in case of high-temperature operation at 65 °C. Calculated ibuprofen solubilities for pure solvents (methanol, toluene, n-hexane; very poor in water) have been validated and are in very good agreement with previous data (Gracin and Rasmuson, 2002). Ibuprofen saturation is never approached in this design, so the focus is on ratios of phase solubilities.

**Table 7.** PFR heating: energy requirements.

PFR #	Reaction	Length (mm)	ID (mm)	$T_{feed}$ (°C)	$\Delta H_{rxn}$ (kJ mol <sup>-1</sup> )	$T_{PFR}$ (°C)	Flowrate in (mg hr <sup>-1</sup> )	Flowrate in (mol hr <sup>-1</sup> )	Duty (W cm <sup>-1</sup> PFR)	Power (W)
PFR 1	Endothermic	326	5.0	25	143.0	150	41.58	0.304	0.103	3.361
PFR 2	Endothermic	102	5.0	0	54.2	50	115.77	1.816	0.299	3.042
PFR 3	Endothermic	1472	5.0	50	207.5	65	306.90	7.710	0.033	4.852



**Fig. 7.** Ternary phase diagrams for water-methanol-solvent (H<sub>2</sub>O-MeOH-S) systems at both temperatures.

The performance of each solvent has been evaluated by considering API solubility and distribution for a wide range of solvent-to-feed mass ratios (0.25 to 5), at two key temperatures (25 °C and 65 °C). API recovery is higher at ambient temperature (25 °C) for n-hexane, and at higher (65 °C) for toluene. Ambient operation better illustrates the scope of CPM and its inherent green chemistry potential, as the option to operate purification separations without high energy requirement is certainly desirable.

Hexane performs best for the lowest solvent-to-feed ratio (0.25) at ambient LLE operation (25 °C), achieving the highest API recovery of 97.4% at the expense of close proximity to the equilibrium line. For a more practical solvent-to-feed ratio (0.5), API recovery remains over 89% at both temperatures.

Toluene achieves a marginally to clearly lower API recovery than hexane, at ambient temperature. Nonetheless, the use of toluene is preferable as it is a classified usable solvent of safe performance, while n-hexane is classified as undesirable due to its high environmental impact (Alfonsi et al., 2008). The lower solvent-to-feed ratios that are feasible for hexane (0.25, 0.5) are not operable for toluene, as they yield a perfectly miscible mixture which cannot induce the necessary phase separation for LLE. However, the minimum feasible solvent-to-feed ratio considered in simulations (0.75) still achieves acceptably high API recovery, which reaches 81.7% and 89.2% at 25 °C and 65 °C, respectively. Ambient operation (25 °C) is recommended as more energy-efficient: it avoids the LLE heating cost and the reported possibility of partial evaporation (which is even more likely at larger scale), thus achieving an annual production of 50.0 kg of ibuprofen (assuming 46 weeks of plant operation).





The environmental (E-) factor is a widely used green chemistry metric with a versatile definition (Sheldon, 2012): in its simplest form, it is the waste mass generated per unit mass or product, so lower E-factor values are more desirable and indicate greener design. Ranging widely, it is typically 0.1 in petrochemical industries, but can be as high as 200 for pharmaceutical products (Ritter, 2013).

The E-factor is computed as a mass ratio of waste (all material not recovered) to API (ibuprofen):

$$E = \frac{m_{waste}}{m_{API}} \quad (7)$$

$$m_{waste} = m_{bpd} + m_{ur} + m_{us} + m_{API} \quad (8)$$

E-factors have been calculated considering a downstream solvent (methanol and toluene/n-hexane) recovery of 90% (Sheldon, 2012), API recovery as reported, and remaining waste comprising all by-products, unreacted feed, 10% waste solvents and un-recovered API. Results are given in Table 8. The solvent feed obviously affects the effluent masses of unrecovered API and solvent waste, both of which in turn have a strong impact on the E-factor, as indicated by Eq. (8); this is particularly evident in case of the highest solvent-to-feed ratio (5), where the E-factor is larger by an order of magnitude.

Toluene is the recommended LLE solvent when considering green chemistry but also performance: for a solvent-to-feed ratio of 0.75 at ambient temperature (25 °C), its use yields an E-factor of 25.4. This is actually a quite acceptable value, rather low for the purpose of pharmaceutical manufacturing, and much lower than E-factors of many pharmaceuticals produced by batch methods (Sheldon, 2012).

## 10 CONCLUSIONS

This paper presents a steady-state process model and simulation for the continuous production of a selected API which has emerged after systematic evaluation of several technical and market criteria. Continuous plug flow reactors and the final purification have been designed for CPM of ibuprofen, via systematic kinetic parameter estimation, solvent selection and thermodynamic property modelling, which is imperative for most organic molecules. Green design efficiency is explored via the E-factor.

Reaction order determination and kinetic parameter estimation have been performed by analysing published data of a continuous flow synthesis pathway (Bogdan et al., 2009). Three plug flow reactors in series have been designed: the computed PFR volumes (6.4, 2.0, 28.9 mL) are relatively small and clearly illustrate the fixed cost and plant footprint advantages of continuous over batch manufacturing. Energy requirements for ensuring high endothermic conversions under isothermal reactor operation are also quite reasonable and imply high potential for energy efficiency, justifying the expectation of fixed as well as operating cost advantages of CPM over batch manufacturing (Schaber et al., 2011). Actual CPM energy requirements will be slightly higher, due to auxiliary (e.g. pumping) functions.

Design of a continuous final separation step after the third reactor is essential for a CPM process, to obtain an aqueous API stream which is pure enough to be fed to secondary (downstream) processing. A systematic evaluation of several candidate organic solvents (via UNIFAC solubility estimation) indicates that toluene is the most efficient extraction solvent, achieving a 81.7% ibuprofen recovery at a solvent-to-feed mass ratio of 0.75 and ambient temperature (25°C), yielding 50.0 kg API annually. This systematic LLE separation design achieves high API recovery with low solvent requirement, thus demonstrating the green chemistry principles of solvent and waste minimisation: a widely used environmental impact metric (the E-factor) has been computed for all candidate designs, and the value corresponding to the selected design (25.4) is quite low compared to many pharmaceutical processes.

Ibuprofen hereby emerges as an ideal API for Continuous Pharmaceutical Manufacturing (CPM): this therapeutic molecule is in high global demand and its production has extreme societal importance, thus representing a significant business opportunity for pilot- and production-scale implementations. The small PFR volumes and heating duties computed illustrate the documented (Roberge et al., 2008) and realised (Mascia et al., 2013) advantages of CPM technology, while the final LLE separation and reasonable solvent duty showcase the potential to design inherently green pharmaceutical processes. The suitability of methanol and a solvent exchange to ethanol (the *de facto* batch processing solvent) require more investigation, to facilitate the integration of CPM with downstream processing practice. The concurrent need for model-based technical and economic optimisation of the entire CPM process also emerges clearly by this process simulation study and must be pursued with systematic methods.

## ACKNOWLEDGEMENT

The authors gratefully acknowledge the financial support of the Engineering and Physical Sciences Research Council (EPSRC) via a Doctoral Training Partnership (DTP) studentship to Mr H.G. Joliffe.

## NOMENCLATURE

### Latin Letters

$A$	Shomate equation parameter
$a_i$	Heat capacity group contribution estimation parameter for functional group $i$
$a^s$	Pure solid component activity
$B$	Shomate equation parameter
$b_i$	Heat capacity group contribution estimation parameter for functional group $i$
$C$	Shomate equation parameter
$C_A$	Concentration of molecule $A$ , mol L <sup>-1</sup>
$C_{A,0}$	Initial concentration of molecule $A$ , mol L <sup>-1</sup>
$C_i$	Group coefficient for functional group $i$ in the total phase change entropy group additivity estimation method
$C_p$	Specific heat capacity at constant pressure, J mol <sup>-1</sup> K <sup>-1</sup>
$c_i$	Heat capacity group contribution estimation parameter for functional group $i$
$D$	Shomate equation parameter
$d_i$	Heat capacity group contribution estimation parameter for functional group $i$
$E$	Shomate equation parameter
$G_i$	Contribution from one functional group $i$ to the total phase change entropy in the group additivity estimation method, J mol <sup>-1</sup> K <sup>-1</sup>
$h_{f0}$	Constant used in estimating the standard enthalpy of formation $\Delta^\circ H_f$
$h_{fi}$	Standard enthalpy of formation group contribution of a single first-order group $i$
$h_{2j}$	Standard enthalpy of formation group contribution of a single second-order group $j$
$ID$	Internal diameter, mm
$k'$	Pseudo-first-order reaction rate constant, s <sup>-1</sup>
$k_i$	Rate constant of reaction $i$ , L mol <sup>-1</sup> s <sup>-1</sup> (second-order reaction)
$L_i$	UNIFAC compound parameter of $r$ , $q$ and $z$
$M_j$	Number of $j$ second-order functional groups
$m_{API}$	Mass of recovered API
$m_{bpd}$	Mass of byproducts
$m_{ur}$	Mass of unreacted reagents
$m_{us}$	Mass of unrecovered solvent
$m_{uAPI}$	Mass of unrecovered API
$MW$	Molecular weight, g mol <sup>-1</sup>
$m_{waste}$	Mass of waste
$N_i$	Number of $i$ first-order functional groups
$n_{CH_2}$	Number of consecutive CH <sub>2</sub> groups
$n_i$	Number of $i$ functional groups
$Q_k$	UNIFAC surface area parameter for functional group $k$
$q_i$	UNIFAC parameter for molecule $i$ , a measure of Van der Waals molecular surface area
$R$	Universal gas constant, 8.3144 J mol <sup>-1</sup> K <sup>-1</sup>
$R^2$	Coefficient of determination
$R_k$	UNIFAC volume parameter for functional group $k$
$r_A$	Rate of reaction of molecule $A$ , mol l <sup>-1</sup> s <sup>-1</sup>
$r_i$	UNIFAC parameter for molecule $i$ , a measure of Van der Waals volume
$T$	Temperature, K
$T^\circ$	Standard temperature, 298 K
$T_{feed}$	Feed temperature, °C
$T_{PFR}$	Reactor temperature, °C
$T_{fus}$	Melting (fusion) point temperature, K

$t$	Time, hr
$U_{mn}$	UNIFAC energy interaction between groups $m$ and $n$
$X_A$	Conversion of molecule $A$ at time $t$
$X_{A,f}$	Final conversion of molecule $A$
$x_i$	Mole fraction of molecule $i$
$x_m$	UNIFAC mole fraction of group $m$
$x_i^{sat}$	Mole fraction at saturation (solubility)
$z$	UNIFAC system coordination number (default value: 10)

### Greek Letters

$\Gamma_k$	UNIFAC residual group activity coefficient for group $k$
$\Gamma_k^{(i)}$	UNIFAC residual group activity coefficient in a reference solution of 100% $i$ molecules
$\gamma_i$	Activity coefficient of component $i$
$\gamma_i^c$	UNIFAC combinatorial component of $\gamma_i$
$\gamma_i^r$	UNIFAC residual component of $\gamma_i$
$\gamma_i^{sat}$	Activity coefficient of component $i$ at saturation
$\Delta^\circ H_f$	Standard enthalpy of formation, J mol <sup>-1</sup>
$\Delta H_{fus}$	Enthalpy of fusion, J mol <sup>-1</sup>
$\Delta^\circ H_{rxn}$	Standard reaction enthalpy, J mol <sup>-1</sup>
$\Delta H_{rxn}$	Reaction enthalpy, J mol <sup>-1</sup>
$\Delta_m^{T_{fus}} H_{tpce}$	Total phase change enthalpy, J mol <sup>-1</sup>
$\Delta S_{fus}$	Entropy of fusion, J mol <sup>-1</sup> K <sup>-1</sup>
$\Delta_0^{T_{fus}} S_{tpce}$	Total phase change entropy, J mol <sup>-1</sup> K <sup>-1</sup>
$\theta_i$	UNIFAC molar-weighted area fractional component for molecule $i$
$\theta_m$	UNIFAC summation of area fraction of group $m$ over all different groups
$\nu_k^{(i)}$	UNIFAC number of occurrences of functional group $k$ in molecule $i$
$\tau_i$	Residence time in reactor $i$ , hr
$\phi_i$	UNIFAC molar-weighted segment fractional component molecule $i$
$\Psi_{mn}$	UNIFAC group energy interaction parameter

### LITERATURE REFERENCES

1. Abboud, C., Berman, E., Cohen, A., Cortes, J., DeAngelo, D. et al., 2013. The price of drugs for chronic myeloid leukemia (CML) is a reflection of the unsustainable prices of cancer drugs: from the perspective of a large group of CML experts. *Blood* **121**(22): 4439–4442.
2. Abrams, D.S., Prausnitz, J.M., 1975. Statistical thermodynamics of liquid mixtures: a new expression for the excess Gibbs energy of partly or completely miscible systems. *AIChE J.* **21**(1): 116–128.
3. Adams, S.S., 1992. The propionic acids: a personal perspective. *J. Clin. Pharmacol.* **32**(4): 317–323.
4. Almario, G.A., Barrague, M., Burnier, P., Enguehard, C., Gao, Z. et al., 2009. 6-cycloamino-3-(pyridin-4-yl)imidazo[1,2-b]pyridazine derivatives, preparation thereof and therapeutic use thereof. WO2009016286 A3.
5. Alfonsi, K., Colberg, J., Dunn, P.J., Fevig, T., Jennings, S. et al., 2008. Green chemistry tools to influence a medicinal chemistry and research chemistry based organisation. *Green Chem.* **10**(1):31–36
6. Anderson, N.G., 2012. Using continuous processes to increase production. *Org. Process Res. Dev.* **16**(5): 852–869.
7. Ashe, R., 2012. From batch to continuous processing. *Chem. Eng.-New York* **119**(10): 34–40.
8. Behr, A., Brehme, V. A., Ewers, C.L.J., Grön, H., Kimmel, T. et al., 2004. New developments in chemical engineering for the production of drug substances. *Eng. Life Sci.* **4**(1): 15–24.
9. Boatright, J.H., Nickerson, J.M., Moring, A.G., Pardue, M.T., 2009. Bile acids in treatment of ocular disease. *J. Ocul. Biol. Dis. Infor.* **2**(3): 149–159.
10. Bogdan, A.R., Poe, S.L., Kubis, D.C., Broadwater, S.J., McQuade, D.T., 2009. The continuous-flow synthesis of ibuprofen. *Ind. Angew. Chem. Int. Edit.* **48**(45): 8547–8550.
11. Chase, M.W., Jr., 1998. NIST-JANAF Thermochemical Tables (4<sup>th</sup> edition), *J. Phys. Chem. Ref. Data*, Monograph 9.
12. Chickos, J.S., Acree, W.E., 1999. Estimating solid-liquid phase change enthalpies and entropies. *J. Phys. Chem. Ref. Data* **28**(6): 1535–1673.

13. Constantinou, L., Gani, R., 1994. New group contribution method for estimating properties of pure compounds. *AICHE J.* **40**(10): 1697–1710.
14. Douglas, J.M., 1988. *Conceptual Design of Chemical Processes*. McGraw-Hill, New York.
15. Fredenslund, A., Jones, R.L., Prausnitz, J.M., 1975. Group-contribution estimation of activity coefficients in nonideal liquid mixtures. *AICHE J.* **21**(6): 1086–1099.
16. Fuse, S., Mifune, Y., Tanabe, N., Takahashi, T., 2012. Continuous-flow synthesis of activated vitamin D<sub>3</sub> and its analogues. *Org. Biomol. Chem.* **10**(27): 5205–5211.
17. Gernaey, K.V., Cervera-Padrell, A.E., Woodley, J.M., 2012. Development of continuous pharmaceutical production processes supported by process systems engineering methods and tools. *Future Med. Chem.* **4**(11): 1371–1374.
18. Gerogiorgis, D.I., Barton, P.I., 2009. Steady-state optimization of a continuous pharmaceutical process. *Comp. Aid. Ch.* **27**(A): 927–932.
19. Gerogiorgis, D.I., Ydstie, B.E., 2005. Multiphysics CFD modelling for design and simulation of a multiphase chemical reactor. *Chem. Eng. Res. Des.* **83**(A6): 603–610.
20. Goldman, J.M., Melo, J.V., 2003. Chronic myeloid leukemia – advances in biology and new approaches to treatment. *N. Engl. J. Med.* **349**(15): 1451–1464.
21. Gracin, S., Brinck, T., Rasmuson, Å.C., 2002. Prediction of solubility of solid organic compounds in solvents by UNIFAC. *Ind. Eng. Chem. Res.* **41**(20): 5114–5124.
22. Gracin, S., Rasmuson, Å.C., 2002. Solubility of phenylacetic Acid, p-hydroxyphenylacetic acid, p-aminophenylacetic acid, p-hydroxybenzoic acid, and ibuprofen in pure solvents. *J. Chem. Eng. Data* **47**(6): 1379–1383.
23. Gruber, P., 2009. Solubilized ibuprofen. US20090175940 A1.
24. Gustafsson, T., Sörensen, H., Pontén, F., 2012. Development of a continuous flow scale-up approach of reflux inhibitor AZD6906. *Org. Process Res. Dev.* **16**(5): 925–929.
25. Hilal, S.H., Karickhoff, S.W., 2003. Verification and validation of the SPARC model (No. EPA/600/R-03/033 (NTIS PB2004-101168)). U.S. Environmental Protection Agency (EPA).
26. Hopkin, M.D., Baxendale, I.R., Ley, S.V., 2010. A flow-based synthesis of imatinib: the API of Gleevec. *Chem. Commun.* **46**(14): 2450–2452.
27. Kockmann, N., Gottsponer, M., Zimmermann, B., Roberge, D.M., 2008. Enabling continuous-flow chemistry in microstructured devices for pharmaceutical and fine-chemical production. *Chem.-Eur. J.* **14**(25): 7470–7477.
28. Kopetzki, D., Lévesque, F., Seeberger, P.H., 2013. A continuous-flow process for the synthesis of artemisinin. *Chem.-Eur. J.* **19**(17): 5450–5456.
29. LaPorte, T.L., Wang, C., 2007. Continuous processes for the production of pharmaceutical intermediates and active pharmaceutical ingredients. *Curr. Opin. Drug Di. De.* **10**(6): 738–745.
30. Lee, H., Chen, R., Lee, Y., Yoo, S., Lee, C., 2009. Essential roles of CKI $\delta$  and CKI $\epsilon$  in the mammalian circadian clock. *P. Natl. Acad. Sci. U.S.A.* **106**(50): 21359–21364.
31. Mascia, S., Heider, P.L., Zhang, H., Lakerveld, R., Benyahia, B. et al., 2013. End-to-end continuous manufacturing of pharmaceuticals: integrated synthesis, purification, and final dosage formation. *Angew. Chem. Int. Ed.* **52**(47):12359–12363.
32. Morgan, S, Grootendorst, P., Lexchin, J., Cunningham, C., Greyson, D., 2011. The cost of drug development: a systematic review. *Health Policy* **100**(1): 4–17.
33. UK Office for National Statistics. 2011. *Business Enterprise Research and Development*, 2011. Statistical Bulletin.
34. Plumb, K., 2005. Continuous processing in the pharmaceutical industry – changing the mind set. *Chem. Eng. Res. Des.* **83**(A6): 730–738.
35. Posner, G.H., Kahraman, M., 2003. Organic chemistry of vitamin D analogues (deltanoids). *Eur. J. Org. Chem.* **2003**(20): 3889–3895.
36. Qian, Z., Baxendale, I.R., Ley, S.V., 2010. A continuous flow process using a sequence of microreactors with in-line IR analysis for the preparation of n,n-diethyl-4-(3-fluorophenylpiperidin-4-ylidenemethyl)benzamide as a potent and highly selective  $\delta$ -opioid receptor agonist. *Chem.-Eur. J.* **16**(41): 12342–12348.
37. Quann, E.J., Khwaja, F., Zavitz, K.H., Djakiew, D., 2007. The aryl propionic acid R-flurbiprofen selectively induces P75NTR-dependent decreased survival of prostate tumor cells. *Cancer Res.* **67**(7): 3254–3262.
38. Ravina, E., Kubinyi, H., 2011. *The Evolution of Drug Discovery*. Wiley.
39. Richardson, G.S., Roehrs, T.A., Rosenthal, L., Koshorek, G., Roth, T., 2002. Tolerance to daytime sedative effects of H1 antihistamines. *J. Clin. Psychopharm.* **2**(5): 511–515.
40. Rihani, D.N., Doraiswamy, L.K., 1965. Estimation of heat capacity of organic compounds from group contributions. *Ind. Eng. Chem. Fund.* **4**(1): 17–21.
41. Ritter, S.K., 2013. Reducing environmental impact of organic synthesis. *Chem. Eng. News.* **91**(15): 22–23.

42. Rivard, A.L., Steer, C.J., Kren, B.T., Rodrigues, C.M.P., Castro, R.E. et al., 2007. Administration of tauroursodeoxycholic acid (TUDCA) reduces apoptosis following myocardial infarction in rat. *Am. J. Chinese Med.* **35**(2): 279–295.
43. Roberge, D.M., Zimmermann, B., Rainone, F., Gottsponer, M., Eyholzer, M., Kockmann, N., 2008. Microreactor technology and continuous processes in the fine chemical and pharmaceutical industry: is the revolution underway? *Org. Process Res. Dev.* **12**(5): 905–910.
44. Schaber, S.D., Gerogiorgis, D.I., Ramachandran, R., Evans, J.M.B., Barton, P.I., Trout, B.L., 2011. Economic analysis of integrated continuous and batch pharmaceutical manufacturing: a case study. *Ind. Eng. Chem. Res.* **50**(17): 10083–10092.
45. Sheldon, R.A., 2012. Fundamentals of green chemistry: efficiency in reaction design. *Chem. Soc. Rev.* **41**(4): 1437–1451.
46. Skrzecz, A., Shaw, D., Maczynski, A., 1999. IUPAC-NIST solubility data series 69. Ternary alcohol-hydrocarbon-water systems. *J. Phys. Chem. Ref. Data* **28**(4): 983–1235.
47. Snead, D.R., Jamison, T.F., 2013. End-to-end continuous flow synthesis and purification of diphenhydramine hydrochloride featuring atom economy, in-line separation, and flow of molten ammonium salts. *Chem. Sci.* **4**(7): 2822–2827.
48. Tambornini, L., Romano, D., Pinto, A., Bertolani, A., Molinari, F., Conti, P., 2012. An efficient method for the lipase-catalysed resolution and in-line purification of racemic flurbiprofen in a continuous-flow reactor. *J. Mol. Catal. B – Enzym.* **84**: 78–82.
49. Tu, Y., 2011. The discovery of artemisinin (qinghaosu) and gifts from Chinese medicine. *Nat. Med.* **17**(10): 1217–1220.
50. Venturoni, F., Gioiello, A., Sardella, R., Natalini, B., Pellicciari, R., 2012. Continuous flow synthesis and scale-up of glycine- and taurine-conjugated bile salts. *Org. Biomol. Chem.* **10**(20): 4109–4115.
51. Wechter, W.J., Leipold, D.D., Quiggle, D.D., McCracken, J.D., Murray, E.D., Loughman, B.E., 2000. R-Flurbiprofen (E-7869), a chemopreventive and treatment of cancer. *Inflammopharmacology* **8**(2): 189–206.
52. Wei, Z.-Y., Brown, W., Takasaki, B., Plobeck, N., Delorme, D. et al., 2000. N,n-diethyl-4-(phenylpiperidin-4-ylidenemethyl)benzamide: a novel, exceptionally selective, potent  $\delta$  opioid receptor agonist with oral bioavailability and its analogues. *J. Med. Chem.* **43**(21): 3895–3905.
53. White, N.J., 2008. Qinghaosu (artemisinin): the price of success. *Science* **320**(5874): 330–334.
54. Wittig, R., Lohmann, J., Gmehling, J., 2003. Vapor-liquid equilibria by UNIFAC group contribution. 6. Revision and extension. *Ind. Eng. Chem. Res.* **42**(1): 183–188.

## APPENDIX

### A1. Reaction kinetics

The differential rate equation for a second order reaction between two substances A and B (first order in A as well as in B) is written with respect to reactant A as follows:

$$-\frac{dC_A}{dt} = kC_A C_B \quad (\text{A1})$$

The following nonlinear equation is obtained by integration:

$$\ln \frac{C_B C_{A,0}}{C_A C_{B,0}} = k(C_{B,0} - C_{A,0})t \quad (\text{A2})$$

A plot of  $\ln \frac{C_B C_{A,0}}{C_A C_{B,0}}$  against time results in a linear correlation, given that the initial concentrations of the reactants ( $C_{A,0}$  and  $C_{B,0}$ ) are generally not equal. The reaction rate constant  $k$  can be determined from the slope of the plot, which is equal to  $k(C_{B,0} - C_{A,0})t$ .

For the first PFR, the reaction rate equation for the key organic feedstock (IBB) is:

$$-r_{IBB} = k_1 C_{IBB} C_{CH_3CH_2COOH} \quad (\text{A3})$$

This can be written as:

$$-r_1 = k_1 (C_{IBB,0} - C_{IBB,0} X_{IBB}) (C_{C_2H_3COOH,0} - C_{IBB,0} X_{IBB}) \quad (A4)$$

The integral form is used to evaluate the required residence time and reactor volume:

$$\tau_{IBB} = \frac{C_{IBB,0}}{k_2} \int_0^{X_{IBB,f}} \frac{dX_{IBB}}{(C_{IBB,0} - C_{IBB,0} X_{IBB}) (C_{C_2H_3COOH,0} - C_{IBB,0} X_{IBB})} \quad (A5)$$

For the second PFR, the reaction rate equation for organic intermediate **2** is:

$$-r_2 = k_2 C_2 C_{PhI(OAc)_2} \quad (A6)$$

Organic intermediate **2** and  $PhI(OAc)_2$  react in a 1:1 stoichiometric ratio, however the feed is not equimolar:

$$-r_2 = k_2 (C_{2,0} - C_{2,0} X_2) (C_{PhI(OAc)_2,0} - C_{2,0} X_2) \quad (A7)$$

The integral form is again used to evaluate the required residence time and reactor volume:

$$\tau_2 = \frac{C_{2,0}}{k_2} \int_0^{X_{2,f}} \frac{dX_2}{(C_{2,0} - C_{2,0} X_2) (C_{PhI(OAc)_2,0} - C_{2,0} X_2)} \quad (A8)$$

For the third PFR, the reaction rate equation is:

$$-r_3 = k_3 C_3 C_{OH^-} \quad (A9)$$

Potassium hydroxide (KOH) is in excess, so this is a pseudo-first order reaction rate equation:

$$-r_3 = k' C_3 \quad (A10)$$

where the pseudo-first order reaction constant  $k'$  is:

$$k' = k_3 C_{OH^-} \quad (A11)$$

The integrated form is used to evaluate the required residence time and reactor volume:

$$\tau_3 = \frac{1}{k' C_{3,0}} \int_0^{X_{3,f}} \frac{dX_3}{(1 - X_3)^2} \quad (A12)$$

## A2. Enthalpy of fusion

The total phase change entropy  $\Delta_0^{T_{fus}} S_{tpce}$  is considered a valid approximation for the entropy of fusion ( $\Delta S_{fus}$ ) and is estimated using equation (A27) for aliphatic and benzenoid aromatic hydrocarbons:

$$\Delta_0^{T_{fus}} S_{tpce} \cong \Delta S_{fus} \quad (A13)$$

$$\Delta_0^{T_{fus}} S_{tpce} = \sum_i n_i G_i + n_{CH_2} C_{CH_2} G_{CH_2} \quad (A14)$$

$$C_{CH_2} = \begin{cases} 1.31 & , \text{ if } n_{CH_2} \geq \sum_i n_i \text{ and } i \neq CH_2 \\ 1.00 & , \text{ otherwise} \end{cases} \quad (\text{A15})$$

Parameter  $G_i$  indicates a group and its associated additivity value, and  $n_i$  is the number of occurrences of  $G_i$ . Group coefficients  $C_i$  are used (i.e.  $n_i G_i C_i$ ) where there are multiple occurrences, with the exception of methylene ( $CH_2$ ) groups, where the group coefficient value is 1.31 if the number of consecutive  $CH_2$  groups exceeds the number of other groups and 1.0 otherwise. The calculation requires tabulated values of  $G_i$  and  $C_i$  (Chickos and Acree, 1999).

With  $\Delta_0^{T_{fus}} S_{tpce}$  taken to be a good estimate of the entropy of fusion, the total phase change enthalpy  $\Delta_m^{T_{fus}} H_{tpce}$  is estimated as follows (assuming no additional solid phase changes):

$$\Delta_m^{T_{fus}} H_{tpce} = \Delta_0^{T_{fus}} S_{tpce} T_{fus} \quad (\text{A16})$$

The total phase change enthalpy is similarly considered a valid approximation of the enthalpy of fusion,  $\Delta H_{fus}$ :

$$\Delta H_{fus} \cong \Delta_m^{T_{fus}} H_{tpce} \quad (\text{A17})$$

### A3. Solubility

The UNIFAC activity coefficient for molecule  $i$  ( $\gamma_i$ ) is the sum of a combinatorial ( $\gamma_i^c$ ) and a residual component ( $\gamma_i^r$ ) (Fredenslund et al., 1975):

$$\ln \gamma_i = \ln \gamma_i^c + \ln \gamma_i^r \quad (\text{A18})$$

The combinatorial component is calculated via the UNIQUAC model (Abrams and Prausnitz, 1975):

$$\ln \gamma_i^c = \ln \frac{\phi_i}{x_i} + \frac{z}{2} q_i \ln \frac{\theta_i}{\phi_i} + L_i - \frac{\phi}{x_i} \sum_j x_j L_j \quad (\text{A19})$$

where  $\phi_i$  and  $\theta_i$  are the molar weighted segment (similar to volume) and area fractional components, respectively. Furthermore,  $L_i$  is a UNIFAC compound parameter defined by  $r_i$ ,  $q_i$  and  $z$ :

$$\phi_i = \frac{x_i r_i}{\sum_j x_j r_j} \quad (\text{A20})$$

$$\theta_i = \frac{x_i q_i}{\sum_j x_j q_j} \quad (\text{A21})$$

$$L_i = \frac{z}{2} (r_i - q_i) - (r_i - 1), \quad (z = 10) \quad (\text{A22})$$

Parameters  $r_i$  and  $q_i$  are calculated from contributions of  $R$  and  $Q$  (the volume and surface area parameters for each functional group, respectively) according to their occurrence  $\nu_k$  on the molecule:

$$r_i = \sum_k \nu_k^{(i)} R_k \quad (\text{A23})$$



$$q_i = \sum_k v_k^{(i)} Q_k \quad (\text{A24})$$

The residual component is calculated as follows:

$$\ln \gamma_i^r = \sum_k v_k^{(i)} [\ln \Gamma_k - \ln \Gamma_k^{(i)}] \quad (\text{A25})$$

where  $\Gamma_k$  and  $\Gamma_k^{(i)}$  is the residual group activity coefficient of group  $k$  in reality and in a reference solution of pure substance  $i$ , respectively. Both  $\Gamma_k$  and  $\Gamma_k^{(i)}$  are calculated via the following equation:

$$\ln \Gamma_k = Q_k \left[ 1 - \ln \sum_m \theta_m \psi_{mk} - \sum_m \frac{\theta_m \psi_{km}}{\sum_n \theta_m \psi_{nm}} \right] \quad (\text{A26})$$

where parameters  $\theta_m$  (summation of the area fraction of group  $m$  over all different groups) and  $x_m$  (the mole fraction of group  $m$ ) are defined as follows:

$$\theta_m = \frac{Q_m x_m}{\sum_n Q_n x_n} \quad (\text{A27})$$

$$x_m = \frac{\sum_j v_m^j x_j}{\sum_j \sum_n v_n^j x_j} \quad (\text{A28})$$

Parameter  $\Psi_{mn}$  is calculated using a modified Arrhenius equation and represents a measure of interaction energy between groups:

$$\Psi_{mn} = \exp \left[ -\frac{U_{mn} - U_{nm}}{RT} \right] \quad (\text{A29})$$

Tabulated data for values of  $Q$ ,  $R$  and group interactions have been published for a multitude of groups (Wittig et al., 2003).

The solubility (mole fraction at saturation  $x_i^{sat}$ ) of components in solvent mixtures is calculated using the following equation (Gracin et al., 2002):

$$a_i^s = x_i^{sat} \gamma_i^{sat} = \exp \left[ \frac{\Delta H_{fus}}{R} \left( \frac{1}{T_{fus}} - \frac{1}{T} \right) \right] \quad (\text{A30})$$

Here,  $H_{fus}$  is the enthalpy of fusion and  $T_{fus}$  is the melting point of pure solid solute  $i$ . The solid activity ( $a_i^s$ ) is first estimated via the exponential of Eq. (A30), and the activity coefficient obtained from Eq. (A18) is used to obtain the solubility of the solid in the solvent mixture ( $x_i^{sat}$ ) via Eq. (A30). An iterative process is required for Eq. (A30), as the activity coefficient depends on mole fraction.