An Automated Dual Modeling Approach to Accelerate Reaction Analysis and Optimization
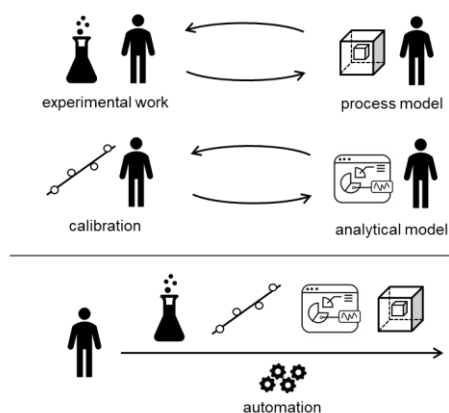
Peter Sagmeister,[†,‡] Lukas Melnizky,[†,‡] Jason D. Williams,*[,†,‡] C. Oliver Kappe*[,†,‡]

†Institute of Chemistry, University of Graz, NAWI Graz, Heinrichstrasse 28, 8010 Graz, Austria

‡Center for Continuous Flow Synthesis and Processing (CC FLOW), Research Center Pharmaceutical Engineering GmbH (RCPE), Inffeldgasse 13, 8010 Graz, Austria

ABSTRACT: In modern pharmaceutical research, the demand for expeditious development of synthetic routes to active pharmaceutical ingredients (APIs) has led to a paradigm shift towards data-rich process development. Conventional methodologies encompass prolonged timelines for reaction and analytical model developments. Both method developments are separated into different departments and often require an iterative process to optimize the models. Addressing this issue, we introduce an innovative dual modeling approach, seamlessly integrating the development of a Process Analytical Technology (PAT) strategy with reaction optimization. This integrated approach is exemplified in diverse amidation reactions and the synthesis of the API benznidazole. The platform, characterized by a high degree of automation and minimal operator involvement, achieves PAT calibration through a "standard addition" approach. Dynamic experiments are executed to screen a broad process space and gather data for fitting kinetic parameters. Employing a Julia-coded software program facilitates rapid kinetic parameter fitting and in-situ optimization within minutes. This highly automated workflow not only expedites the understanding and optimization of chemical processes, but also holds significant promise for time and resource savings within the pharmaceutical industry.

INTRODUCTION

The modern synthetic chemist has a wider range of techniques at their disposal than ever before for developing synthesis routes to complex molecules.[1] These technologies drive innovation in advanced materials, agrochemicals and life-saving drugs, allowing products to reach the market in timelines that were previously thought to be unachievable.[2–5] Developed manufacturing processes, particularly for the synthesis of active pharmaceutical ingredients (APIs), must consist of a scalable reaction, but also control strategies and analytical methods to ensure suitable product quality. Modeling and simulation play an increasingly important role in development workflows, granting a holistic and data-rich overview of the system in question.

Digital and data-rich methods are becoming omnipresent in process development workflows. During preliminary optimization efforts, synthetic route scouting is often augmented by computer-aided synthesis planning (CASP) tools, including machine learning (ML) algorithms.[6] Reaction optimization within the chosen route can be performed using a myriad of data-driven approaches, including closed-loop optimization.[7] Further detailed reaction optimization and modeling then provides additional understanding toward a robust procedure. Finally, by combining reaction models with reactor models, the entire process stream can be simulated, resulting in a digital twin that can be used to track the state of the product at any point in space and time.

Due to this wide variety of data-driven tools, process development is an interdisciplinary exercise, requiring expertise from synthetic chemists, analytical chemists, data scientists, chemical engineers and more (**Figure 1, top**). The synthetic route, process models, and control strategy are typically developed individually, resulting in overlapping tasks and a significant extent of repetition. Although continuous processing (flow chemistry) offers many benefits, such as enhanced product quality, increased efficiency, and cost savings,[8–12] developing such processes requires even more of a multidisciplinary team of scientists and engineers from various backgrounds. By unifying and
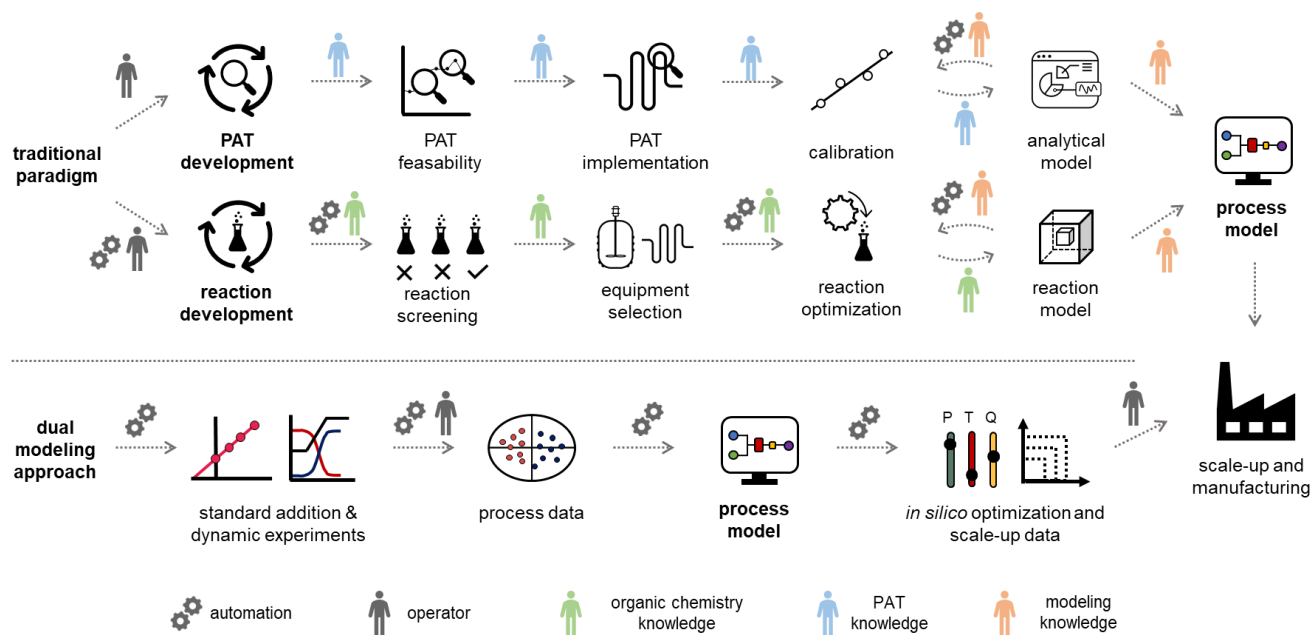
Figure 1 Top: Example of a traditional optimization approach for process development. Typically, the reaction and the PAT models are developed separately by several team members with different background knowledge. **Bottom:** The dual modeling approach, utilizing a high degree of automation to combine PAT and reaction optimization, resulting in a full process model.

automating the experimental and modeling steps in this workflow, we anticipate significant savings in time and materials, alongside enhanced control of the resulting processes.

Despite the range of advanced methods available, many reactions, particularly in flow, are still optimized by changing one variable at a time (OVAT) with offline analysis of the results.[7] Automation, data-rich experiments and real-time process analytics can drastically accelerate development and collect process-relevant data quickly and efficiently.[13–15] Self-optimization and automated Design of Experiments (DoE) have been proven to be excellent methodologies for screening a broad process space and finding optimal process conditions.[16–21] The material consumption for these optimization methodologies can be drastically reduced by the implementation of droplet or slug flow platforms.[22–25]

The aforementioned methods generally rely on single data points, which are measured once the reactor has reached steady state. Utilizing dynamic experiments can drastically accelerate the collection of dense datasets. In these experiments, ramps over time are executed to explore the design space. This dynamic change is followed with process analytical technology (PAT) and the collected data must be corrected to match the corresponding input setpoints. Several approaches have been described in the literature in recent years, including ramps for a single process variable (often residence time)[26–28] or multiple process variables simultaneously.[29–37]

The implementation of PAT and data processing models is still a substantial hurdle for non-specialist chemists.[38–40] This typically requires expert knowledge to develop and calibrate data processing (chemometric) models, in a time- and resource intensive procedure. Automating the workflow for collection, data processing and model generation

from dynamic experiment data will drastically accelerate process development timeframes.

The use of data-rich workflows also plays an important role in advancing sustainability and green practices across the pharmaceutical industry. There exists a large number of sustainable synthesis procedures that do not receive their warranted attention, often due to a lack of data and understanding around their operation and broader applicability. An excellent example of such a methodology is the 1,5,7-triazabicyclo[4.4.0]dec-5-ene (TBD)-catalyzed amidation of esters,[41–44] which obviates an ester hydrolysis step and the use of a stoichiometric coupling agent.[45,46] To increase uptake of such sustainable synthesis methodologies, data-rich experimentation facilitates rapid scoping, allowing chemists to make informed, data-driven decisions. Real-time monitoring and advanced data analytics can then accelerate process optimization and the development of science-based control strategies, which secure medicine supplies and prevent waste in manufacturing.

To fulfill all of these requirements, we endeavored to establish a platform that calibrates process analytics, collects optimization data in a dynamic flow regime, and parameterizes a process model for scale-up in less than a working day (<8 h, **Figure 1, bottom**). This optimization platform has a high degree of automation, but still includes the operator in the loop. The calibration of PAT is performed using a standard addition approach[47] in continuous flow. The standard addition data trains a partial least squares (PLS) regression model. This model is then applied to the collected dynamic data in a broad process space. The processed analytical data is fed with all process inputs into a software program which is coded in Julia. The program is capable of fitting kinetic parameters and creating a process model which can be used for *in-silico* optimization and scale-up. The developed workflow is showcased with a

2

broad scope of amidation reactions between esters and amines and the two-step synthesis of the API benznidazole (alkylation followed by amidation).

RESULTS AND DISCUSSION

Platform and Approach

In traditional process development, reaction models are built mainly using data from offline analysis (**Figure 1, top**). This workflow includes the initial reaction screening and feasibility study, reaction optimization and the reaction model development. In parallel to (or even after) the reaction development, a PAT and control strategy is often developed to control critical quality attributes (CQAs) in the final process. This encompasses the feasibility and implementation of PAT within the process, calibration of the process analytics, and the real-time processing of raw data into meaningful information (e.g., species concentrations and performance metrics). Multiple scientists with individual expertise will work on both workflows to combine them into a final process model. The digitalization of chemical development work provides opportunities to automate parts of the aforementioned workflows. This ultimately will save resources and accelerate the development of processes.

The proposed "dual modeling approach", detailed in this manuscript, combines the development of a PAT strategy and reaction optimization in one synergistic workflow. This highly digitalized workflow is separated into two parts, which are simultaneously executed on one platform (**Figure 1, bottom**). The platform can be comprised of any flow chemistry equipment and different process analytics. A degree of automation is necessary to at least send new setpoints to the flow equipment over pre-determined time intervals. To follow this precise workflow, two valves should be present in the setup, allowing the reactor to be bypassed, dosing the product to the reaction stream before or after the reactor. This allows for maximum flexibility in terms of quick PAT calibration and investigation of the reaction design space.

A standard addition approach in continuous flow is utilized to calibrate the PAT. Similar to a batch standard addition,[47–49] different concentration levels are measured by varying the pump flow rates accordingly. Bypassing the reactor enables the rapid achievement of steady-state conditions and fast acquisition of the concentration levels. Product calibration is achieved by continuously spiking a known concentration of product to the reactor outlet.

The standard addition is utilized to assign concentration values to each recorded spectrum. The labelled spectra with concentration values are then used to train and validate a PLS regression model. The PLS model is generated within this workflow through either automated means using a Python program or by an operator using chemometrics software (in our case PEAXACT software from S-PACT).

Dynamic experiments are automatically executed after the calibration stage to explore the reaction design space. Different reaction conditions are explored with single parameter or multiple parameter ramps. Steady states between the dynamic ramps facilitate the data evaluation and act as points of validation in the results.

Process models from the acquired dynamic experiments are generated using software coded in the programming language Julia. Julia focuses on scientific computing, data analysis and statistical programming.[50] It is an open-source programming language and utilizes precompiled code, which accelerates the execution of the code. The developed software fulfills the data handling, the definition of
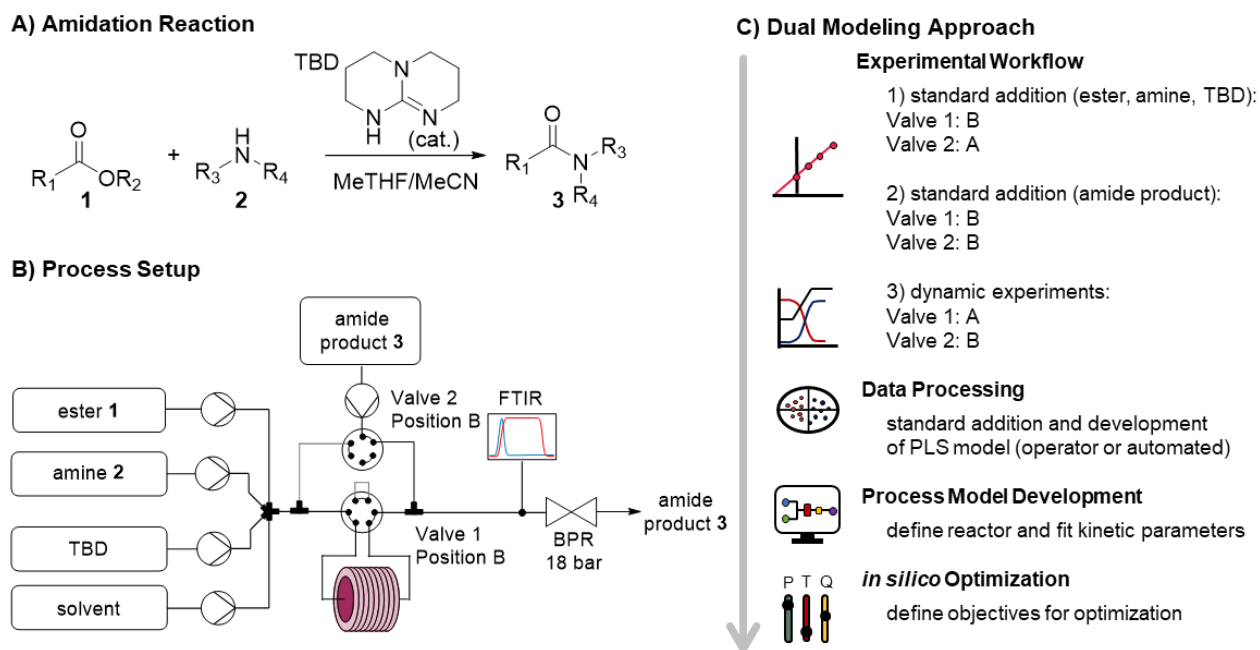

Figure 2 A) General reaction scheme of the investigated TBD-catalyzed amidation reaction. B) The optimization platform used to perform the dual modeling approach. C) The automated workflow of the dual modeling approach.

3

the chemical reactor, the identification of kinetic parameters, and performs *in-situ* optimization with the generated process model within minutes.

The input parameters from the reactor and measured concentration values from the dynamic experiments are read by the software. The operator has only to define the reaction network and the reactor configuration in the software. The kinetic parameters are fitted by employing a cost function, which compares the measured results to the computed results from the kinetic parameters and inputs. A global optimization algorithm (NLopt-BOBYQA)[51] is employed to find the global best fit for the kinetic parameters. The global optimum is then refined using a simplex algorithm (Nelder-Mead). The obtained kinetic parameters are combined into a process model, which can be used for in-situ optimization, such as identifying the Pareto front of competing objectives or simulating any other point of interest.

Dual Modeling for Sustainable Amidation

The formation of an amide bond is one of the most important reactions in the pharmaceutical industry.[45,46,52] TBD has been shown to be an effective catalyst in facilitating the amidation of esters by primary and secondary amines (**Figure 2A**), including demonstration on >10 kg scale.[41–44,53] Despite its tremendous potential in improving sustainability, TBD-catalyzed amidation has received surprisingly little attention. Accordingly, this provides a perfect opportunity to incorporate a data-rich workflow for reaction optimization and understanding of substrate effects. Continuous flow processing can act as an enabling technology to apply high pressure and temperature and intensify the reaction. In our investigation, combinations of 5 esters (**1a-e**) and 3 amines (**2a-c**) have been examined with this dual modeling approach.

The continuous flow setup was comprised of commercially available flow equipment, automated and controlled with an orchestrating software (**Figure 2B**). In order to minimize the void volume of the transfer lines, polytetrafluoroethylene (PTFE) tubing with an inner diameter of 0.3 mm was utilized. The feed solutions of ester, amine, TBD, solvent, and the amide product were prepared in MeTHF/MeCN (9+1 v/v) and introduced by HPLC pumps. The ester, amine, TBD and solvent feeds were mixed in a 7-port mixing unit (with 2 blocked ports). An automatic 6-
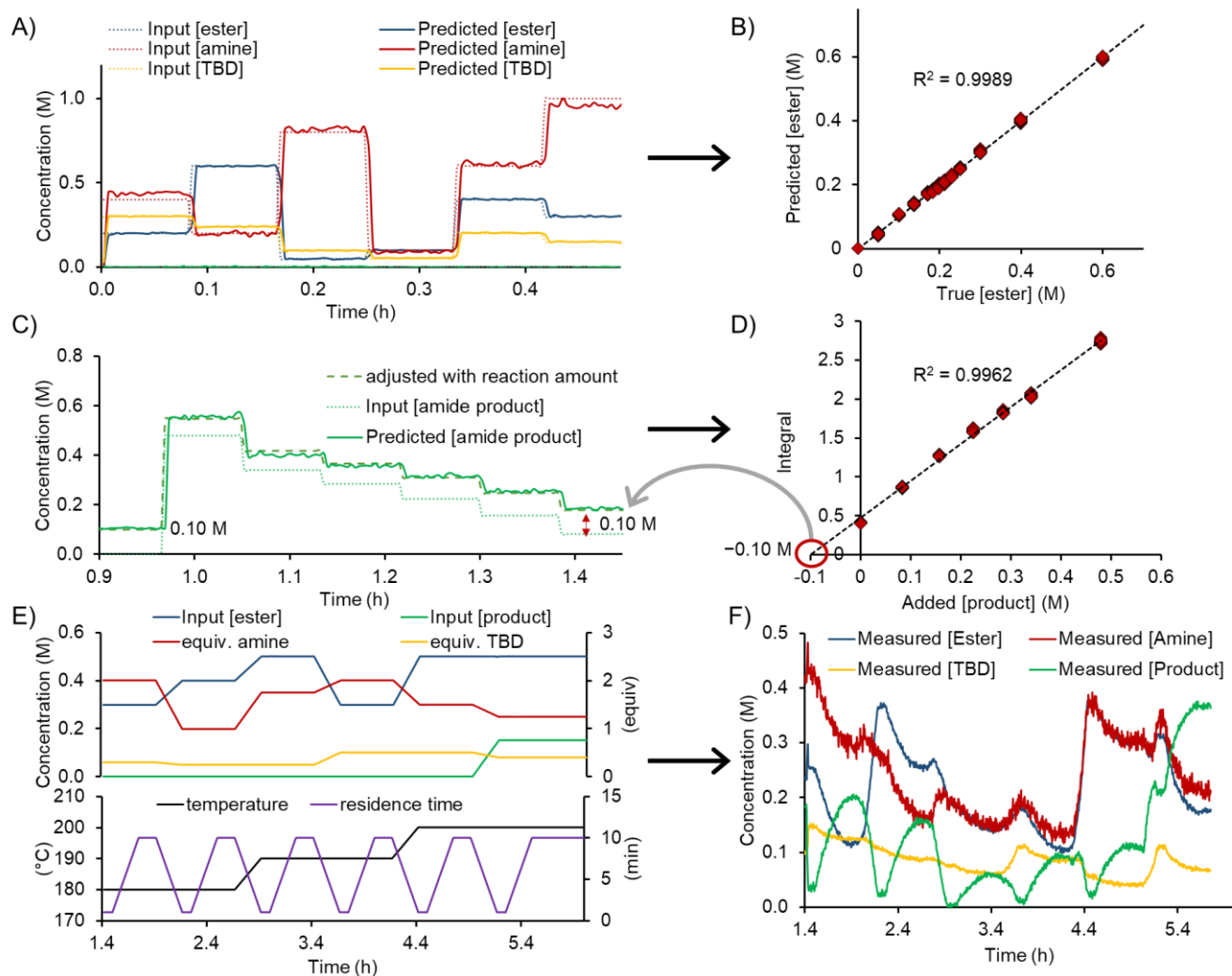

Figure 3. A) 6 different concentration levels for ester, amine, TBD to obtain data for training the PLS model. B) Representative graph for the amide product standard addition. C) Pre-determined process parameter ramps over time for the dynamic experiments.

4

port valve (valve 1) was utilized to either (position A) direct the reaction mixture through a stainless steel reactor coil (5.67 mL, 0.8 mm i.d.) placed on a heating block or (position B) guide it directly to a Fourier transform infrared spectrometer (FTIR). Another automatic 6-port valve (valve 2) was used for calibration purposes, to dose the amide product either directly before the reactor (position A) or after the reactor (position B) via a T-piece. The system was pressurized with a membrane-based back-pressure regulator (set to 19 bar).

The experimental workflow for the calibration of the FTIR was automatically executed and a standard addition approach was used. First, 6 different levels of ester, amine, TBD and solvent were generated and directly analyzed by FTIR, without passing through the reactor (**Figure 3A**). The product standard addition was performed by allowing a certain reaction composition to pass through the reactor and dosing 6 different concentrations of product to the outcoming reaction mixture of the reactor (**Figure 3B**). The assignment of the concentration values with the standard addition is outlined in detail in the supporting information.

The amine pump was also turned off during the standard addition process to investigate if an intermediate species between the ester and the TBD could be observed, with or without passing through the heated reactor coil (Valve 1 in position A or B). In each case, no significant decrease in the ester concentration was observed, implying that the active intermediate does not form (in appreciable quantities) in the absence of an amine nucleophile. This is in agreement with previous reports, whereby the formation of an acylated TBD intermediate is reversible[54] and has only been isolated in reaction with specific lactone substrates.[55]

Two methods of building a PLS model were investigated: (1) automated PLS model generation in Python and (2) manual PLS model generation in PEAXACT (by an operator). Typically, 12 different concentration levels from the standard addition were used as training data for the PLS model. Both approaches provide similar results, however the automated approach reduced the PLS model generation time from an hour (depending on experience of the user) to a few minutes. Applying a first derivative pretreatment to the raw spectra provided improved results compared to second derivative or no derivative. Multiple PLS models were generated for the different substrate combinations with both processing approaches. The obtained results (ranks for the PLS model, RMSEs and mass balances) from both approaches provided similar results and are explained in detail in the supporting information.

The root mean square error of cross validation ($RMSE_{CV}$) for the different esters **1**, amines **2**, TBD and the formed amide products **3** were between 18 – 68 mM, 30 – 214 mM, 9 – 50 mM, and 12 – 42 mM, respectively. In most dynamic flow experiments, the relative error for the mass balance of ester and amine was below 10%, which is within the uncertainty of the PLS model. In almost all experiments a higher error in mass balance was observed for TBD. Predictions for the amine typically had more noise compared to product and ester predictions. This might be explained by limited characteristic bands in the spectrum (only two bands observed in the fingerprint region below 800 cm^{-1}) for the amine substrates.

The dynamic experiments were executed as a set of 6 different experimental points (**Figure 3C**) to explore a broad chemical space. Three different temperature levels (180 °C, 190 °C, and 200 °C) were investigated, with two dynamic ramp sets for each. The ester concentration was varied between 0.3 M and 0.5 M. The equivalents of amide and TBD were varied between 1.0 – 2.0 and 0.25 – 0.50, respectively. Product inhibition was investigated in the last experiment by adding 0.15 M (0.3 equivalents) of product to the reaction mixture. At the beginning of the dynamic experiment the system was equilibrated to steady state (for 5 min) with a residence time of 1 min. Then, a residence time ramp over 15 min was executed, from the shortest residence time of 1 min to the longest of 10 min. At the end of this ramp, the system was allowed to reach steady state again for 10 min. Then, a change for the next reaction conditions was achieved by dynamically changing all reaction parameters (residence time, temperature, equiv amine, equiv TBD, and concentration of ester) over 15 min. The resulting experimental program consisted of 4.5 h, resulting in a total time of 6 h, when combined with the initial calibration/standard addition levels.

The input data and processed FTIR data from the dynamic runs were passed on to the developed software in Julia (**Figure 4**). The timestamps of the process parameters and measurement points were automatically interpolated to a uniform time axis. The operator must simply define the reactor setup (volume) in the software. In our case the reactor setup was comprised of two different segments. The first segment was the heated reactor part and the second segment was the tubing to the measurement point. Another input for the software is the reaction network. The change of reaction component concentrations, as a function of the
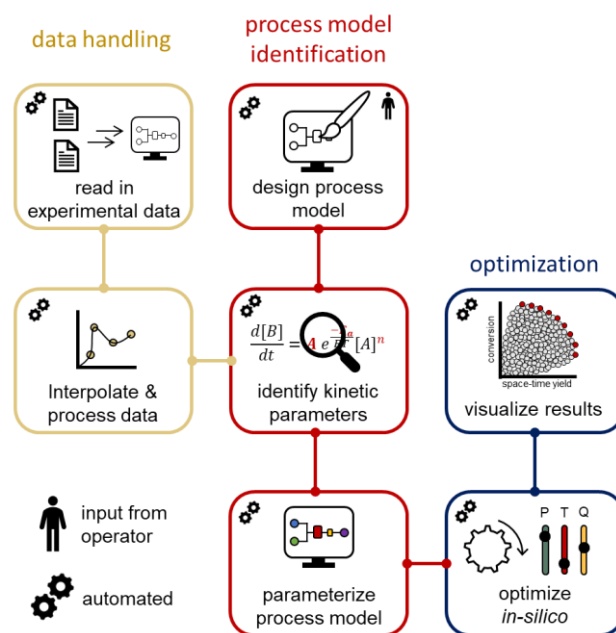

Figure 4. Schematic overview of the software developed in Julia. Data is automatically read in, interpolated and processed. Based on a process model (designed by an operator), the kinetic parameters are identified and a model parameterized. Using this model, in silico optimization can be performed.

5

transport and the kinetics, can be simulated with differential equations. Different global optimizers have been tested resulting the NLopt-BOBYQA algorithm as the most favored one in terms of performance and speed (see supporting information). Additionally, the refinement of the optimized parameters is achieved using a Nelder-Mead algorithm. The boundaries for the refinement are within +5% to -5% of the obtained global optimum.

The reaction kinetic parameters were fitted with a single- and two-step reaction network (**Figure 5A**). The developed Julia software is able to fit reaction orders, the Arrhenius pre-exponential factor (A) and activation energy (E_a). For better comparison between each amidation reaction a two-step reaction network was used and the reaction orders were fixed to 1 for ester, amine, TBD and intermediate. In first reaction the ester is activated by TBD to form an active *N*-acyl intermediate. This intermediate then reacts with the amine to corresponding the amide product, regenerating TBD. In total 10 different amidation reactions were investigated, including a comparison of different benzoic

esters: methyl vs ethyl vs isopropyl ester (**Figure 5B**). The activation energy (E_a1) and pre-exponential factor (A_1) for the rate determining step (activation of ester) is displayed in **Figure 5C**.

The activation energy for the reaction of pyridine-based ester **1a** with the different amines **2a**, **2b** and **2c** is 21.0 kJ/mol, 29.1 kJ/mol, and 24.6 kJ/mol, respectively. This reactivity follows a trend for amine nucleophiles: primary > cyclic secondary > acyclic secondary (**2a** > **2c** > **2b**). The same order of reactivity was also observed for carbocyclic ester **1b.** Activation energies of 26.7 kJ/mol, 34.0 kJ/mol, and 31.9 kJ/mol were fitted for its reaction with **2a, 2b** and **2c**, respectively. Although the reactivity order of cyclic vs acyclic secondary amines follows known nucleophilicity scales, the primary amine would be expected to be less reactive.[56] This implies that other factors, such as steric effects or hydrogen bonding capabilities, have an increased influence in this reaction.
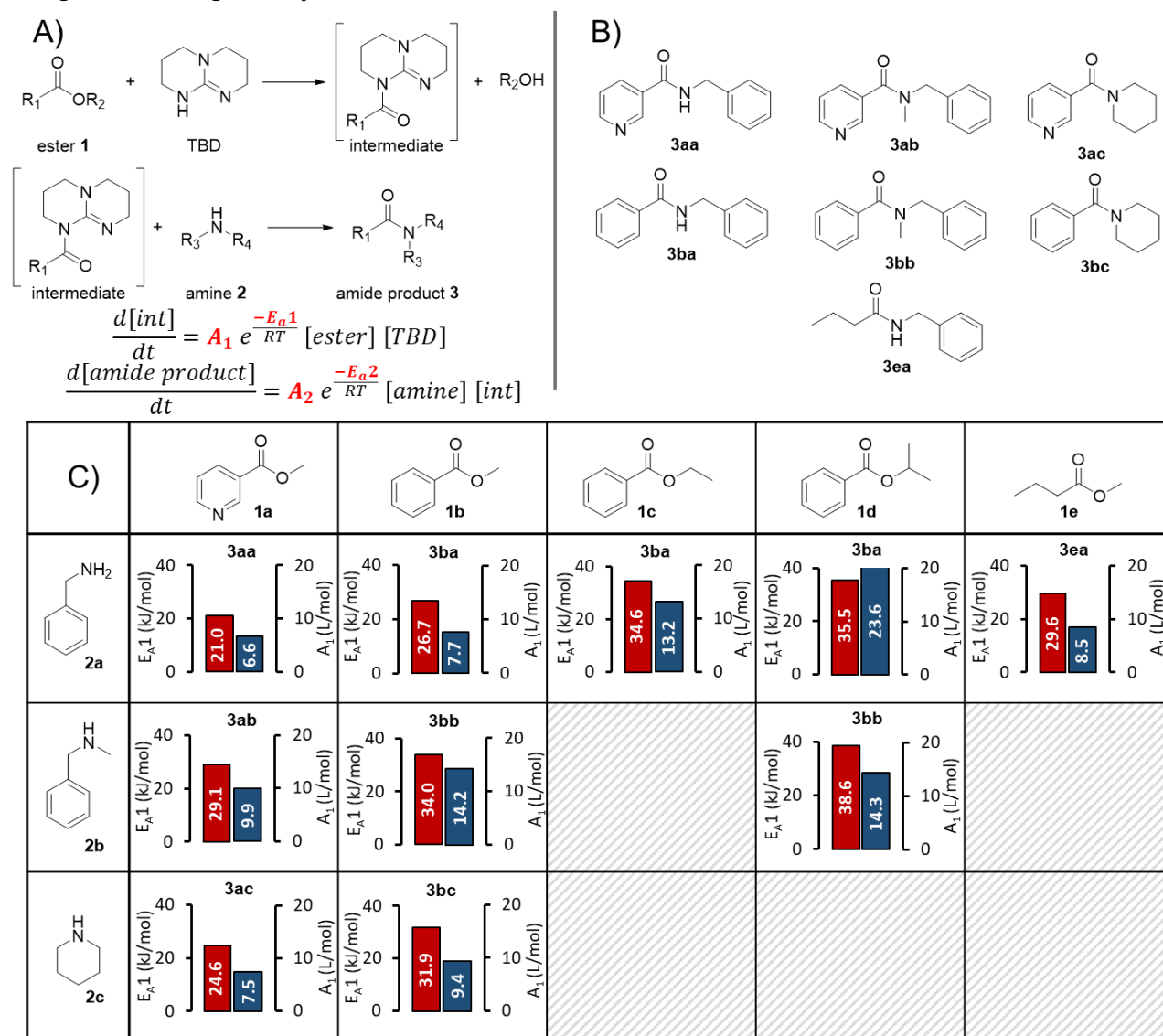

Figure 5 A) Reaction network and equations used to fit kinetic parameters for the amidation reactions. B) The formed amide products in the amidation reactions. C) Pre-exponential factors in blue and activation energies in red for the first step of the amidation reactions with different ester and amine starting materials.

6

The electron-deficient nature of heterocyclic ester **1a** significantly enhances the reaction rate compared to carbocyclic ester **1b**. The increased reactivity was reflected in the lower activation energies for **1a** in reactions with all three different amines. This is in agreement with the reported large positive Hammett substituent constant for a 3-pyridyl substituent (σ = +0.55).[57]

A slightly lower activation energy of 26.7 kJ/mol was obtained for **1b** compared to **1e** (29.6 kJ/mol) in the reaction with **2a**, quantifying the difference in reactivity between aromatic and aliphatic esters. By means of comparing the influence of the alkoxy group on ester reactivity, the methyl moiety (**1b**) has a lower activation energy than ethyl (**1c**), which is lower than the sterically bulky isopropyl group (**1d**). Although this trend would be expected (methyl > ethyl > isopropyl), we have quantified the relationship between these esters. Interestingly, this study has also demonstrated that the reactivity of both the amine and ester partners are of key importance to achieving successful reaction, which is not immediately intuitive, when considering the mechanism proceeding via a TBD *N*-acyl intermediate.

Finally, we have also demonstrated that the kinetic parameter fitting is not restricted to a specific set of experimental setpoints. This was evidenced by the fact that similar kinetic parameters were obtained for the reaction of **1a** with **2a** with a dynamic experiment using different (less forcing) experimental setpoints (see supporting information).

Application to API Synthesis

To examine and demonstrate the dual modeling approach further, the two-step synthesis of an API was performed. benznidazole (**7**) is used for the treatment of Chagas disease (American trypanosomiasis) and is on the World Health Organization (WHO) list of essential medicines.[58] The API is synthesized by alkylation of 2-nitroimidazole (**4**) with ethyl bromoacetate (**5**) to form ethyl ester **6** (**Figure 6A**), followed by TBD-catalyzed amidation with benzylamine (**2a**) (**Figure 6C**). Both transformations would benefit in terms of process intensification by developing a continuous flow protocol. The alkylation step also provides an opportunity to demonstrate the utility of the dual modeling approach for a different reaction type.

The alkylation step was carried out in a similar setup as described for the aforementioned amidation reactions. The feed solutions of **4**, **5,** triethylamine (TEA) and **6** were prepared in ethanol. The solubility of **4** was increased by adding 1.1 equiv TEA to the stock solution. The reactor coil was switched to a 2.79 mL perfluoroalkoxy alkane (PFA) coil and the BPR was set to 18 bar. During the dynamic experiments for the alkylation step, the concentration of **4** was varied between 0.3 and 0.5 M and the temperature was varied between 60 and 120 °C. The loading of **5** and TEA was varied between 1.0 – 2.0 and 1.35 – 2.0 equivalents, respectively. The residence time ramps were between 1 and 10 minutes.

Using the same workflow described above, the FTIR calibration was performed by the standard addition
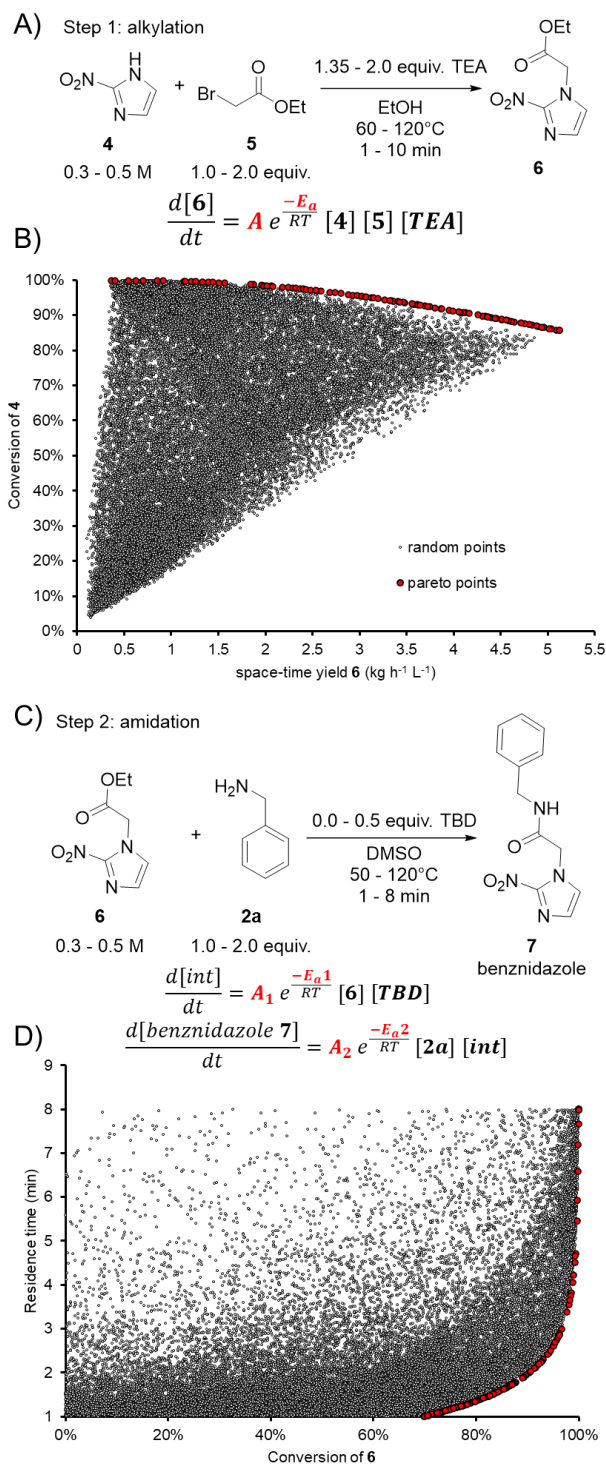

$$\frac{d[6]}{dt} = A\, e^{\frac{-E_a}{RT}}\,[4]\,[5]\,[TEA]$$

$$\frac{d[int]}{dt} = A_1\, e^{\frac{-E_a 1}{RT}}\,[6]\,[TBD]$$

$$\frac{d[benznidazole\ 7]}{dt} = A_2\, e^{\frac{-E_a 2}{RT}}\,[2a]\,[int]$$

Figure 6 A) Reaction scheme for the alkylation step yielding ethyl ester **6**. B) NSGA-II optimization to determine the Pareto optimal front, balancing the trade-off between conversion and space-time yield. Pareto optimal front depicted by red points, contrasting with gray points (from random inputs). C) Reaction scheme for the amidation yielding benznidazole **7**. D) Visualization of residence time vs. conversion. Pareto optimal front depicted by red points, contrasting with gray points (from random inputs).

approach, prior to performing dynamic experiments. The auto-generated PLS model had a RMSE$_{CV}$ of 43 mM, 42 mM,

7

18 mM, and 46 mM for **4**, **5**, TEA, and **6**, respectively. This PLS model was applied to the dynamic run, then the process parameters and FTIR results were fed into the Julia software. The kinetic parameters were fitted as a trimolecular single step reaction, with fixed reaction orders of 1 for **4**, **5** and TEA. The fitted pre-exponential factor (A) was 1.986×10^5 L^2/mol^2 with an activation energy (E_a) of 49.5 kJ/mol. The activation energy in this case was higher than that of all previously-examined amidation reactions, meaning that the rate of the alkylation has a comparatively higher temperature-dependence.

Another advantage of building a process model, such as this, is the opportunity to perform further in silico optimization. As an example, optimization was performed using the NSGA-II algorithm to find a Pareto front of the two opposing objectives: conversion and space-time yield (**Figure 6B)**. Points in red show the Pareto front, which is a series of optimal points showing the trade-off between the two objectives. A large number of random results (gray points) were also generated to demonstrate that none of these would surpass the Pareto front. Following this optimal front, a maximum space-time yield of 3.2 kg L^{-1} h^{-1} of **6** can be achieved when ensuring conversion >95%. Should a higher conversion of >99% be required, a maximum space-time yield of 1.9 kg L^{-1} h^{-1} of **6** would be attainable. It should be noted that this type of analysis can readily be carried out for any desired combination of objectives. The corresponding process setpoints can be obtained and utilized in future experiments.

In the amidation step the feed solutions of **6**, **2a**, TBD and **7**, were prepared in DMSO, due to poor solubility in the previously-used MeTHF/MeCN solvent system. The same reactor setup was used as for the previous reaction step. The dynamic runs screened the following chemical space: 0.3 – 0.5 M of **6**, 1.0 – 2-.0 equiv **2a**, 0.0 –0.5 equiv TBD, temperature 50 –120°C, and residence time 1 – 8 min. In the last experimental point, the concentration of TBD was completely omitted to confirm that no reaction was observed.

The PLS model for the FTIR measurements to follow the dynamic runs was developed with the previously described methodology. The RMSE$_{CV}$ for the compounds **6**, **2a**, TBD, and **7** were 24 mM, 33 mM, 11 mM, and 28mM, respectively. The kinetics for the amidation step were fitted with the same reaction mechanism as described above (**Figure 5A**). The following kinetic parameters for the first step (intermediate formation) were fitted: A_1 = 7.524 L/mol and E_a1 = 14.19 kJ/mol. In the second step (formation of product), the best-fitting parameters were: A_2 = 6.659 L/mol and E_a2 = 11.5 kJ/mol.

The optimal operating area (Pareto front) was identified using an *in-silico* multi-objective optimization for residence time and conversion (**Figure 6D**). It can be observed that residence times of 4.25 min can still provide conversion >99% with a space-time yield of 1.8 kg L^{-1} h^{-1}. Increasing the space-time yield to 2.8 kg L^{-1} h^{-1} results in a decrease of residence time to 2.6 min at 95% conversion.

CONCLUSION

In summary, we have developed and demonstrated a holistic "dual modeling" approach to reaction optimization. This approach is capable of developing both an analytical model and reaction model within one experimental day. First, standard addition is utilized to collect calibration data for the FTIR and a python program automatically generates a PLS model to determine species concentrations. Dynamic flow experiments are then used to quickly cover a broad chemical space and collect high density data for reaction model development. The obtained data is fed into a software program that fits kinetic parameters for different reaction networks within minutes. Additionally, the software can use the fitted kinetic parameters and the reactor geometries to generate a digital twin of the synthesis process. This process model is utilized for *in-silico* optimization of multiple objectives, such as conversion, space-time yield or measures of environmental impact. This drastically reduces optimization timescales and saves resources during process development.

This dual modeling approach was applied to a rarely-used protocol for sustainable amidation reactions. By examining a range of different substrate combinations, insights into the different kinetics for each substrate were gained. Accordingly, a better understanding of the compatible reaction substrates can be inferred. We then expanded this approach to the two-step synthesis of the API benznidazole, where both alkylation and amidation reactions proved to be amenable to this approach. The developed reaction models were also utilized for *in silico* optimization, whereby different objectives, and the trade-offs between them, can be rapidly explored without additional experimental effort.

Following on from this initial dissemination, we expect the dual modeling approach to be utilized in a wide range of future studies. For example, the method can be used to rapidly determine the Hammett reaction constant, ρ, for new synthetic methodologies and sustainable alternatives, providing predictable applicability for different substrate classes. Furthermore, we anticipate exploration of additional PAT instruments, such as chromatographic analysis, to provide more insight into impurities and allow more precise quantification of reaction intermediates.

AUTHOR INFORMATION

Corresponding Author

* Email: jason.williams@rcpe.at, oliver.kappe@uni-graz.at

and sustainable chemistry throughout the global chemistry enterprise. The ACS GCI Pharmaceutical Roundtable is composed of pharmaceutical and biotechnology companies and was established to encourage innovation while catalyzing the integration of green chemistry and green engineering in the pharmaceutical industry. The activities of the Roundtable reflect its members' shared belief that the pursuit of green chemistry and engineering is imperative for business and environmental sustainability.

ABBREVIATIONS

API, active pharmaceutical ingredient; BOBYQA, bound optimization by quadratic approximation; BPR, back pressure regulator; CASP, computer-aided synthesis planning; CQA, critical quality attribute; DoE, design of experiments; EMA, European Medicines Agency; FTIR, Fourier transform infrared spectroscopy; HPLC, high-performance liquid chromatography; ML, machine learning; NSGA, non-dominated sorting genetic algorithm; OVAT, one variable at a time; PAT, process analytical technology; PFA, perfluoroalkoxy alkane, PLS, partial least squares; PTFE, polytetrafluoroethylene; RMSE, root mean square error; TBD, 1,5,7-triazabicyclo[4.4.0]dec-5-ene; TEA, triethylamine.

REFERENCES

(1) Ley, S. V.; Fitzpatrick, D. E.; Ingham, R. J.; Myers, R. M. Organic Synthesis: March of the Machines. *Angew. Chem. Int. Ed.* **2015**, *54*, 3449–3464.

(2) Allais, C.; Connor, C. G.; Do, N. M.; Kulkarni, S.; Lee, J. W.; Lee, T.; McInturff, E.; Piper, J.; Place, D. W.; Ragan, J. A.; Weekly, R. M. Development of the Commercial Manufacturing Process for Nirmatrelvir in 17 Months. *ACS Cent. Sci.* **2023**, *9*, 849–857.

(3) Anderson, A. S. A Lightspeed Approach to Pandemic Drug Development. *Nat. Med.* **2022**, *28*, 1538–1538.

(4) Halford, B. How to Scale Up a Lifesaving Molecule in a Matter of Months. *ACS Cent. Sci.* **2023**, *9*, 1715–1717.

(5) Halford, B. The Path to Paxlovid. *ACS Cent. Sci.* **2022**, *8*, 405–407.

(6) Coley, C. W.; Thomas, D. A.; Lummiss, J. A. M.; Jaworski, J. N.; Breen, C. P.; Schultz, V.; Hart, T.; Fishman, J. S.; Rogers, L.; Gao, H.; Hicklin, R. W.; Plehiers, P. P.; Byington, J.; Piotti, J. S.; Green, W. H.; Hart, A. J.; Jamison, T. F.; Jensen, K. F. A Robotic Platform for Flow Synthesis of Organic Compounds Informed by AI Planning. *Science* **2019**, *365*, eaax1566.

(7) Taylor, C. J.; Pomberger, A.; Felton, K. C.; Grainger, R.; Barecka, M.; Chamberlain, T. W.; Bourne, R. A.; Johnson, C. N.; Lapkin, A. A. A Brief Introduction to Chemical Reaction Optimization. *Chem. Rev.* **2023**, *123*, 3089–3126.

(8) Britton, J.; Raston, C. L. Multi-Step Continuous-Flow Synthesis. *Chem. Soc. Rev.* **2017**, *46*, 1250–1271.

(9) Plutschack, M. B.; Pieber, B.; Gilmore, K.; Seeberger, P. H. The Hitchhiker's Guide to Flow Chemistry. *Chem. Rev.* **2017**, *117*, 11796–11893.

(10) Gutmann, B.; Cantillo, D.; Kappe, C. O. Continuous-Flow Technology - A Tool for the Safe Manufacturing of Active Pharmaceutical Ingredients. *Angew. Chem. Int. Ed.* **2015**, *54*, 6688–6728.

(11) Baumann, M.; Moody, T. S.; Smyth, M.; Wharry, S. A Perspective on Continuous Flow Chemistry in the Pharmaceutical Industry. *Org. Process Res. Dev.* **2020**, *24*, 1802–1813.

(12) Breen, C. P.; Nambiar, A. M. K.; Jamison, T. F.; Jensen, K. F. Ready, Set, Flow! Automated Continuous Synthesis and Optimization. *Trends Chem.* **2021**, *3*, 373–386.

(13) Schneider, G. Automating Drug Discovery. *Nat. Rev. Drug Discov.* **2018**, *17*, 97–113.

(14) Nambiar, A. M. K.; Breen, C. P.; Hart, T.; Kulesza, T.; Jamison, T. F.; Jensen, K. F. Bayesian Optimization of Computer-Proposed Multistep Synthetic Routes on an Automated Robotic Flow Platform. *ACS Cent. Sci.* **2022**, *8*, 825–836.

(15) Sagmeister, P.; Williams, J. D.; Kappe, C. O. The Rocky Road to a Digital Lab. *Chimia* **2023**, *77*, 300–306.

(16) Schweidtmann, A. M.; Clayton, A. D.; Holmes, N.; Bradford, E.; Bourne, R. A.; Lapkin, A. A. Machine Learning Meets Continuous Flow Chemistry: Automated Optimization towards the Pareto Front of Multiple Objectives. *Chem. Eng. J.* **2018**, *352*, 277–282.

(17) Sagmeister, P.; Ort, F.; Jusner, C. E.; Hebrault, D.; Tampone, T.; Buono, F. G.; Williams, J. D.; Kappe, C. O. Autonomous Multi-Step and Multi-Objective Optimization Facilitated by Real-Time Process Analytics. *Adv. Sci.* **2022**, *9*, 2105547.

(18) Clayton, A. D.; Pyzer-Knapp, E. O.; Purdie, M.; Jones, M. F.; Barthelme, A.; Pavey, J.; Kapur, N.; Chamberlain, T. W.; Blacker, A. J.; Bourne, R. A. Bayesian Self-Optimization for Telescoped Continuous Flow Synthesis. *Angew. Chem. Int. Ed.* **2023**, *62*, e202214511.

(19) Jorayev, P.; Russo, D.; Tibbetts, J. D.; Schweidtmann, A. M.; Deutsch, P.; Bull, S. D.; Lapkin, A. A. Multi-Objective Bayesian Optimisation of a Two-Step Synthesis of p-Cymene from Crude Sulphate Turpentine. *Chem. Eng. Sci.* **2022**, *247*, 116938.

(20) Müller, P.; Clayton, A. D.; Manson, J.; Riley, S.; May, O. S.; Govan, N.; Notman, S.; Ley, S. V; Chamberlain, T. W.; Bourne, R. A. Automated Multi-Objective Reaction Optimisation: Which Algorithm Should I Use? *React. Chem. Eng.* **2022**, *7*, 987–993.

(21) Knoll, S.; Jusner, C. E.; Sagmeister, P.; Williams, J. D.; Hone, C. A.; Horn, M.; Kappe, C. O. Autonomous Model-Based Experimental Design for Rapid Reaction Development. *React. Chem. Eng.* **2022**, *7*, 2375–2384.

(22) Slattery, A.; Wen, Z.; Tenblad, P.; Pintossi, D.; Sanjose-Orduna, J.; den Hartog, T.; Noel, T. Automated Self-Optimization, Intensification and Scale-up of Photocatalysis in Flow. *Sci.* **2024**, *382*.

(23) Reizman, B. J.; Wang, Y. M.; Buchwald, S. L.; Jensen, K. F. Suzuki-Miyaura Cross-Coupling Optimization Enabled by Automated Feedback. *React. Chem. Eng.* **2016**, *1*, 658–666.

(24) Baumgartner, L. M.; Coley, C. W.; Reizman, B. J.; Gao, K. W.; Jensen, K. F. Optimum Catalyst Selection over Continuous and Discrete Process Variables with a Single Droplet Microfluidic Reaction Platform. *React. Chem. Eng.* **2018**, *3*, 301–311.

(25) Wagner, F.; Sagmeister, P.; Jusner, C. E.; Tampone, T. G.; Manee, V.; Buono, F. G.; Williams, J. D.; Kappe, C. O. A Slug Flow Platform with Multiple Process Analytics Facilitates Flexible Reaction Optimization. *Adv. Sci.* **2024**, *2308034*.

(26) Hone, C. A.; Holmes, N.; Akien, G. R.; Bourne, R. A.; Muller, F. L. Rapid Multistep Kinetic Model Generation from Transient Flow Data. *React. Chem. Eng.* **2017**, *2*, 103–108.

(27) Moore, J. S.; Jensen, K. F. "Batch" Kinetics in Flow: Online IR Analysis and Continuous Control. *Angew. Chem. Int. Ed.* **2014**, *53*, 470–473.

(28) Van Herck, J.; Junkers, T. Rapid Kinetic Screening via Transient Timesweep Experiments in Continuous Flow Reactors . *Chemistry–Methods* **2022**, *2*, 1–6.

(29) Florit, F.; Nambiar, A. M. K.; Breen, C. P.; Jamison, T. F.; Jensen, K. F. Design of Dynamic Trajectories for Efficient and Data-Rich Exploration of Flow Reaction Design Spaces. *React. Chem. Eng.* **2021**, *6*, 2306–2314.

(30) Wyvratt, B. M.; McMullen, J. P.; Grosser, S. T. Multidimensional Dynamic Experiments for Data-Rich Process Development of Reactions in Flow. *React. Chem. Eng.* **2019**, *4*, 1637–1645.

(31) Zhang, B.; Mathoor, A.; Junkers, T. High Throughput Multidimensional Kinetic Screening in Continuous Flow Reactors. *Angew. Chem. Int. Ed.* **2023**, *62*, 1–7.

(32) Waldron, C.; Pankajakshan, A.; Quaglio, M.; Cao, E.; Galvanin,

F.; Gavriilidis, A. Model-Based Design of Transient Flow Experiments for the Identification of Kinetic Parameters. *React. Chem. Eng.* **2020**, *5*, 112–123.

(33) Schrecker, L.; Dickhaut, J.; Holtze, C.; Staehle, P.; Wieja, A.; Hellgardt, K.; Hii, K. K. M. An Efficient Multiparameter Method for the Collection of Chemical Reaction Data via 'One-Pot' Transient Flow. *React. Chem. Eng.* **2023**, *8*, 3196–3202.

(34) Sagmeister, P.; Schiller, C.; Weiss, P.; Silber, K.; Knoll, S.; Horn, M.; Hone, C. A.; Williams, J. D.; Kappe, C. O. Accelerating Reaction Modeling Using Dynamic Flow Experiments, Part 1: Design Space Exploration. *React. Chem. Eng.* **2023**, *8*, 2818–2825.

(35) Silber, K.; Sagmeister, P.; Schiller, C.; Williams, J. D.; Hone, C. A.; Kappe, C. O. Accelerating Reaction Modeling Using Dynamic Flow Experiments, Part 2: Development of a Digital Twin. *React. Chem. Eng.* **2023**, *8*, 2849–2855.

(36) Martinuzzi, S.; Tranninger, M.; Sagmeister, P.; Horn, M.; Williams, J. D.; Kappe, C. O. Dynamic Experiments in Flow Accelerate Reaction Network Definition in a Complex Hydrogenation Using Catalytic Static Mixers. *React. Chem. Eng.* **2024**, *9*, 132–138.

(37) Schaber, S. D.; Born, S. C.; Jensen, K. F.; Barton, P. I. Design, Execution, and Analysis of Time-Varying Experiments for Model Discrimination and Parameter Estimation in Microreactors. *Org. Process Res. Dev.* **2014**, *18*, 1461–1467.

(38) Sagmeister, P.; Lebl, R.; Castillo, I.; Rehrl, J.; Kruisz, J.; Sipek, M.; Horn, M.; Sacher, S.; Cantillo, D.; Williams, J. D.; Kappe, C. O. Advanced Real-Time Process Analytics for Multistep Synthesis in Continuous Flow. *Angew. Chem. Int. Ed.* **2021**, *60*, 8139–8148.

(39) Morin, M. A.; Zhang, W.; Mallik, D.; Organ, M. G. Sampling and Analysis in Flow: The Keys to Smarter, More Controllable, and Sustainable Fine-Chemical Manufacturing. *Angew. Chem. Int. Ed.* **2021**, *60*, 20606–20626.

(40) Rodriguez-Zubiri, M.; Felpin, F.-X. Analytical Tools Integrated in Continuous-Flow Reactors: Which One for What? *Org. Process Res. Dev.* **2022**, *26*, 1766–1793.

(41) Lan, C. B.; Auclair, K. 1,5,7-Triazabicyclo[4.4.0]Dec-5-Ene: An Effective Catalyst for Amide Formation by Lactone Aminolysis. *J. Org. Chem.* **2023**, *88*, 10086–10095.

(42) Fritz-Langhals, E. Unique Superbase TBD (1,5,7-Triazabicyclo[4.4.0]Dec-5-Ene): From Catalytic Activity and One-Pot Synthesis to Broader Application in Industrial Chemistry. *Org. Process Res. Dev.* **2022**, *26*, 3015–3023.

(43) Weiberth, F. J.; Yu, Y.; Subotkowski, W.; Pemberton, C. Demonstration on Pilot-Plant Scale of the Utility of 1,5,7-Triazabicyclo[4.4.0]Dec-5-Ene (TBD) as a Catalyst in the Efficient Amidation of an Unactivated Methyl Ester. *Org. Process Res. Dev.* **2012**, *16*, 1967–1969.

(44) Sabot, C.; Kumar, K. A.; Meunier, S.; Mioskowski, C. A Convenient Aminolysis of Esters Catalyzed by 1,5,7-Triazabicyclo[4.4.0]Dec-5-Ene (TBD) under Solvent-Free Conditions. *Tetrahedron Lett.* **2007**, *48*, 3863–3866.

(45) El-Faham, A.; Albericio, F. Peptide Coupling Reagents, More than a Letter Soup. *Chem. Rev.* **2011**, *111*, 6557–6602.

(46) Dunetz, J. R.; Magano, J.; Weisenburger, G. A. Large-Scale Applications of Amide Coupling Reagents for the Synthesis of Pharmaceuticals. *Org. Process Res. Dev.* **2016**, *20*, 140–177.

(47) Hutchinson, G.; Welsh, C. D. M.; Burés, J. Use of Standard Addition to Quantify in Situ FTIR Reaction Data. *J. Org. Chem.* **2021**, *86*, 2012–2016.

(48) Liu, J.; Sato, Y.; Yang, F.; Kukor, A. J.; Hein, J. E. An Adaptive Auto-Synthesizer Using Online PAT Feedback to Flexibly Perform a Multistep Reaction. *Chemistry–Methods* **2022**, *2*, 1–8.

(49) Sato, Y.; Liu, J.; Kukor, A. J.; Culhane, J. C.; Tucker, J. L.; Kucera, D. J.; Cochran, B. M.; Hein, J. E. Real-Time Monitoring of Solid-Liquid Slurries: Optimized Synthesis of Tetrabenazine. *J. Org. Chem.* **2021**, *86*, 14069–14078.

(50) Bezanson, J.; Edelman, A.; Karpinski, S.; Shah, V. B. Julia: A Fresh Approach to Numerical Computing. *SIAM Rev.* **2017**, *59*, 65–98.

(51) Negri, G. H.; Cavalca, M. S. M.; de Oliveira, J.; Araújo, C. J. F.; Celiberto, L. A. Evaluation of Nonlinear Model-Based Predictive Control Approaches Using Derivative-Free Optimization and FCC Neural Networks. *J. Control. Autom. Electr. Syst.* **2017**, *28*, 623–634.

(52) Roughley, S. D.; Jordan, A. M. The Medicinal Chemist's Toolbox: An Analysis of Reactions Used in the Pursuit of Drug Candidates. *J. Med. Chem.* **2011**, *54*, 3451–3479.

(53) Levina, M. A.; Zabalov, M. V.; Krasheninnikov, V. G.; Tiger, R. P. Kinetics and Quantum Chemical Aspects of the Mechanism of the Guanidine (TBD) Catalyzed Aminolysis of Cyclocarbonate Containing Soybean Oil Triglycerides as the Model Process of Green Chemistry of Polyurethanes. *React. Kinet. Mech. Catal.* **2020**, *129*, 65–83.

(54) Kiesewetter, M. K.; Scholten, M. D.; Kirn, N.; Weber, R. L.; Hedrick, J. L.; Waymouth, R. M. Cyclic Guanidine Organic Catalysts: What Is Magic about Triazabicyclodecene? *J. Org. Chem.* **2009**, *74*, 9490–9496.

(55) Lohmeijer, B. G. G.; Pratt, R. C.; Leibfarth, F.; Logan, J. W.; Long, D. A.; Dove, A. P.; Nederberg, F.; Choi, J.; Wade, C.; Waymouth, R. M.; Hedrick, J. L. Guanidine and Amidine Organocatalysts for Ring-Opening Polymerization of Cyclic Esters. *Macromolecules* **2006**, *39*, 8574–8583.

(56) Brotzel, F.; Ying, C. C.; Mayr, H. Nucleophilicities of Primary and Secondary Amines in Water. *J. Org. Chem.* **2007**, *72*, 3679–3688.

(57) Blanch, J. H. Determination of the Hammett Substituent Constants for the 2-, 3-, and 4-Pyridyl and -Pyridinium Groups. *J. Chem. Soc. B Phys. Org.* **1966**, 937–939.

(58) Losada Galván, I.; Alonso-Padilla, J.; Cortés-Serra, N.; Alonso-Vega, C.; Gascón, J.; Pinazo, M. J. Benznidazole for the Treatment of Chagas Disease. *Expert Rev. Anti. Infect. Ther.* **2021**, *19*, 547–556.

10