Automated Optimization under Dynamic Flow Conditions

Jonathan P. McMullen[‡] and Brian M. Wyvratt[‡]

Process Research and Development, Merck & Co., Inc., 126 East Lincoln Avenue, Rahway, New Jersey, USA, 07065

ABSTRACT: Automated optimization in flow reactors is a technology that continues to gain interest in academic and industrial research. For drug substance applications, where limited material is available for extensive studies, it is imperative that the automated optimization procedure identify ideal conditions for manufacturing in a resource sparing manner. It is equally as important that these investigations provide data-rich results so that the information can be used for process understanding. Achieving these two objectives in parallel is challenging with traditional automated optimization systems that rely on steady-state data. Dynamic flow systems, which adjust process inputs in a controlled manner to collect transient reaction results, maximize reaction information content. In this work, the gains in reaction knowledge by performing the automated optimization in a dynamic flow system are demonstrated using a nucleophilic aromatic substitution as a case study. A gradient-based search algorithm is used to optimize a multi-faceted objective function that accounts for yield, material input, and productivity. The immense dataset from the automated dynamic optimization was used to establish a reaction model to provide greater insight to the reaction kinetics and selectivity.

Introduction

Flow chemistry and continuous reactors have proven to be enabling technologies for pharmaceutical reaction development and drug substance manufacturing.[1-5] For process development, the ability to perform experiments in sequential order has rendered flow reactors as powerful tools for reaction optimization.[6-10] In flow reactors, where new experimental conditions can be achieved simply by adjusting flow rates or temperature, one factor at a time (OFAT) techniques can be used to individually select the residence time, reagent equivalents, or catalyst loading to achieve a desired reaction outcome.[11-13] This strategy can be extended to multiple factors and used to explicitly enumerate the reaction performance over a defined design space.[14, 15] While these approaches provide results that are easy to visualize, a more insightful and statistically relevant approach toward reaction optimization can be achieved by applying a design of experiment strategy to flow experiments.[16] Factorial designs offer reaction sensitivity information and suggest how factors should be adjusted to improve reaction performance. As one nears the optimum, central composite designs are used to approximate the reaction response surface as a quadratic model with the proposed optimal conditions located at the model's supremum. Each of these optimization methodologies described above are iterative, where data from the prior flow experiments are used to update experimental conditions in search of the optimal conditions. As such, the reaction optimization procedure becomes laborious and increasingly complex as the number of process variables grows.

The speed and efficiency of these optimization studies can be greatly enhanced with automated optimization flow systems. Here, the reactor system is equipped with Process Analytical Technology (PAT) for reaction monitoring, automation technology to adjust reaction conditions, and a feedback algorithm to guide sequential experiments toward the optimal conditions. Early examples of this technology focused on self-optimization of single-step reactions using black-box optimization algorithms.[17-20] Growing interest in flow chemistry research and applications, along with advancements in machine learning,[21, 22] have led to the expansion of automated optimization systems.[23] Besides optimizing the continuous variables of complex reactions,[24] discrete optimization approaches have been used to identify the best solvent,[25] base,[26] and catalyst[10, 27] for reactions. By cascading flow reactors and including work-up modules in the flow system, automated systems have been used to optimize the conditions for an overall process of a multi-step synthesis.[28, 29]

All automated optimization flow reactor systems require accurate and informative experimental data to arrive at the optimal conditions. Due to limited availability of starting materials, converging on the optimal conditions in a minimum number of experiments is imperative for drug substance process development. A judicious selection among the myriad of algorithms that have been implemented in automated optimization systems can improve efficiency of the optimization investigations,[30] though this decision is also influenced by the number of variables, the constraints on the reaction system, accuracy of analytical methods, and prior knowledge of the reaction kinetics. Additionally, because the procedure is material intensive, it is unlikely that multiple automated optimization trials will be performed during development; therefore, the obtained solution must account for all important reaction features, such as yield, purity, and sustainability. Using a multi-faceted objective function, implementing an algorithm that can optimize multiple objective functions simultaneously, or using a model-based optimization approach[31, 32] ensures superior utility of the optimization results. When these techniques cannot be used directly, it is important that the automated optimization provide sufficient data to generate accurate process models in post-run processing. These models can be used to predict reaction performance, to aid in process characterization experiments, and to identify the appropriate design space for manufacturing.

Despite advancements in the efficiency of automated optimization, these systems have continued to rely on steady-state experiments. This data is collected inefficiently because the reaction material that exits the reactor as the flow systems transitions from one experiment to the next, approximately three to five reactor volumes, is discarded. Dynamic flow systems, on the other hand, have proven to be a data-rich alternative to conventional steady-state approaches by adjusting reaction inputs in a controlled manner to collect transient data.[33-36] In these dynamic systems, experimental information content is maximized as each element of the reaction outlet stream provides a new reaction data point. Linear, one-dimensional dynamic experiments, where a single variable is ramped, have shown to be efficient tools for kinetic studies.[37] More recent work has illustrated increased efficiency by performing orthogonal 2-dimensional dynamic experiments.[38] Multi-dimensional trajectories that combine linear and non-linear ramps have shown to yield higher levels of data density by efficiently transgressing large areas of the reaction design space.[39, 40]

In this work, the enhancement in data-rich experimentation from a dynamic flow system is demonstrated for an automated optimization application. The aromatic nucleophilic substitution (S_NAr) shown in Scheme 1 was selected as a case study.[41] Features of this reaction are common in drug substance applications, such as the parallel side reaction and the product degradation pathway. Here, the desired product (2) is the ortho-nitro addition of morpholine. The para-nitro isomer (3) and the overreaction double addition impurity (4) are considered impurities. To further relate this model chemistry to drug substance applications, a hypothetical scenario was created where the reaction was required to produce less than 5.0 LCAP[42] of the double addition impurity to generate acceptable quality material, mimicking situations with limited downstream purge capability. A conjugate gradient method was used to optimize an objective function that considered reaction yield, double addition impurity level, process mass intensity, and residence time. The dynamic optimization was performed from two initial conditions to better understand if the global optimum was achieved. Data from these investigations were used to estimate kinetic parameters from a mechanistic model to provide further process insight and map the design space.

Equipment & Experimental Methods

Reactor Setup

The flow platform (Figure 1) utilized in this work was designed to be reaction agnostic to afford flexibility for future applications. Substrate and reagent solutions were positioned on Mettler Toledo balances (ME4301) to monitor solution feed rates. The reactor system had three inlet feed streams that were preheated with heat exchangers, and a quench stream that entered downstream of the FTIR. Solutions were charged to the flow reactor using a manifold of high-pressure MilliGAT HF pumps (MG-2-CER-XT-NM, PTFE stator) via tubular heat exchangers (0.04" ID PFA, 25 ft, 6.2 mL each) that were submerged in a bath held at 75°C. The substrate stream (1) was mixed with the solvent stream using a tee-mixer (IDEX, PEEK, 0.050" ID,
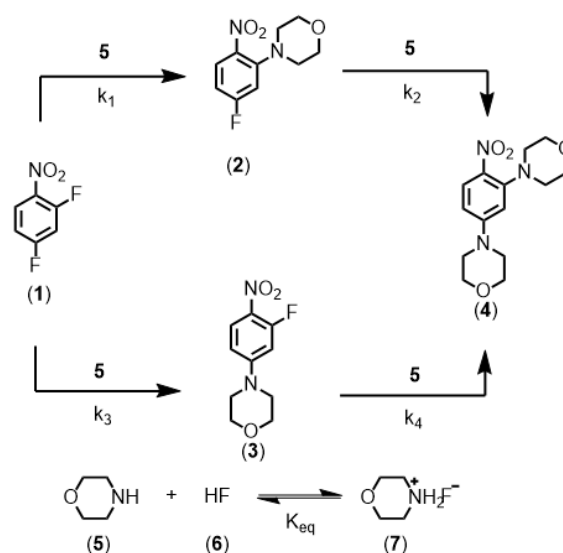
Scheme 1. S_NAr reaction between 2,4-difluoronitrobenze and morpholine

1/8") and then immediately combined with the morpholine stream using an interdigital mixer (MiChS Co., LTD.) before entering the reactor (0.04" ID PFA, 12.36 mL). Residence time distribution studies for this reactor have been reported previously.[40] The reactor temperature was controlled using a Polar Bear Plus (Cambridge Reactor Design) equipped with custom metal mandrals to secure the reactor, promote heat transfer, and improve experimental reproducibility. In-line monitoring of the reactor outlet stream was achieved with a Fourier Transform Infrared Spectroscopy (FTIR,
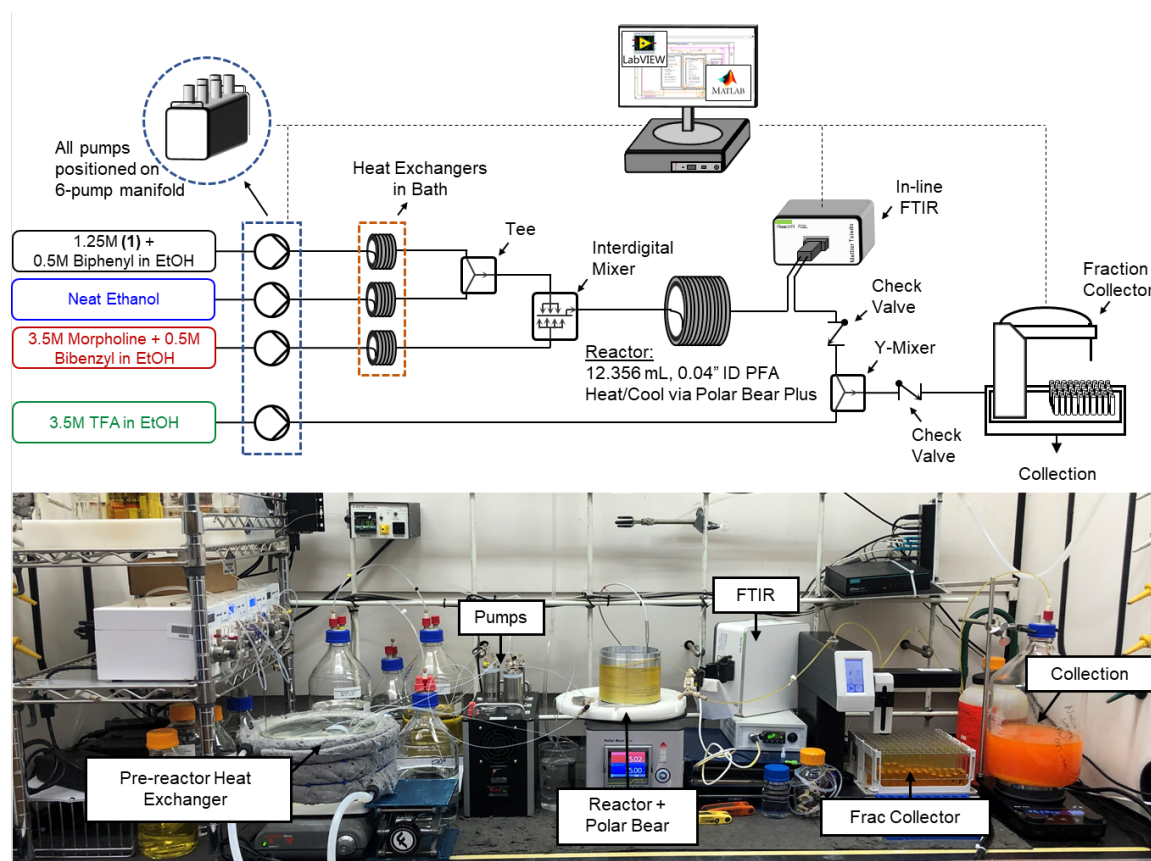
Figure 1. Reactor schematic and picture of reactor in operation.

Mettler Toledo ReactIR 702L, D-Sub Micro Flow Cell DiComp). The reaction was quenched with a trifluoroacetic acid stream in a Y-mixer (IDEX, PEEK, 0.040" ID, 1/8") prior to flowing through the fraction collector (Teledyne Isco Foxy R1) where off-line samples were obtained and analyzed using Ultra High Pressure Liquid Chromatography (UPLC, Agilent 1290 Infinity II). Check valves (IDEX, 3 psi, 0.020" ID, 1/8") were leveraged between the FTIR and quench mixer to prevent backflow.

Reactor Automation

Labview (National Instruments, 2018 SP1 Full Developer Suite) was used to automate the reactor platform, leveraging a state machine architecture from JKI (2018). LabVIEW-Matlab interfacing capabilities and the Open Platform Communication United Architecture (OPC UA) toolkit were utilized for various aspects of the automation sequence, optimization algorithm, and in-line analytics. New LabVIEW subVIs were generated or existing product architecture was leveraged to control the analytical balances, milliGAT pumps, Polar Bear Plus, and Foxy R1 fraction collector. Additionally, OPC UA functionality was leveraged to read data from the FTIR spectrometer for use in the model regressions and optimization searches to evaluate performance in real-time. Additional details are provided in the ESI.

Analytical Methods

To maximize data capture for a single experiment, both in-line analysis (FTIR) and off-line analysis (UPLC) were performed to provide orthogonal datasets for comparison and validation. The fraction collector used repeated fixed time intervals, specified at the start of the experiment for sample collection and diversion to waste. As discussed in previous work,[40] to avoid averaging a broad range of dynamic reaction data, a so-called sampling bias, and associated decreased efficiency in data collection, it is recommended to sample based on time rather than by volume. UPLC analyses were performed on an Agilent 1290 Infinity II using an analytical method from previously reported work[40] that was leveraged and modified as a fit-for-purpose method for the S_NAr chemistry investigated here. Additional details around sample preparation, method parameters, calibrations and analysis are provided in the ESI.

Infrared (IR) spectra were analyzed in the MettlerToledo iC IR software environment (version 7.1). Off-line UPLC results were used to calibrate the FTIR data to train and validate a quantitative model. This predictive model auto-selected the number of factors using spectral information in two regions, $3800 – 2500$ cm^{-1} and $1800 – 750$ cm^{-1}, due to the ATR bands for diamond.

Experimental Conditions & Procedure

All experiments were performed in the reactor setup described above. The LabVIEW code required multiple inputs prior to an automated experiment, including the reaction parameters (solution concentrations, the desired reaction concentration, reactor volume) and the optimization parameters, such as the initial point, bounding constraints, maximum search step size, and termination criteria. To set the design space, the lower and upper bounds on the residence time (τ) were set to 1 and 20 min, respectively, while the bounds for morpholine equivalents (equiv) were set from 0.5 to 4.0. These design space constraints along with the feasible range of pump flowrates helped to establish the various stock stream concentrations and reactor volume.

The stock solutions of 2,4-difluoronitrobenzene and morpholine were prepared at 1.25 M and 3.5 M, respectively, in ethanol. Solutions were prepared in advance of experiments and were found to be stable via HPLC analysis upon extended age. Biphenyl and bibenzyl were used as internal standards in the starting material and morpholine streams, respectively to confirm experimental reaction concentration and morpholine equivalents. Biphenyl (0.3M) and bibenzyl (0.4M) were included in the 2,4-difluoronitrobenzene and morpholine streams, respectively. To allow residence time and morpholine equivalents to vary independently, a stream of neat ethanol was included, affording an additional degree of freedom to the system. Additionally, to enable accurate offline analysis of the collected fractions, a 3.5M trifluoroacetic acid (TFA) in ethanol stream was used to quench the reaction downstream of the FTIR, with the TFA equivalents held constant at 4.0 throughout. All reactions were performed at 75°C. This temperature was selected from a range of values explored in previous works.[16, 41] Please refer to the ESI for additional experimental details.

Correction of Conditions

The correct interpretation of experimental results obtained under dynamic flow conditions requires a correlation between the recorded data points as a function of time on stream and the achieved reaction conditions determined by the history of input flowrates. Due to the finite holdup volume of the flow reactor, reaction conditions set at the reactor inlet (e.g., equivalents) as well as the instantaneous residence time based on flowrates are not immediately realized at the reactor effluent. Instead, the actual conditions achieved must be calculated knowing the instantaneous flowrate profile provided as inputs. This correction of experimental conditions is performed for each inline (FTIR) or offline (UPLC) analytical data point to directly correlate reaction results to experienced reaction conditions.

The realized residence time for a single reaction element eluting from the reactor is calculated from Eq. 1, where $Q(t)$ is the flowrate profile, $V_{Reactor}$ is the total reactor volume, and t_i and t_f are the initial and final reaction time, respectively. The residence time, τ, is then defined as the difference between the final and initial reaction times, as expressed in Eq. 2.

$$V_{Reactor} = \int_{t_i}^{t_f} Q(t)dt \qquad (1)$$

$$\tau = t_f - t_i \qquad (2)$$

In practice, the effluent of the reactor does not always represent the point of inline or offline sampling, particularly if there is a quench to halt reaction progress for more accurate reaction analysis. When dead volumes and quench streams are applicable, Eq. 3 is solved in parallel with Eq. 1 and Eq. 2, where t_{meas}, V_d, and $Q_{quench}(t)$ represent the time of measurement, dead volume, and quench flowrate profile, respectively.

$$V_d = \int_{t_f}^{t_{meas}} [Q(t) + Q_{quench}(t)]dt \qquad (3)$$

Eqs. 1 – 3 relate the residence time, τ, to the measurement time, t_{meas}, which can be specified to reflect the periodic sampling of either the FTIR or the fraction collector. Similarly, the morpholine equivalents can also be calculated for each t_{meas} by comparing the instantaneous flowrates of the 2,4-difluoronitrobenzene and morpholine at the start of the reaction, t_i. The experimental equivalents are given by Eq. 4, where $C_{Morpholine}$, C_{SM}, $Q_{Morpholine}$, and Q_{SM} are the morpholine stream concentration, starting material stream concentration, morpholine stream flowrate, and starting material stream flowrate, respectively.
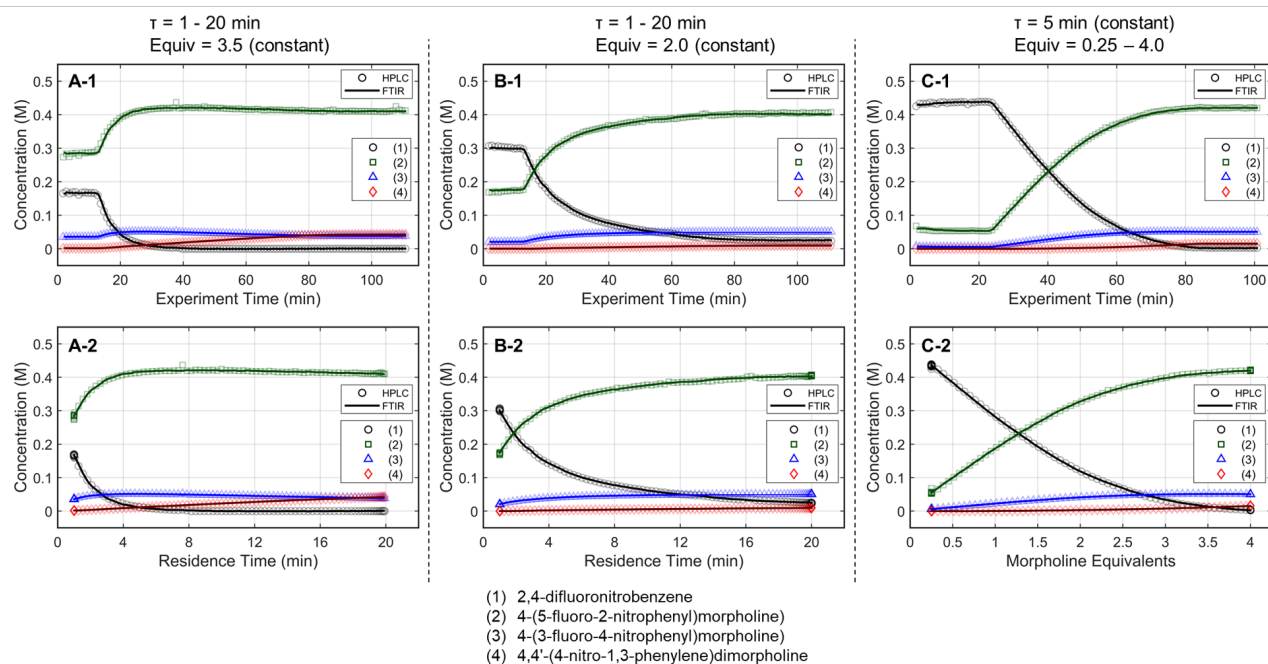
$$Morpholine\ Eq(t_{meas}) = \frac{C_{Morpholine}Q_{Morpholine}(t_i)}{C_{SM}Q_{SM}(t_i)} \qquad (4)$$

While analytical solutions exist for linear dynamic variations to residence time,[33] more complex flow rate profiles require numerical integration. Conditions for all experiments performed in this work were corrected for each measurement sample using custom Matlab code. More information is provided in the Electronic Supplemental Information.

Optimization Methodology & Results

Optimization Strategy and Objective Function Definition

The optimization strategy was assembled through preliminary dynamic experiments, consideration for real-time reaction monitoring capability, and elements of process development that are typically important to drug substance manufacturing. First, the process variables for reaction optimization were identified and limited to residence time and morpholine equivalents for this S_NAr case study. Although temperature is included often in optimization studies, dynamic trajectories in temperature can be challenging as correcting the experimental conditions requires a temperature-dependent kinetic

Figure 2. One-dimensional dynamic reaction experimental results, showing the FTIR (−) and UPLC (○,□,△,◊) results for the various species. The figures portray results from (A-1, A-2) dynamic residence time (1 – 20 min) at constant morpholine equivalents (3.5), (B-1, B-2) dynamic residence time (1 – 20 min) at constant morpholine equivalents (2.0), and (C-1, C-2) dynamic equivalents of morpholine (0.25 – 4.0) at constant residence time (5 min). Plots X-1 represent the reaction performance data as a function of experimental time, while plots X-2 show the reaction performance against the corrected reaction conditions from the 1-dimensional dynamic ramps.

model *a priori*,[38] or a sluggish trajectory to accurately estimate the reaction temperature experienced in the dynamic experiment.[39] Because the focus of this work revolves around increasing data-density of automated optimization systems and minimizing resource requirements, it was decided to not include reaction temperature in this initial demonstration.

Next, preliminary one-dimensional dynamic flow experiments were performed to characterize reaction performance over the design space and correlate results to process parameters. These initial experiments also provided an appreciation for data reproducibility and were used in the calibration of the FTIR model. Two experiments dynamically varied residence time from 1 to 20 minutes at morpholine equivalents of 2.0 and 3.5, while a third experiment dynamically varied morpholine equivalents from 0.25 to 4.0 at a constant residence time of five minutes. The dynamic transient time (60 min), temperature (75°C), and sampling strategy (one FTIR scan and one UPLC sample per minute) were held constant in each run. During these dynamic experiments, the in-line FTIR recorded spectra of the reactor outlet stream while the fraction collector aliquoted quenched reaction samples for off-line UPLC analysis. During post-run processing, the reaction conditions corresponding to each acquired data point were corrected using Eqs. (1) – (4), enabling accurate correlation between reaction performance and the two parameters of interest across the design space.

Results for the three one-dimensional dynamic experimental runs are provided in Figure 2. Of note, over 370 UPLC results were obtained from these three dynamic experiments highlighting the data-density gains associated with this experimental methodology. The reaction profiles for the various experiments are plotted against time on stream (Figure 2, *X*-1 plots) and the corrected conditions (Figure 2, *X*-2 plots). Experimental reaction data followed the expected trends in conversion and impurity generation with respect to residence time and morpholine equivalents. Excess morpholine results in faster conversion rates to the desired product (**2**), but generates stability concerns as product degradation and double addition impurity (**4**) formation occur with increasing residence time. This trend is best viewed in Figure 2, A-2, where 3.5 equivalents of morpholine resulted in near reaction completion and a product yield of approximately 84%; however, significant quantities of the double addition impurity formed with slightly longer residence times. In aggregate, these dynamic experiments reveal a balance between residence time, morpholine equivalents, yield, and purity.

The mass balance closure was greater than 97% in all samples, validating the sampling, analytical, and post-processing methods. Experimental reproducibility was assessed by comparing the reaction results at conditions that overlapped across the dynamic trajectories. There was a maximum error of 1.0% in the reaction species concentrations for the duplicate measurements at 2.0 morpholine equivalents and five-minute residence time, and for those at 3.0 equivalents and five-minute residence time. These data reinforce the accuracy and reproducibility of the platform and methodology allowing the work to reliably transition to the automated optimization runs.

These experimental data were also leveraged to calibrate a multivariate, partial least squares (PLS) FTIR model for the main reaction species and the bibenzyl internal standard. FTIR trends for the reaction species derived from the resulting model are provided against the UPLC data results in Figure 2. Parity plots of measured UPLC concentrations and the model predicted FTIR concentrations are provided in the ESI. Overall, there was excellent agreement between the FTIR models and the UPLC measurements, giving confidence that the FTIR could be used for accurate real-time reaction monitoring and provide quantitative feedback in the optimization routine.

This preliminary knowledge and analytical capability enabled definition of an objective function suitable for many drug substance manufacturing applications. Because comparable reaction performance was observed across different conditions, these experiments suggested that reaction inputs such as residence time and morpholine equivalents could be included in the optimization objective function to search for ideal operating conditions that also achieve target performance. This feature can be important in drug substance manufacturing as multiple aspects, such as process mass

intensity, robustness, and cycle time, are considered in addition to yield and purity. Using the objective function (F_{obj}) defined in Eq. 5, the optimization routine aimed to maximize yield while achieving other desired properties through several penalty functions, P_i. Here, piecewise quadratic penalty functions corresponding to P_1, P_2, and P_3 were used to maintain target thresholds on double addition impurity formation, morpholine use, and residence time, respectively. The penalty functions are presented in Figure 3 with the mathematical formulae given in the ESI. As common with many optimization routines, the maximization problem was solved by minimizing the negative objective function.

$$F_{Obj} = \left[\frac{[2]}{[1]_{initial}} - \sum P_i \right] \quad (5)$$

Optimization Procedure

Given the nature of dynamic flow experiments, where conditions are gradually adjusted to move along a trajectory, a gradient-based optimization approach was deemed the most appropriate to implement in this work. Similar to how numerical solvers perform iterative calculations along a gradient-based search direction to locate the optimum, an automated dynamic flow system continually profiles the reaction along the search trajectory. The paths taken by the gradient-based search also provide a straightforward, visual understanding of the objective function response surface. As such, the optimization procedure provides a solution and an interpretation of the process sensitivities across the search areas. Furthermore, gradient-based search algorithms are typically efficient at finding a local optimum and may minimize the material quantity required to obtain an experimental solution.

What follows is the procedure for how a gradient-based search algorithm was applied to dynamic flow experimentation. For
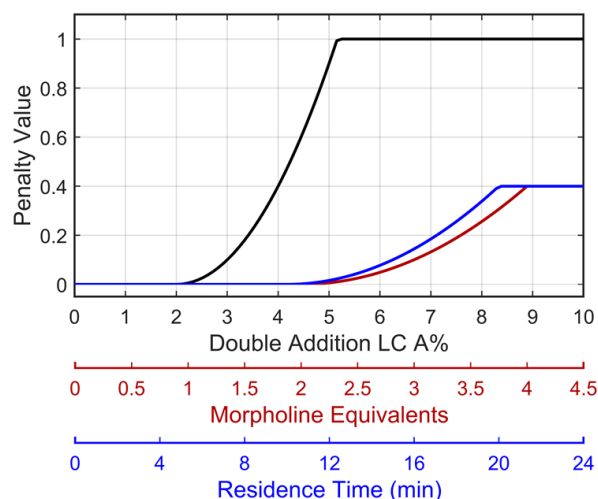

Figure 3. Piecewise quadratic penalty functions visualizing the penalty value as a function of double addition impurity levels (−), morpholine equivalents (−), and residence time (−).

reference, a comprehensive flowchart of the approach is provided in Figure 4.

Response Surface Model with Transient Datasets

Experimentally, gradients are estimated by first fitting a response surface model of the objective function over a subset of the design space and then taking the derivative with respect to the factors. A variety of experimental designs are suitable for this purpose, but
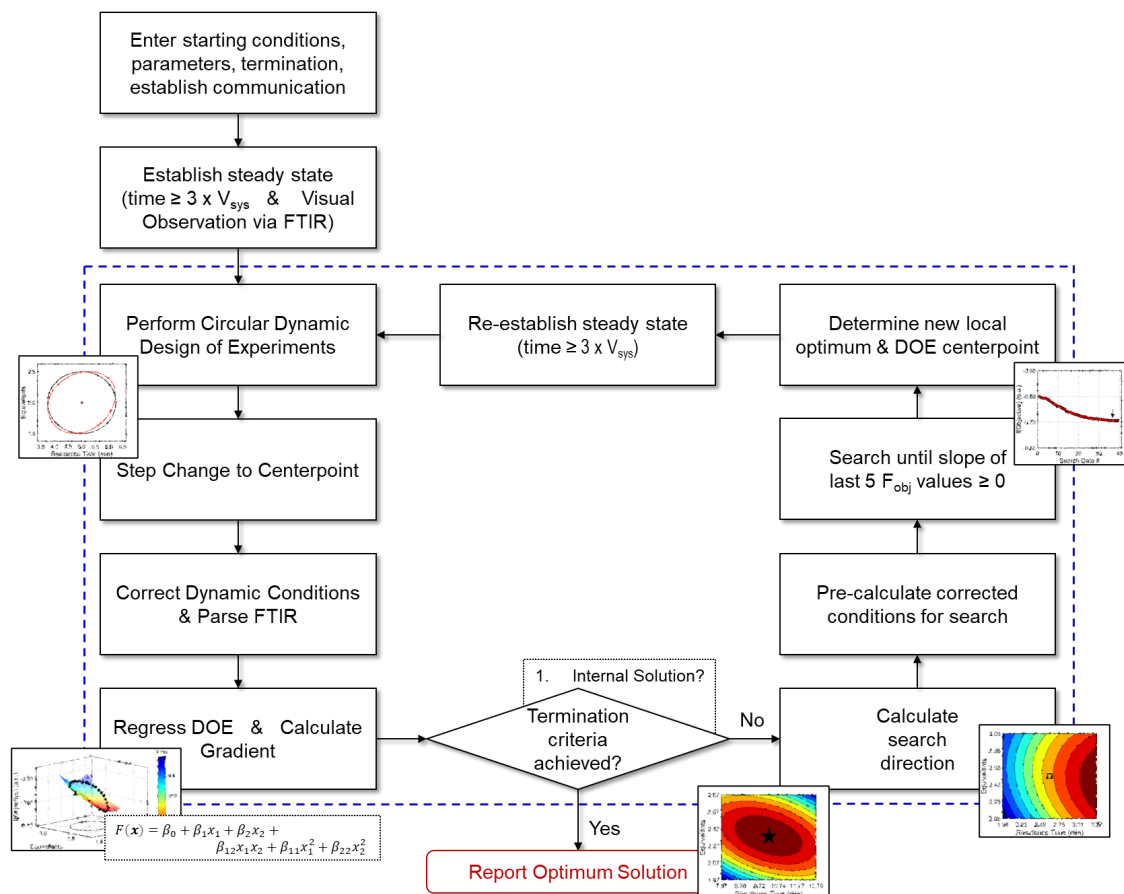
Figure 4. Flowchart describing the optimization algorithm along with associated decision points. Contents within dashed lines represent iterative, automated operations.

factorial and central composite designs are most common in drug substance process characterization. For two-factor designs, a 2^2 full-factorial design forms a square with experiments performed at the vertices and an experiment at the center to evaluate curvature. While this design is efficient for steady-state experiments, dynamic experiments can better utilize material streams and provide more dense datasets by continuously collecting data along a trajectory that outlines a standard experimental design. While a few continuous function forms were evaluated for establishing a response surface model (see ESI), a circular path inscribed within the vertices of a 2^2 factorial design was selected. To evaluate curvature over the local design space, the system underwent a step-change to the center point and collected a single steady-state data point. While this discontinuity in experimental path was undesired as a fully dynamic optimization was pursued, compared to alternative methods considered, this simplistic design offered a straightforward approach to achieving corrected conditions that more closely resembled the intended dynamic design. A more detailed discussion of these dynamic design of experiments pattern and the corresponding corrected conditions are provided in the ESI.
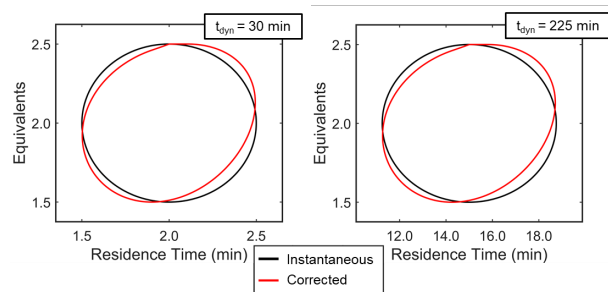
With the dynamic design of experiment structure established, other key parameters governing the operations were defined. The dynamic transient time ($t_{dynamic}$), defined as the time for dynamic operations, and the reaction residence time center point (t_{CP}) both impact how closely the corrected conditions match the targeted input conditions. Over the course of this work, the ratio $t_{dynamic}/t_{CP}$ was identified as a tuning parameter to balance experimental throughput and roundness of the corrected conditions. Simulations were performed at various $t_{dynamic}/t_{CP}$ values (see ESI), and a ratio of 15 struck the appropriate balance for this application. Example dynamic design of experiments simulations with a $t_{dynamic}/t_{CP}$ of 15 are provided in Figure 5, showing the consistency across differing center point and dynamic transient time conditions. Additionally, parameter ranges around the design center point, [τ_{CP}, Equiv$_{CP}$], were set at Equiv$_{CP}$ ± 0.5 for morpholine equivalents while the range in residence time was variable for each center point expressed as τ_{CP} ± 0.25 · τ_{CP}. These ranges were selected to ensure a measurable response in both parameters regardless of the center point. Prior to performing the dynamic design of experiments, residence time and equivalents ranges were transformed into coded variables, x1 and x2, respectively, which ranged from -1 to 1. These coded variables were leveraged through the dynamic design of experiments, regression, and gradient calculation; coded variables were then renormalized prior to the search operation.

The above dynamic design of experiments methodology was implemented in practice in the following manner:

1. Experimental factors [τ, Equiv] were converted to coded factors [x_1, x_2] using a specified or calculated center point [τ_{CP}, Equiv$_{CP}$] and the pre-defined factor ranges.

2. Performed initial steady state experiment at [x_1, x_2] = [0,1].

3. Dynamically progressed clockwise around the circular design, returning to [x_1, x_2] = [0,1].

4. Re-equilibrated system at [x_1, x_2] = [0,1].

5. Performed step change to the center point [x_1, x_2] = [0,0] conditions and flushed three system volumes to achieve steady-state to obtain reaction performance at center point

After performing the circular design of experiments and recording reaction results via the inline FTIR, the automated optimization algorithm corrected reaction conditions (Eq. 1 - 4) using imbedded Matlab code. The yield and penalty functions were evaluated at each corrected condition to calculate the corresponding objective function value. Stepwise linear regression was then used to model the objective function as a quadratic function of the two parameters, corrected residence time and morpholine equivalents, as shown in Eq. 6.

$$F(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 \qquad (6)$$

Figure 5. Simulation results showing the instantaneous (−) and associated corrected (−) conditions for two example circular design of experiments centered at [τ, Equiv] = [2 min, 2.0] (left) and [15 min, 2.0] (right). The dynamic transient time, calculated using the $t_{dynamic}/\tau_{reaction}$ ratio of 15, is shown for each simulation as well.

A quadratic model was selected as the basis for regression as this structure is often applied to response surface models around an optimum and capture curvature in systems without overfitting. When far from the from the optimum, curvature becomes less pronounced and the model reduces to linear terms.

Before implementing in a fully autonomous optimization, the above protocol was tested in a series of experiments in two different regions of the design space. These validation experiments are provided in the ESI. Results indicate that the algorithm responsible for the experimental design, results manipulation/transformation, and regression to generate a model performed adequately. These results not only validate the algorithm to generate a response surface model, but also ensure an accurate gradient is obtained from which to inform a path forward towards the reaction optimum.

Search Direction and Termination Criteria
From the objective function response surface model (Eq. 6), the gradient vector, $g(x)$, can be estimated as the partial derivative with respect to input factors (Eq. 7 − 8).

$$g(x) = \nabla F(x) = \left[\frac{\partial F}{\partial x_1}, \quad \frac{\partial F}{\partial x_2}\right] \qquad (7)$$

$$g(x) = [\beta_1 + \beta_{12} x_2 + 2\beta_{11} x_1, \quad \beta_2 + \beta_{12} x_1 + 2\beta_{22} x_2] \qquad (8)$$

The gradient can be used as the search direction for sequential experiments, as in the case for method of steepest descent. While this search direction is a logical approach for the initial search ($k=1$), successive search iterations ($k>1$) can lead to slow convergence. Alternatively, conjugate gradient techniques may offer faster convergence as these methods use a mixture of the new steepest descent direction (g^{k+1}) and the previous search direction (p^k). This weighted search prevents large shifts in the search direction, making conjugate gradient methods less prone to slow convergence for corrugated objective function responses surfaces.

In this work, the search direction vector, $p(x)$, was calculated using the Polak-Ribiere conjugate gradient (PR-CG) method. This protocol provides favorable convergence for quadratic cost functions and biases the search direction toward the direction of steepest descent when the cost function is not quadratic.[43] As defined by Eq. 9 - 10, the PR-CG method leverages objective function trends from the current gradient (g^{k+1}), the previous gradient (g^k), and the prior search direction (p^k) in calculation of the new search vector to avoid the "zig-zag" behavior in the steepest descent. In these calculations, the gradient of the current iteration, $g(x)^{k+1}$, is calculated using the current center point, x^{k+1}, and the previous gradient, $g(x)^k$, which was measured at the previous center point, x^k.

$$p^{k+1} = g(x)^{k+1} + \gamma p^k \qquad (9)$$

$$\gamma = \frac{g(x)^{k+1} \cdot (g(x)^{k+1} - g(x)^k)}{g(x)^k \cdot g(x)^k} \qquad (10)$$

With a known search trajectory, a path for dynamic experiments was calculated from the current center point. The automated optimization performed dynamic experiments along this search direction. To ensure that the corrected conditions closely followed the targeted search condition pathway while simultaneously considering experimental efficiency, a series of conditional statements were used to keep the maximum rate of change for the residence time, $\Delta \dot{t}_{max}$, at 2.5 s/s and the maximum change in morpholine equivalence, $\Delta \dot{E} q_{max}$, at 2.5 x 10^{-4} equivalents/s. As the FTIR data became available, the reaction yield and the penalty functions were computed for the objective function with the pre-corrected line search conditions. Dynamic line search experiments continued until the slope of the objective function for the five most recent experiments became non-negative. Using the slope from a series of data points made the line search termination more robust to experimental noise and error. After terminating the dynamic line search, the system identified the search conditions that provided the best objective function value. These conditions were selected as the new center point, and the system performed a new circular, dynamic design of experiment using updated coded variables, as defined above.

This iterative procedure of performing dynamic design of experiments and line searches continued until the optimal conditions resided within the response surface of the current dynamic design of experiment. These conditions are defined as a point, $\mathbf{x}^*$, which lies within the bounds of the current dynamic design of experiment range, has $\mathbf{g}(\mathbf{x}^*)$ = 0, and a positive Hessian, $H(\mathbf{x}^*) > 0$. When these conditions are satisfied, the program terminated.
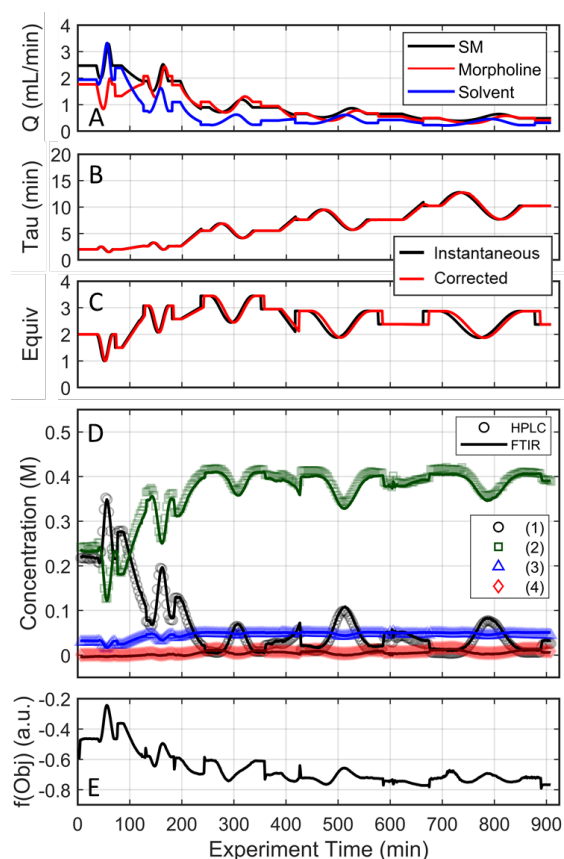
Automated Optimization Results

Building from the preliminary experimental work and methodology described above, automated dynamic experiments were performed from two initial conditions. The first investigation started with a residence time of 2 minutes and 1.5 morpholine equivalents. These initial conditions achieved incomplete reaction conversion allowing for investigation into how the system would improve the reaction performance while balancing the penalty functions. The second investigation started at a 15 minute residence time and 3.5 equivalents of morpholine. Under these conditions, yield was near maximum levels, but double addition impurity levels were elevated and the penalty functions greatly diminished the objective function value. Here, optimization progress would aid in understanding how the system would move toward preferred operating conditions when final conversion levels were already realized. Additionally, performing these independent experiments provided a sense of system robustness and increased the likelihood of identifying a global optimum. As shown in Table 1, the results from both dynamic optimization investigations converged to an optimum near 10 minutes and ~2.3 equivalents of morpholine.

As the experimental approach was consistent across the two investigative experiments sans the initial conditions, an in-depth dissection of the experimental results is limited to the run initialized at a residence time of 2 minutes and 1.5 morpholine equivalents (Investigation #1); these additional details for Investigation #2 can be

Figure 6. Conditions and experimental results from the automated optimization run starting from [2 min, 1.5 equiv] presented as a function of experimental time on stream. (A) Flowrates for the 3 streams combined to drive the reaction and control conditions: 2,4-difluoronitrobenzene (SM), Morpholine (Base) and Ethanol (Solvent). (B,C) Instantaneous/input conditions and subsequently corrected conditions as a function of time on stream. (D) UPLC and FTIR data results for the 4 species: (1) 2,4-difluoronitrobenzene, (2) 4-(5-fluoro-2-nitrophenyl)morpholine), (3) 4-(3-fluoro-4-nitrophenyl)morpholine), (4) 4,4'-(4-nitro-1,3-phenylene)dimorpholine. (E) Calculated objective function based on the FTIR results and corrected conditions as a function of experiment time.

Figure 7. Objective function plotted against the design space conditions for the optimization run starting from [2 min, 1.5 equiv].
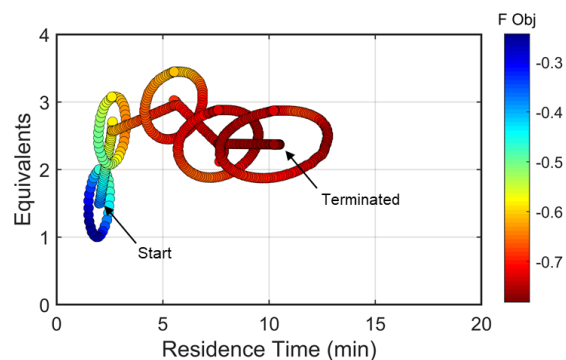
Table 1. Initial and final conditions for the two automated dynamic optimization investigations along with the calculated objective function for each conditions set

	Initial Conditions			Final Conditions		
	τ (min)	Equiv	F[#]	τ (min)	Equiv	F[#]
Investigation #1	2.0	1.5	-0.36	10.2	2.39	-0.77
Investigation #2	15.0	3.5	+0.50	10.0	2.26	-0.78

[#] Gradient algorithm operated as a minimizer with more negative responses correlating to more desirable outcomes

found in the Electronic Supplemental Information. In Investigation #1, starting with the initial conditions of two minute residence time and 1.5 equivalents, the dynamic flow system continuously adjusted flow rates (Figure 6A) to tune the residence time and morpholine equivalents (Figure 6B-C) to improve the reaction performance (Figure 6D) and objective function (Figure 6E). Closer inspection of the data reveals that the flow reaction conditions were predominantly dynamic throughout the experiment and operated under steady-state only during the step changes to the center point (Figure 6A). The corrected conditions (Figure 6B-C) matched the targeted input conditions due to the well-informed selection of dynamic transient times, as discussed previously. The FTIR model that was used for feedback was further validated through off-line UPLC analysis of collected, quenched samples (Figure 6D). Overall, the automated dynamic optimization collected 907 FTIR and 302 UPLC data points during its search for the optimal conditions.

The approach that the automated dynamic optimization system used to arrive at the optimal conditions for the first investigation can be best understood through a series of figures. The procedure began with the dynamic design of experiment around the initial conditions of two minute residence time and 1.5 equivalents of morpholine, corresponding to the circular subset of data surrounding the initial conditions in Figure 7. Note, the circular dynamic design of experiment generated in coded factors is elongated when plot with real variables. Step-wise regression was then performed on the objective function output to generate the response surface model (RSM) shown in Figure 9, RSM #1. The RSM #1 parity plot in Figure 9 shows the excellent agreement between the experimental objective function values and those predicted by the empirical model. This agreement assures that the optimization algorithm can leverage the response surface model to accurately identify the appropriate search direction.

As the contours of RSM #1 illustrate, the gradient in this region of the design space showed that improvements to the objective function could be achieved by increasing morpholine equivalents and residence time. The lack of active penalty functions in this region indicated that this improvement was driven by increased product yield with longer residence time and higher equivalents. For the first dynamic line search, the autonomous optimization procedure used the method of steepest descent to move along the gradient, starting from the center point. In Figure 7, this search corresponds to the line of dynamic experiments originating from the initial center point. The objective

function values obtained during this dynamic search are provided in Figure 8, Search #1. As this profile shows, significant improvement in the objective function was achieved during this first search and terminated once a series of successive dynamic experiments showed little improvement. Following termination, parsing the search dataset identified new local best conditions of 2.62 minutes residence time and 2.58 equivalents morpholine.

The second dynamic design of experiment was performed around these new best conditions, resulting in the circular dataset centered around 2.62 minutes residence time and 2.58 equivalents morpholine
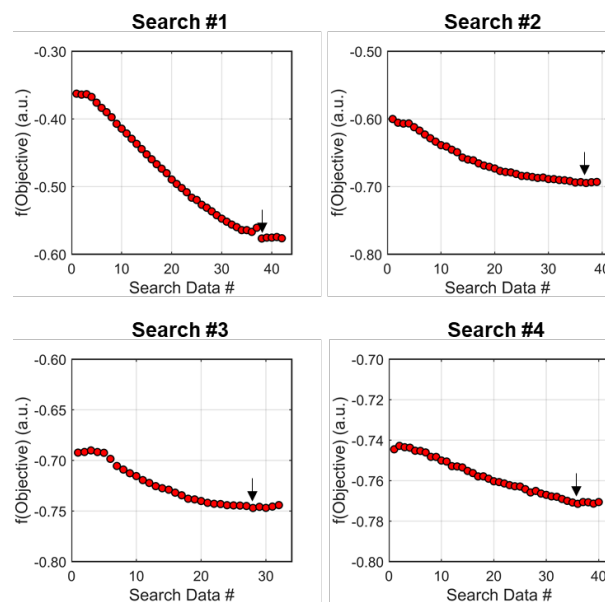
Figure 8. Progression towards minimizing objective function values during the four searches performed during Investigation #1, the run starting from [τ, Equiv] = [2 min, 1.5 equiv]. The black arrow in each figure signifies the minimum value throughout the search informing the next dynamic DOE center point conditions chosen.
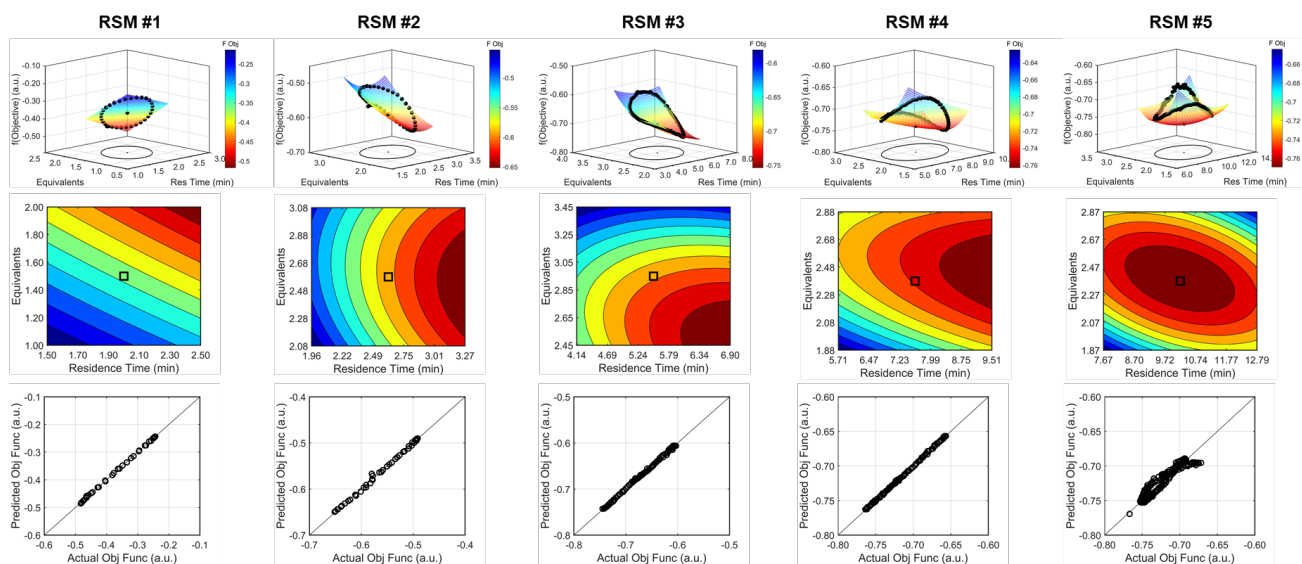

Figure 9. Response surface model results for the various DOEs performed during the Investigation #1, the run starting from [τ, Equiv] = [2 min, 1.5 equiv]. Each DOE and regression is represented within each column. (Top Row) Comparison of the experimental objective function results and the response surface model mesh against the design space region. (Middle Row) Contour plot describing the response surface model and directionality. (Bottom Row) Parity plots of the experimental data against the predicted values based on the RSM. Color Scale: Objective function improves as color transitions from blue to red.

in Figure 7. Similar to the operations described above, the algorithm used this data to regress a response surface model (Figure 9, RSM#2) which also showed excellent agreement with the experimental data (Figure 9). Here, as in subsequent iterations, the conjugate gradient method was used to calculate the search direction. Consequently, although the gradient in this region suggested that greatest improvement in the objective function would be achieved by increasing residence time alone, information from the prior design of experiments was leveraged to perform a dynamic line search with increasing residence time and equivalents. Objective function values during this dynamic line search are provided in Figure 8 and correspond to the line of dynamic experiments that originated from 2.62 minutes residence time and 2.58 equivalents morpholine in Figure 7, ultimately terminating with a new F_{Obj} best at 5.52 minutes residence time and 2.95 equivalents morpholine.

The autonomous system continued these iterative operations of dynamic design of experiments and subsequent conjugate gradient line searches for three more iterations. During these later optimization experiments, the penalty functions became active and steered the dynamic routine toward high yielding conditions with more desirable operating conditions. In the last dynamic design of experiments enveloping 10.23 minutes and 2.37 equivalents, the response surface model exhibited curvature with a calculated root within the space explored (Figure 9, RSM#5). After verifying the response surface model had a positive definite Hessian, the algorithm concluded that the calculated root was indeed an optimum and terminated. Under these conditions, the reaction achieved a 78% yield and a double addition impurity level of 1.2 area% with ~8% residual starting material and ~10% side-product impurity present. Additional detail regarding the response surface models terms, gradient vectors and search vectors can be found in the ESI.

Using similar operations, a second automated optimization (Investigation #2) was performed starting at a residence time of 15 min and 3.5 morpholine equivalents (see the results in Figure 10). Additional details and data for this investigation can be found in the ESI. In contrast to the previous optimization run, the dynamic design of experiment in this initial region captured product and isomer overreaction to the double addition impurity. Consequently, objective function values in this region were heavily penalized for double addition impurity levels as well as excessive values of morpholine equivalents and residence time. The resulting gradient calculated from the response surface model pushed the system towards shorter residence times and lower morpholine equivalents to ameliorate the penalty functions while maintaining high reaction yield. The subsequent first dynamic line search terminated at a residence time of approximately 8 minutes and 2.6 morpholine equivalents. The algorithm similarly iterated between the dynamic DOE/search operations until termination criteria were satisfied.
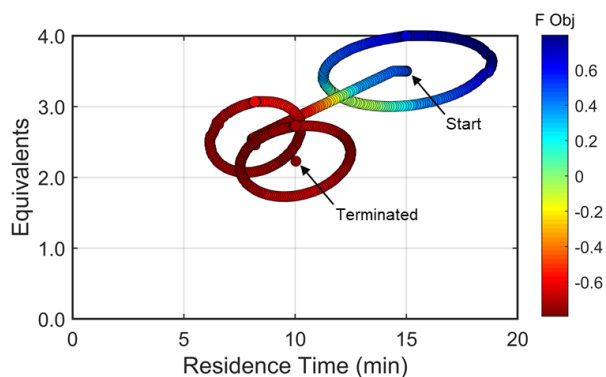
Figure 10. Objective function plotted against the design space conditions for Investigation #2, the optimization run starting from [τ, Equiv] = [15 min, 3.5 equiv].

Results from the second DOE (RSM#2) of Investigation #2 indicated that, while the optimal conditions were not found *within* the circular DOE, the predicted optimal conditions *was* within the rectangular design space subset studied during that DOE. As per the coded algorithm, no additional search was performed, but rather, the system shifted to a new center point at the hypothesized optimum and performed one additional confirmatory DOE and RSM generation. The resulting RSM identified that a minimum existed within the design space subset, reporting out an optimum at 10.0 min residence time and 2.26 equivalents morpholine (F_{Obj} = -0.78). This identified optimum was close to the final center point [10.0 min, 2.23 equiv] indicating good agreement with the predicted optimum from the second RSM. At the center point of the final DOE, the reaction produced a 79% yield to the desired product, ~7% starting material, 10% side-reaction impurity, and ~1.6A% overreaction double addition impurity. While the reaction performance along with the identified optimal conditions for the observed optimum from both optimization experiments do not exactly match, despite comparable conditions, these results are within experimental error. Optimization results were ultimately confirmed using steady state experiments, results of which can be found in the ESI.

Kinetic Modelling

At a cursory glance, results from the automated dynamic optimization provide a set of preferred operating conditions that achieve the desired reaction outcome. Additionally, the dense data set from these investigations can also be used to build a mechanistic model to gain further process insight. This model can be leveraged to better define the design space, establish process robustness, and identify process sensitivities through numerical methods. To demonstrate this potential, the automated dynamic optimization data from this S_NAr case study was used to build and regress a precise kinetic model.

The set of equations (Eqs. 11-22) describing the overall reaction network was established from the set of reactions given in Scheme 1. To facilitate kinetic modeling, the equilibrium constant, K_{eq}, for the acid neutralization was estimated from published pKa values for hydrofluoric acid (pKa = 3.8)[44] and morpholine (pKa = 8.36).[45] The rate constants were regressed by minimizing the weighted sum of squared errors using the quantitative FTIR data and corrected conditions obtained from the dynamic optimization trials. Additional modeling and regression information is provided in the ESI. The resulting best fit parameters and the associated 95% confidence parameters are listed in Table 2. Parity plots for the desired product (**2**) and the double addition impurity (**4**), key species influencing the optimization objective function, are shown in Figure 11 and highlight the excellent agreement between experimental and predicted reaction performance.

This kinetic model was used to profile the product yield, the double addition impurity formation, and the objective function (Eq. 5) over a

$$r_1 = -k_1[1][5] \tag{11}$$
$$r_2 = -k_2[2][5] \tag{12}$$
$$r_3 = -k_3[1][5] \tag{13}$$
$$r_4 = -k_4[3][5] \tag{14}$$
$$r_5 = -k_5\left([6][5] - \frac{1}{K_{eq}}[7]\right) \tag{15}$$
$$\frac{d[1]}{dt} = r_1 + r_3 \tag{16}$$
$$\frac{d[2]}{dt} = -r_1 + r_2 \tag{17}$$
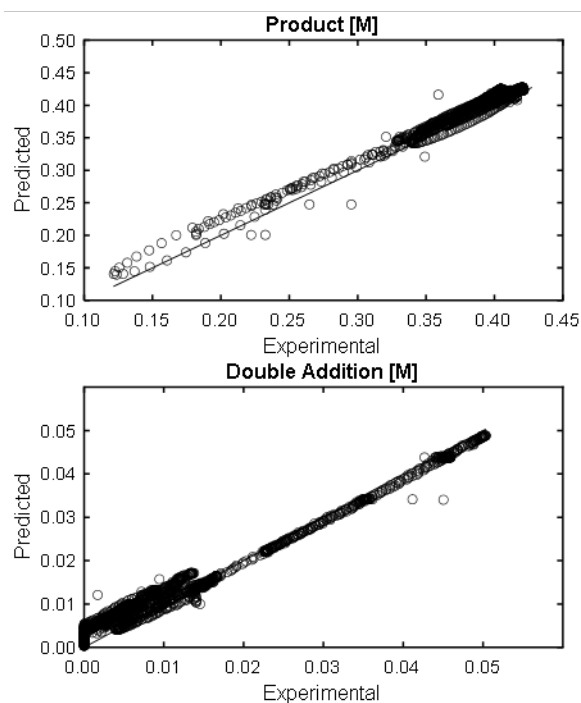$$\frac{d[3]}{dt} = -r_1 + r_4 \tag{18}$$
$$\frac{d[4]}{dt} = -r_3 - r_4 \tag{19}$$
$$\frac{d[5]}{dt} = r_1 + r_2 + r_3 + r_4 + r_5 \tag{20}$$
$$\frac{d[6]}{dt} = -r_1 - r_2 - r_3 - r_4 - r_5 \tag{21}$$
$$\frac{d[7]}{dt} = -r_5 \tag{22}$$

Table 2. Regressed rate constant estimates

k_1 (1/M·min)	$k_2 \times 10^{-3}$ (1/M·min)	$k_3 \times 10^{-2}$ (1/M·min)	$k_4 \times 10^{-2}$ (1/M·min)
$0.55 \pm 0.7\%$	$3.84 \pm 1.9\%$	$7.42 \pm 0.7\%$	$3.16 \pm 1.8\%$

Figure 11. Parity plots for Product (**2**) and Double Addition Impurity (**4**) using proposed kinetic model for predicted values.

range of residence times and morpholine equivalents (Figure 12). These modelled response surfaces illustrate the reaction features that reinforce the experimental results from the automated dynamic optimization. At short residence times and low morpholine equivalents, increasing the morpholine charge leads to the most significant gains in the objective function as more morpholine is needed to counter the neutralization of hydrofluoric acid and increase yield. High yields at minimum residence times can be achieved with super stoichiometric morpholine but similar yields at more modest morpholine levels are realized by extending the residence time. The penalty function, P_2, builds this latter preference into the automated dynamic optimization and results in the downward sloping objective values observed at the upper left portion of the design space. In the later searches of the first automated dynamic optimization trial, the algorithm reaches the optimum by balancing residence time and morpholine equivalents while maintaining high yields in the search.

At long residence times and high morpholine equivalents, such as those used in Investigation #2, the reaction exhibits high yield and elevated double addition impurity levels. These double addition levels, the excessive amounts of morpholine, and the unnecessarily long residence times cause the penalty functions to dominate the objective function value. From this design space area, the automated dynamic optimization searches for conditions that minimize these penalty terms while improving reaction yield. Additionally, as multiple penalty functions are active in this portion of the design space, the objective function improves quickly and significantly along the initial search for the second automated dynamic optimization. In fact, the algorithm quickly approached a region of optimal conditions described by the elliptical plateau of the objective function within the design space of $2.0 - 2.6$ equivalents of morpholine and residence times between 6 and 14 minutes.
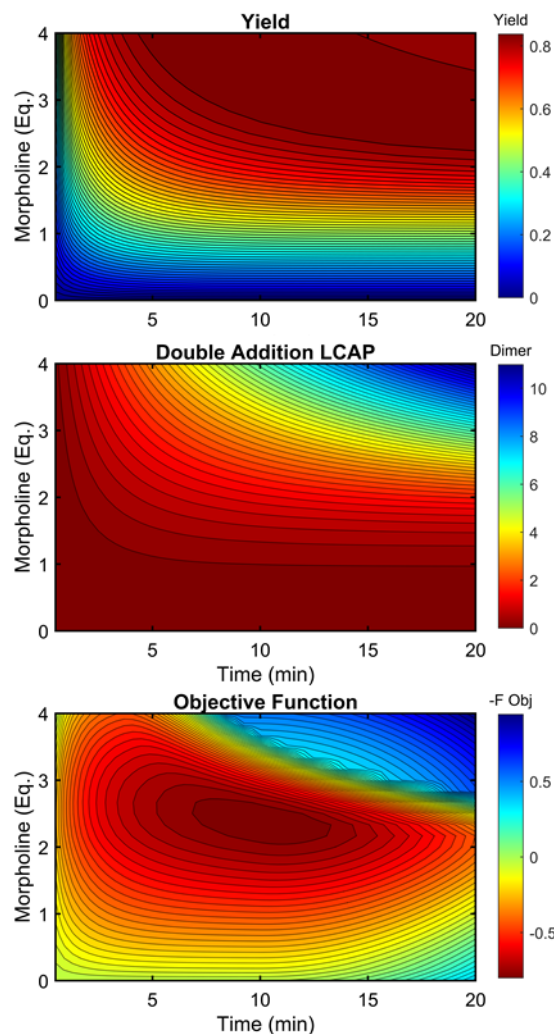
Figure 12. Contour plots of product yield, double addition impurity LCAP, and objective function as functions of residence time and morpholine equivalents using regressed kinetic parameters in **Table 2.**

Conclusions

As shown in this work, performing autonomous optimization with dynamic flow operations provides a unique, data-rich solution to a common process development problem. The data-dense, multi-dimensional trajectories obtained between the initial and optimal conditions capture the reaction sensitivities over a broad design space, information that would require extensive time and materials under traditional steady state conditions. For drug substance development and manufacturing, this wealth of multi-variate data provides invaluable process knowledge and can be used directly in establishing the operating ranges as process development timelines are further accelerated. Further process understanding is gained by leveraging the automated dynamic optimization datasets to build mechanistic or data-driven process models. As demonstrated, these models can be used to accurately predict trends in reaction product yield and impurity generation. Furthermore, as new process information becomes available, these models can be utilized to update optimal conditions given a new objective function or processing constraints.

Growing interest in autonomous optimization in flow chemistry will lead to applications in dynamic optimization for complex reactions and multi-step synthesis. As the size of the optimization problem expands, the knowledge obtained from the data-rich experimentation of

dynamic optimization will become increasingly vital for rapid and efficient process development of drug substance compounds. In these situations, alternative algorithms, such as SNOBFIT, Bayesian optimization algorithms, multi-objective optimization algorithms, or model-based optimization algorithms, may be more efficient than gradient-based approaches. Adapting these more sophisticated optimization algorithms for dynamic operations to balance efficiency and data-rich experimentation will be of significant value for future research in flow chemistry optimization. Incorporating integer optimization with dynamic flow operations may lead to data-rich routines that simultaneously optimize the selection of reagents and solvents as well as the associated operating conditions. As automated dynamic optimization becomes more prevalent in lab development, it has the potential to extend into manufacturing. At its core, automated dynamic optimization around the target manufacture behaves similarly as several well-established feedback process control strategies. Growing familiarity with feedback control systems in the lab will drive feedback control in production operations, thereby improving the robustness and productivity of these processes in the manufacturing setting.

AUTHOR INFORMATION

Corresponding Authors
‡ Email: jonathan.mcmullen@merck.com
‡ Email: brian.wyvratt@merck.com

The authors contributed equally to this work.

CONFLICTS OF INTEREST

There are no conflicts to declare.

ACKNOWLEDGMENTS

NOTES AND REFERENCES

1. McMullen, J. P.; Marton, C. H.; Sherry, B. D.; Spencer, G.; Kukura, J.; Eyke, N. S., Development and Scale-Up of a Continuous Reaction for Production of an Active Pharmaceutical Ingredient Intermediate. *Organic Process Research & Development* **2018**, *22*, 1208-1213.

2. Salehi Marzijarani, N.; Snead, D. R.; McMullen, J. P.; Lévesque, F.; Weisel, M.; Varsolona, R. J.; Lam, Y.-h.; Liu, Z.; Naber, J. R., One-Step Synthesis of 2-Fluoroadenine Using Hydrogen Fluoride Pyridine in a Continuous Flow Operation. *Organic Process Research & Development* **2019**, *23*, 1522-1528.

3. Otte, D. A. L.; Basu, K.; Jellett, L.; Whittington, M.; Spencer, G.; Burris, M.; Corcoran, E. B.; Stone, K.; Nappi, J.; Arvary, R. A.; Donoghue, D.; Ren, H.; Maloney, K. M.; Naber, J. R., Development of a Green and Sustainable Manufacturing Process for Gefapixant Citrate (MK-7264) Part 4: Formylation–Cyclization as a Flow–Batch Process Leads to Significant Improvements in Process Mass Intensity (PMI) and CO Generated versus the Batch–Batch Process. *Organic Process Research & Development* **2020**, *24*, 2478-2490.

4. Thaisrivongs, D. A.; Naber, J. R.; McMullen, J. P., Using Flow To Outpace Fast Proton Transfer in an Organometallic Reaction for the Manufacture of Verubecestat (MK-8931). *Organic Process Research & Development* **2016**, *20*, 1997-2004.

5. Bottecchia, C.; Lévesque, F.; McMullen, J. P.; Ji, Y.; Reibarkh, M.; Peng, F.; Tan, L.; Spencer, G.; Nappi, J.; Lehnherr, D.; Narsimhan, K.; Wismer, M. K.; Chen, L.; Lin, Y.; Dalby, S. M., Manufacturing Process Development for Belzutifan, Part 2: A Continuous Flow Visible-Light-Induced Benzylic Bromination. *Organic Process Research & Development* **2022**, *26*, 516-524.

6. Sagmeister, P.; Ort, F. 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. *Advanced Science* **2022**, *9*, 2105547.

7. Mateos, C.; Nieves-Remacha, M. J.; Rincón, J. A., Automated platforms for reaction self-optimization in flow. *Reaction Chemistry & Engineering* **2019**.

8. 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. *Chemical Engineering Journal* **2018**, *352*, 277-282.

9. Bédard, A.-C.; Adamo, A.; Aroh, K. C.; Russell, M. G.; Bedermann, A. A.; Torosian, J.; Yue, B.; Jensen, K. F.; Jamison, T. F., Reconfigurable system for automated optimization of diverse chemical reactions. *Science* **2018**, *361*, 1220-1225.

10. Reizman, B. J.; Wang, Y.-M.; Buchwald, S. L.; Jensen, K. F., Suzuki–Miyaura cross-coupling optimization enabled by automated feedback. *Reaction Chemistry & Engineering* **2016**, *1*, 658-666.

11. von Keutz, T.; Williams, J. D.; Kappe, C. O., Flash Chemistry Approach to Organometallic C-Glycosylation for the Synthesis of Remdesivir. *Organic Process Research & Development* **2021**, *25*, 1015-1021.

12. Uhlig, N.; Martins, A.; Gao, D., Selective DIBAL-H Monoreduction of a Diester Using Continuous Flow Chemistry: From Benchtop to Kilo Lab. *Organic Process Research & Development* **2020**, *24*, 2326-2335.

13. Kim, H.; Inoue, K.; Yoshida, J.-i., Harnessing [1,4], [1,5], and [1,6] Anionic Fries-type Rearrangements by Reaction-Time Control in Flow. **2017**, *56*, 7863-7866.

14. Musci, P.; Colella, M.; Sivo, A.; Romanazzi, G.; Luisi, R.; Degennaro, L., Flow Microreactor Technology for Taming Highly Reactive Chloroiodomethyllithium Carbenoid: Direct and Chemoselective Synthesis of α-Chloroaldehydes. *Organic Letters* **2020**, *22*, 3623-3627.

15. Usutani, H.; Tomida, Y.; Nagaki, A.; Okamoto, H.; Nokami, T.; Yoshida, J.-i., Generation and Reactions of o-Bromophenyllithium without Benzyne Formation Using a Microreactor. *Journal of the American Chemical Society* **2007**, *129*, 3046-3047.

16. Taylor, C. J.; Baker, A.; Chapman, M. R.; Reynolds, W. R.; Jolley, K. E.; Clemens, G.; Smith, G. E.; Blacker, A. J.; Chamberlain, T. W.; Christie, S. D. R.; Taylor, B. A.; Bourne, R. A., Flow chemistry for process optimisation using design of experiments. *Journal of Flow Chemistry* **2021**, *11*, 75-86.

17. McMullen, J. P.; Jensen, K. F., An Automated Microfluidic System for Online Optimization in Chemical Synthesis. *Organic Process Research & Development* **2010**, *14*, 1169-1176.

18. McMullen, J. P.; Stone, M. T.; Buchwald, S. L.; Jensen, K. F., An Integrated Microreactor System for Self-Optimization of a Heck Reaction: From Micro- to Mesoscale Flow Systems. *Agnew. Chem. Int. Ed.* **2010**, *49*, 7076-7080.

19. Krishnadasan, S.; Brown, R. J. C.; deMello, A. J.; deMello, J. C., Intelligent routes to the controlled synthesis of nanoparticles. *Lab on a Chip* **2007**, *7*, 1434-1441.

20. Bourne, R. A.; Skilton, R. A.; Parrott, A. J.; Irvine, D. J.; Poliakoff, M., Adaptive Process Optimization for Continuous Methylation of Alcohols in Supercritical Carbon Dioxide. *Organic Process Research & Development* **2011**, *15*, 932-938.

21. Frey, D.; Shin, J. H.; Musco, C.; Modestino, M. A., Chemically-informed data-driven optimization (ChIDDO): leveraging physical models and Bayesian learning to accelerate chemical research. *Reaction Chemistry & Engineering* **2022**.

22. Pomberger, A.; Pedrina McCarthy, A. A.; Khan, A.; Sung, S.; Taylor, C. J.; Gaunt, M. J.; Colwell, L.; Walz, D.; Lapkin, A. A.,

The effect of chemical representation on active machine learning towards closed-loop optimization. *Reaction Chemistry & Engineering* **2022**.

23. Konan, K. E.; Abollé, A.; Barré, E.; Aka, E. C.; Coeffard, V.; Felpin, F.-X., Developing flow photo-thiol–ene functionalizations of cinchona alkaloids with an autonomous self-optimizing flow reactor. *Reaction Chemistry & Engineering* **2022**.

24. Nandiwale, K. Y.; Hart, T.; Zahrt, A. F.; Nambiar, A. M. K.; Mahesh, P. T.; Mo, Y.; Nieves-Remacha, M. J.; Johnson, M. D.; García-Losada, P.; Mateos, C.; Rincón, J. A.; Jensen, K. F., Continuous stirred-tank reactor cascade platform for self-optimization of reactions involving solids. *Reaction Chemistry & Engineering* **2022,** *7*, 1315-1327.

25. Reizman, B. J.; Jensen, K. F., Simultaneous solvent screening and reaction optimization in microliter slugs. *Chemical Communications* **2015,** *51*, 13290-13293.

26. Hsieh, H.-W.; Coley, C. W.; Baumgartner, L. M.; Jensen, K. F.; Robinson, R. I., Photoredox Iridium–Nickel Dual-Catalyzed Decarboxylative Arylation Cross-Coupling: From Batch to Continuous Flow via Self-Optimizing Segmented Flow Reactor. *Organic Process Research & Development* **2018,** *22*, 542-550.

27. 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. *Reaction Chemistry & Engineering* **2018,** *3*, 301-311.

28. 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.

29. Clayton, A. D.; Schweidtmann, A. M.; Clemens, G.; Manson, J. A.; Taylor, C. J.; Niño, C. G.; Chamberlain, T. W.; Kapur, N.; Blacker, A. J.; Lapkin, A. A.; Bourne, R. A., Automated self-optimisation of multi-step reaction and separation processes using machine learning. *Chemical Engineering Journal* **2020,** *384*, 123340.

30. 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? *Reaction Chemistry & Engineering* **2022**.

31. 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. *Reaction Chemistry & Engineering* **2020,** *5*, 112-123.

32. McMullen, J. P.; Jensen, K. F., Rapid Determination of Reaction Kinetics with an Automated Microfluidic System. *Organic Process Research & Development* **2011,** *15*, 398-407.

33. Moore, J. S.; Jensen, K. F., "Batch" Kinetics in Flow: Online IR Analysis and Continuous Control. *Angewandte Chemie International Edition* **2014,** *53*, 470-473.

34. Mozharov, S.; Nordon, A.; Littlejohn, D.; Wiles, C.; Watts, P.; Dallin, P.; Girkin, J. M., Improved Method for Kinetic Studies in Microreactors Using Flow Manipulation and Noninvasive Raman Spectrometry. *Journal of the American Chemical Society* **2011,** *133*, 3601-3608.

35. Schwolow, S.; Heikenwalder, B.; Abahmane, L.; Kockmann, N.; Roder, T., Kinetic and Scale-up Investigations of a Michael Addition in Microreactors. *Organic Process Research & Development* **2014,** *18*, 1535-1544.

36. Hone, C. A.; Holmes, N.; Akien, G. R.; Bourne, R. A.; Muller, F. L., Rapid multistep kinetic model generation from transient flow data. *Reaction Chemistry & Engineering* **2017,** *2*, 103-108.

37. Taylor, C. J.; Manson, J. A.; Clemens, G.; Taylor, B. A.; Chamberlain, T. W.; Bourne, R. A., Modern advancements in continuous-flow aided kinetic analysis. *Reaction Chemistry & Engineering* **2022**.

38. Aroh, K. C.; Jensen, K. F., Efficient kinetic experiments in continuous flow microreactors. *Reaction Chemistry & Engineering* **2018,** *3*, 94-101.

39. 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. *Reaction Chemistry & Engineering* **2021,** *6*, 2306-2314.

40. Wyvratt, B. M.; McMullen, J. P.; Grosser, S. T., Multidimensional dynamic experiments for data-rich process development of reactions in flow. *Reaction Chemistry & Engineering* **2019,** *4*, 1637-1645.

41. O'Brien, A. G.; Horváth, Z.; Lévesque, F.; Lee, J. W.; Seidel-Morgenstern, A.; Seeberger, P. H., Continuous Synthesis and Purification by Direct Coupling of a Flow Reactor with Simulated Moving-Bed Chromatography. *Angewandte Chemie International Edition* **2012,** *51*, 7028-7030.

42. LCAP is an acronym for liquid chromatography area percent. It is measurement commonly used throughout drug substance research and development.

43. Beers, K. J., *Numerical Methods for Chemical Engineering Applications in Matlab*. Cambridge University Press: New York, 2007.

44. Carey, F. A., *Organic Chemistry Fourth Edition*. McGraw Hill: New York, 2000.

45. Hall, H. K., Correlation of the Base Strengths of Amines1. *Journal of the American Chemical Society* **1957,** *79*, 5441-5444.