An all-in-one multipurpose robotic platform for the self-optimization, intensification and scale-up of photocatalysis in flow

Aidan Slattery,[†,1] Zhenghui Wen,[†,1] Pauline Tenblad,[†,1] Diego Pintossi,[1]

Jesús Sanjosé-Orduna,[1] Tim den Hartog,[1,2,3] Timothy Noël[1,]*

[1] Flow Chemistry Group, van 't Hoff Institute for Molecular Sciences (HIMS), University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands.

[2] Zuyd University of Applied Sciences, Nieuw Eyckholt 300, 6419 DJ Heerlen, The Netherlands.

[3] The Netherlands Organisation for Applied Scientific Research (TNO), High Tech Campus 25, 5656AE Eindhoven, The Netherlands.

[†]*These authors contributed equally.*

** Email: t.noel@uva.nl*

Abstract

The optimization, intensification, and scaling up of chemical processes are essential and time-consuming aspects of contemporary chemical manufacturing, necessitating expertise and precision due to their intricate and sensitive nature. However, these process development problems are often carried out independently and consecutively, which can exacerbate the already significant consumption of time and resources involved in the process. In this work, we present a versatile, all-in-one robotic platform for the autonomous optimization, intensification, and scaling up of photocatalytic reactions in flow. This platform overcomes associated challenges through the integration of readily available hardware and custom software, offering a hands-off solution. Our open source platform combines a liquid-handler, syringe pumps, a tunable high-powered photoreactor, cheap IoT devices and an in-line NMR to enable automated, data-rich optimization using a Closed-Loop Bayesian Optimization strategy. The use of a high-power continuous-flow capillary photoreactor enables highly reproducible data to be obtained, as it mitigates issues related to mass, heat, and photon transport that are often the main sources of irreproducibility in photocatalytic transformations. A user-friendly graphical interface allows chemists without programming or machine learning expertise to easily optimize, monitor, and analyze photocatalytic reactions for chemical spaces of both continuous and discrete variables. The system's effectiveness was demonstrated by testing it on challenging photocatalytic transformations, which resulted in increased overall reaction yields and an impressive up to 550-fold improvement in space-time yields compared to batch processes. Additional tests on literature-reported reactions previously optimized in flow yielded substantial increases in both yield and space-time yield. Overall, our studies demonstrate that combining flow-based reactor technology with Bayesian optimization yields superior and unbiased results compared to human effort and intuition in terms of pace, precision, and outcomes for the optimization of photocatalytic reactions. Finally, due to its ability to autonomously generate datasets that include both optimal and suboptimal conditions, our RoboChem platform also contributes to advancing the field towards a digitally-driven era in synthetic chemistry.

Introduction

Over the past century, organic chemists have made significant advancements in developing a multitude of synthetic methods, empowering researchers to efficiently prepare increasingly intricate organic molecules.[1] The progress achieved in this field has broadened our access to a vast chemical space and facilitated the creation of molecules endowed with distinctive properties, which are crucial for the development of future pharmaceuticals, agrochemicals, and materials.[2] However, streamlining and optimizing a compound trace into a widely applicable synthetic method is both demanding and time-intensive, requiring meticulous attention to detail and significant investment of time and resources.

A typical strategy for the development of synthetic methodologies consists of several stages (Figure 1A). Initially, chemists employ random screening based on existing literature to identify potential reaction hits, which serve as starting points. Once an initial hit is discovered, the focus shifts towards optimizing the reaction conditions for a single substrate using a method known as "one-factor-at-a-time" (OFAT) .[3] In this approach, various reaction variables, such as ligands, bases, solvents, and temperature, are systematically screened one by one. The best result is retained, and the next variable is subsequently optimized. However, it is crucial to recognize that OFAT optimization fails to capture potential synergistic interactions between multiple variables,[4,5] which can lead to suboptimal outcomes. Furthermore, it is crucial to acknowledge that following this approach can result in over-optimization for a specific substrate, necessitating subsequent re-optimization of the reaction conditions to discover more general conditions that can be applied across the entire reaction scope. As a result, when these more "general" reaction conditions are uniformly applied to all members of the reaction scope, it may lead to less than optimal reaction yields for a significant portion of the scope.

Laboratory automation is increasingly being utilized by both academia and industry as a driver to increase expediency of chemical reaction discovery and optimization.[6] This movement has been a boon to the area of high-throughput experimentation (HTE) over the last few years,[7] and has fundamentally changed the way that chemists approach optimization and discovery.[8–12] However, there are drawbacks to the systems currently being utilized. Traditional, batch-based approaches have excelled at screening discrete variables (e.g. solvent, base, catalyst) however they are less effective at screening continuous variables (e.g. temperature, reaction time, photon equivalents[13] concentration) (Figure 1B).[14] On the flip-side automated continuous-flow reactor platforms, traditionally, are effective at changing continuous variables, but struggle to screen discrete parameters (Figure 1C).[15] While much work has been done in the development of automated flow reactors for reaction screening and optimization they nevertheless tend to be material inefficient and limited in the screening of discrete chemical parameters, or need to be transferred to another reactor system in order to perform scale up and intensification operations adding redundancies to the process.[16–20]

In response to the challenge of rapidly optimizing reaction conditions, we sought to develop an all-in-one multipurpose robotic platform, called RoboChem, that enables the self-optimization, intensification and scale up of photocatalytic transformations (Figure 1D). This innovative platform overcomes associated challenges by integrating off-the-shelf hardware and customized software, providing a hands-off solution. Our open-source platform combines a liquid handler, syringe pumps, a high-powered photoreactor, inexpensive Internet of Things (IoT) devices, and an in-line nuclear magnetic resonance (NMR) system to enable automated and data-rich optimization. By utilizing a high-power continuous-flow capillary photoreactor, our platform ensures highly reproducible data, effectively mitigating issues related to mass, heat, and photon transport that often contribute to

irreproducibility in photocatalytic transformations.[21,22] To account for complex intercorrelations between reaction variables, optimization algorithms such as design of experiments (DoE) and statistical modeling can be integrated into the platform. However, for complex non-linear relationships, such as those encountered in photocatalytic reactions, machine learning proves to be an even more effective approach.[23] Its ability to rapidly and efficiently analyze vast amounts of data enables the identification of underlying patterns and the extraction of meaningful conclusions.[24] Thus, combining machine learning with reaction automation is advantageous.[25] Given that our platform operates as a linear system (i.e., not parallelized), minimizing the number of experiments required to reach optimal conditions was crucial. For this reason, we turned to Bayesian Optimization, which has gained popularity in the chemistry community due to its ability to optimize black-box functions.[26,17,27] As an automated flow chemistry setup, our platform is capable of exploring large regions of the experimental and chemical space within a relatively short period, making it well-suited for addressing complex optimization problems encountered in photocatalysis. The Robochem platform distinguishes itself from current reporting methods by optimizing every substrate, thereby enabling a clear evaluation of the applicability and limitations of the reported transformations, resulting in increased value for industrial implementation. We demonstrate the general applicability of RoboChem to the optimization of a diverse set of photocatalytic transformations, including Hydrogen Atom Transfer (HAT) photocatalysis and photoredox catalysis, which are relevant to medicinal and crop protection chemistry.

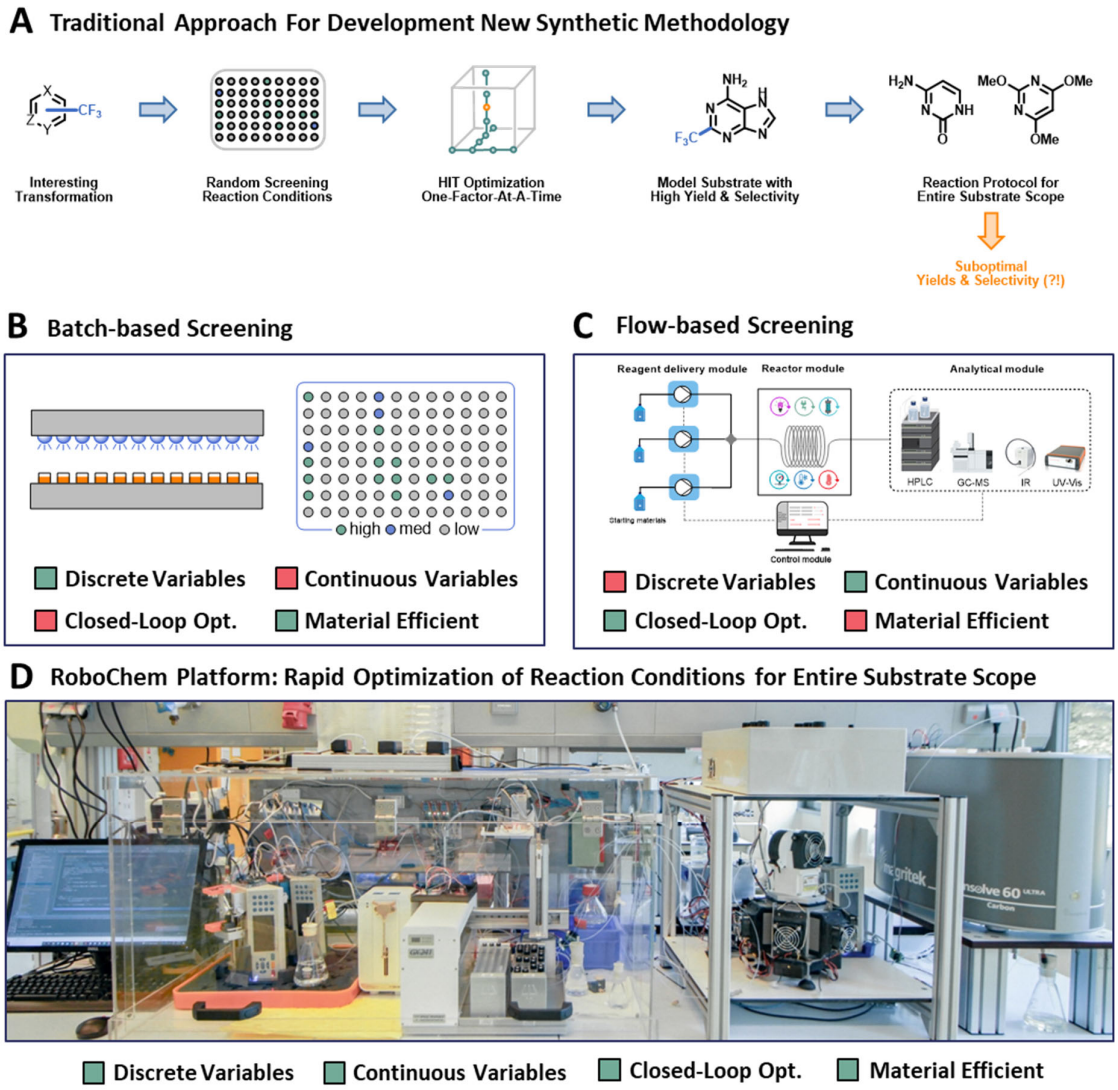

Figure 1 (A) Classical approach to develop a new synthetic method with subsequent scope elaboration. (B) Batch High-throughput experimentation platform for the rapid optimization of continuous flow systems using batch technology[28] (C) General representation of some popular flow chemistry optimization platforms which use pumps and algorithmic control (D) RoboChem – a benchtop platform for the optimization of photochemical systems which accommodates the ability to optimize both discrete and continuous variables, perform in a closed-loop optimization workflow while remaining material efficient.

RoboChem platform

The RoboChem platform can be divided into three distinct workflows: the controller, the planner and the receptionist (Figure 2A). The hardware controller oversees the management of the physical platform, encompassing tasks such as preparing the reaction mixture, executing the experiment, and conducting subsequent in-line analysis. The planner, which is a machine learning model, is responsible for determining the optimal experiments to run. It selects the experiment parameters and communicates them to the controller. The results are then fed back to the machine learning model,

which subsequently decides on the next experiment. Lastly, the receptionist functions as the Graphical User Interface (GUI), providing users with the interface to input the necessary parameters, launch the optimization campaign, and initiate the process.

Platform – Controller

RoboChem is controlled by custom Python code and uses open-source libraries (Figure 2A) with off-the-shelf instruments and devices. By coupling a liquid handler, syringe pumps, switching valves, a high power photoreactor as well as simple Internet of Things (IoT) devices such as phase sensors and ultrasonic detectors with an in-line 60 MHz NMR for data-rich optimization, we have come up with a workflow to easily and efficiently optimize and intensify photochemical processes (Figure 2B). Each generated 'reaction slug' (650 µL) represents a discrete set of reaction conditions, and the reactions are executed sequentially: sample preparation, followed by reaction under the specified conditions, and finally, automated analysis and processing. The use of NMR for data analysis allowed for the accurate gathering of yields without the need to first calibrate the analytics with a pure product, and allowed us to run most reactions with no internal standard. While a 60 MHz benchtop NMR was selected as the analytical technique, the platform is easily adaptable to accommodate other analytical techniques such as IR spectroscopy, HPLC-UV-Vis, HPLC-MS, or GC-MS. As the volume of the reaction slug is determined by the analytical method, it could be significantly reduced by an order of magnitude if an alternative technique like HPLC were utilized.[29]

The RoboChem platform operates in a closed-loop manner, driven by the Bayesian Optimization (BO) algorithm. This iterative process involves the BO algorithm proposing experiments, which are automatically executed and analyzed. The obtained results are then fed back into the BO algorithm, which generates a new set of conditions for further optimization.[30,31] By harnessing the capabilities of a photo-microreactor equipped with high-intensity LEDs, known for scaling up photochemistry to productivities exceeding 2 kg/day,[32] and its ability to make instantaneous computer-controlled power output adjustments, we can minimize reaction times and significantly enhance the throughput of the platform. This increased efficiency reduces the time required for a comprehensive optimization run.

By employing a series of phase sensors and a dedicated algorithm to detect the passage of a reaction slug, the reactor can efficiently track the position of the reaction slug within the platform (Figure 3C). This cost-effective approach enables precise control over the reaction as it traverses the system. The ability to accurately monitor the reaction's location throughout eliminates the need to hardcode pump volumes, allowing for seamless compatibility with reactors of varying sizes without requiring any code modifications. Moreover, this tracking capability facilitates precise "parking" of the reaction slug in the NMR for analysis, optimizing reagent usage by minimizing the required quantity for each reaction condition.

The entire system is conveniently located on a standard laboratory benchtop and is enclosed within a custom-designed, closed suction box, eliminating the need for placement within a fume hood during reaction runs. The system's design facilitates three distinct operating modes:
(i) Single experiment: Conducting a reaction under specific conditions, whether for the purpose of yield/productivity discovery or as part of a scope entry.
(ii) Self-optimization: Automating the optimization process for a single reaction or multiple reactions consecutively.
(iii) Scale-up: Exploiting the optimized conditions obtained through self-optimization for efficient scaling up of the reaction.

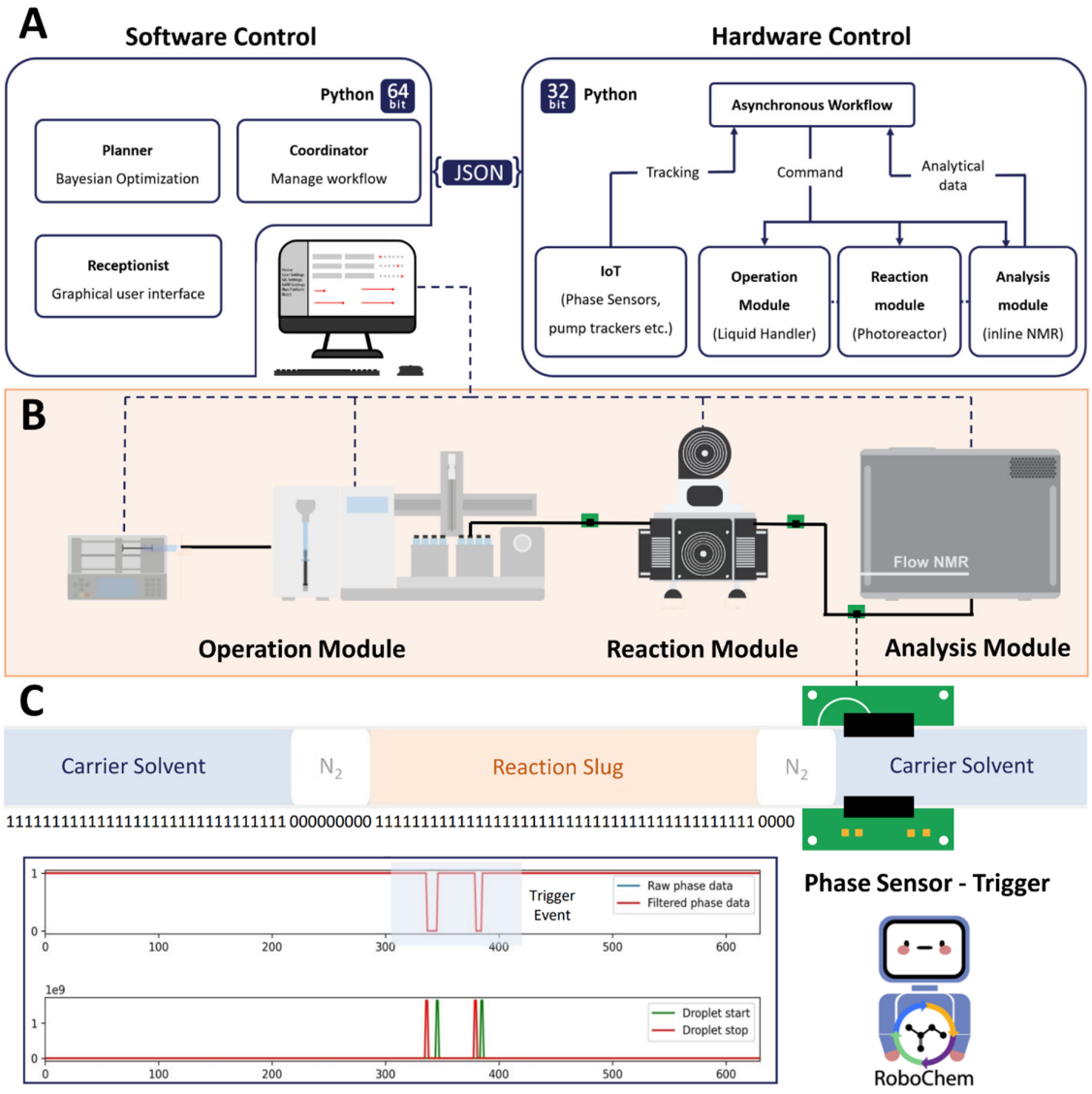

Figure 2. All-in-one multipurpose robotic platform for the self-optimization, intensification and scale-up of photochemistry in flow. (A) High-level view of platform architecture (B) Simplified-view of physical platform layout combining an operation module, reaction module and analysis module (C) Reaction tracking on the platform carried out by phase sensors to allow for triggering of events. Control by tracking a reaction slug as it passes over phase sensor and investigated by an algorithm to form a trigger for a next phase of the optimization cycle.

Bayesian Optimization – Planner

The platform leverages Bayesian Optimization (BO), a machine learning-based approach, to optimize chemical reactions. BO is a probabilistic model-based method designed to efficiently identify the maximum (or minimum) of an unknown black-box function.[33] It constructs a probability model of the function using carefully selected samples, which guides the search process by suggesting the next point to evaluate. The BO model incorporates both exploitation and exploration strategies. Exploitation involves investigating areas predicted to have the highest value, while exploration focuses

on exploring points where the model has limited knowledge. This dual approach prevents the model from becoming trapped in local maxima or minima. The iterative process continues with the model being updated after each new evaluation of the function until a predetermined threshold is reached or a specific number of experiments have been conducted.

The BO model is implemented using the open-source Python package Dragonfly, developed by Kandasamy et al.[31,34,35] The initial runs are decided using Latin-hypercube sampling.[36,37] The researchers define the input variables (parameters to be changed) and the objective to be optimized. The platform supports both single-objective and multi-objective optimization, targeting yield and/or throughput. In single-objective optimization, the model identifies the global maximum of the reaction. In multi-objective optimization, the model finds a set of non-dominated solutions known as the Pareto front.[38] To assess the progress of the optimization problem, the platform tracks the hypervolume after each run.[39] In cases of interrupted runs or the desire to build upon previously executed experiments, the platform allows for further optimization from that point (For further details, see Supplementary Information).

The platform's integration of machine learning effectively reduces the reliance on human resources.[40,41] Once the experiments are set up and the optimization process is initiated, the platform operates independently. The machine learning model autonomously determines the next set of experiments to run, and the corresponding commands are automatically transmitted to the platform. As a result, the platform can run continuously, including overnight, freeing up the chemist to focus on other tasks.

GUI – Receptionist

A key aspect of the platform's design is the development of an intuitive Graphical User Interface (GUI) that enables chemists without programming or machine learning expertise to easily navigate the system. The GUI provides functionality for creating new experiments, which store all the settings and results for an optimization run. It also allows users to generate the required positional and sample data utilized by the platform and liquid handler to prepare reactions (Figure 3A).

The liquid handler can accommodate multiple stock solutions of the same type. As the platform consumes stock solutions, it automatically tracks the quantity utilized. When one stock solution is depleted, the platform seamlessly transitions to the next vial containing a remaining stock solution. The GUI defines the entire chemical space to be explored under the Machine Learning Settings page.

In the Run Platform tab (Figure 3B), a button initiates the platform, and the GUI continuously tracks the results. For single-objective optimizations, the GUI presents a chart displaying the objective function (yield or throughput) against the number of runs. In multi-objective optimizations, it provides a plot of yield versus throughput and includes a graph tracking the hypervolume.

The GUI performs validations to ensure all necessary files and chemical spaces are properly defined before allowing the platform to run. If multiple reagents are added to a subcategory, the GUI automatically treats them as discrete variables for the optimization process.

After each run, the data consisting of both the input parameters and the output values, are automatically stored in a JSON file. These JSONs have been converted to .CSV files for easier data manipulation and are available in the supplementary information as both .CSV and in table format. They have been stored like this in guidance with the FAIR guiding principles for scientific data.[42] These datasets have the potential to be used in future projects; the data is of high quality due to the absence

of mass, heat and photon transport issues. Since they all are run on the same platform, the error between them is reduced. Another advantage is the presence of negative data which is commonly not published and thus hard to get, but nevertheless important for the development of machine learning models.[43]

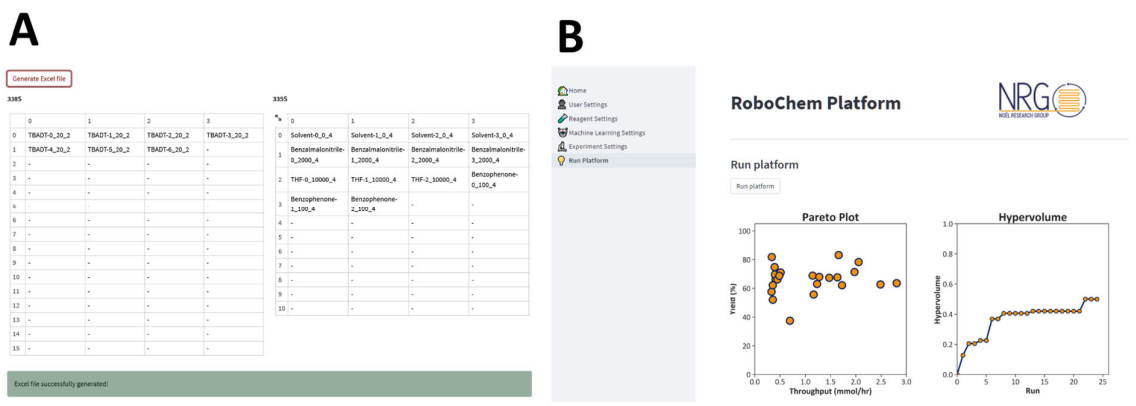

Figure 3. Graphical User Interface (A) Automated generation of sample vial positions for tracking of stock solutions (B) Plot generation for a multi-objective optimization tracking the pareto front and dominated points and the associated Hypervolume over time.

Performance Benchmarking

The primary focus of the RoboChem platform is to identify optimal reaction conditions for photocatalytic transformations. Our versatile platform caters to both single- and multi-objective optimization problems, offering synthetic chemists the ability to maximize yield, productivity, and other relevant objective functions. To accomplish this, we selected four distinct photocatalytic reactions, covering a total of 16 substrates, for optimization. For each case, we compared the yield and productivity reported in the literature with the conditions determined by the AI-assisted RoboChem platform. The reaction conditions discovered by the AI were subsequently employed to scale up the transformations.

Case Study 1: Single Objective Optimization for the photocatalytic HAT alkylation

We began our testing and validation of the platform with a Giese-type reaction via photocatalytic HAT activation of hydrocarbons.[44] The reaction was conducted in the Signify Eagle flow photo-microreactor, utilizing tunable 0-144 W, 365 nm light LEDs. This choice allowed us to evaluate a robust and well-established chemistry in our laboratory (Figure 4). Four optimization variables were selected for the reaction (Benzalmalononitrile concentration, THF loading, catalyst (TBADT) loading, and residence time). A total of nineteen experiments were conducted in a closed-loop fashion continuously for four hours. The initial phase involved six experiments, serving as a preliminary scan of the reaction space. Subsequently, the BO algorithm recommended one new experimental condition at a time, aiming to maximize the objective function (yield (%)). Within nine experiments, the platform achieved a yield of over 90% and began converging on the optimal conditions for the chemistry, resulting in a yield exceeding 95% for the desired product. Notably, the reaction demonstrated a detrimental effect of high light intensity, with the optimal range found to be between 20-50% of full power (28–72 W optical input power). These optimal conditions were then utilized for scaling up the transformation, confirming the AI-determined yield with an isolated yield of 99% (3.7 mmol scale).

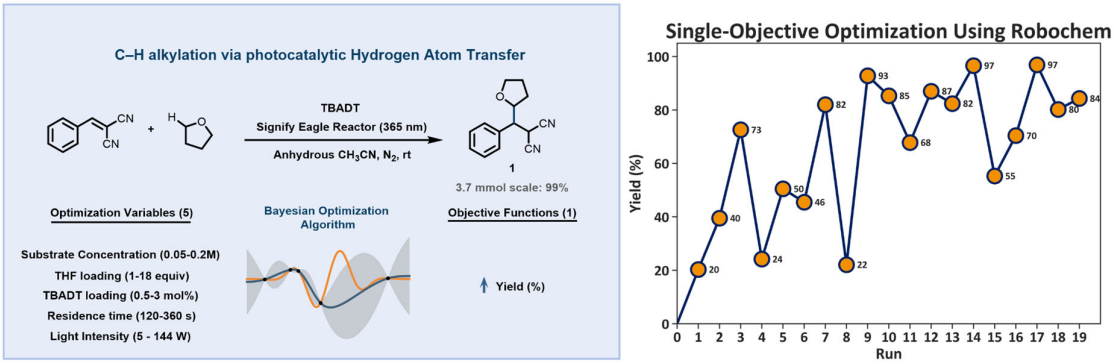

Figure 4. Single-objective optimization of the photocatalytic HAT-alkylation of benzalmalonitrile and THF.

Case Study 2: Single and Multi-Objective Optimization of C–H Trifluoromethylthiolation of C(sp^3)–H and C(sp^2)–H bonds via decatungstate-enabled hydrogen atom transfer (HAT)

Having validated the automated AI-driven photochemical platform in a single objective optimization problem, we aimed next to investigate its capability for optimizing various photocatalytic processes in multi-objective fashion, seeking to simultaneously optimize yield and throughput. Consequently, the reaction conditions that are found by the AI model are readily suitable for subsequent scale-up. As an initial benchmark, we selected the decatungstate-mediated trifluoromethylthiolation of C(sp^3)–H and C(sp^2)–H bonds *via* hydrogen atom transfer (HAT), as reported by Koenig et al.[45] The incorporation of the SCF_3 group in drug-like molecules holds significant value in medicinal chemistry. It offers high lipophilicity (as indicated by the Hansch parameter of π_R = 1.44) and notable electronegativity, which enhances the pharmacokinetic properties and optimizes the interaction between the active compound and its target.

In the trifluoromethylthiolation campaign (Figure 5), five reaction parameters and two objective functions were optimized simultaneously. The photochemistry was conducted in the Signify Eagle Reactor, which utilizes PFA tubing with a 0.8 mm I.D. and a total volume of 2.85 mL. To provide the necessary light source, a chip-on-board (COB) UV LED system with a tunable light intensity ranging from 0 to 144 W of optical power was employed. The screening chemical space encompassed five continuous parameters: Phth-SCF_3 concentration, H-donor equivalents, TBADT photocatalyst loading, residence time and light intensity. The objective functions chosen for optimization were either yield [%] or simultaneously the yield [%] and throughput [mmol h^{-1}] of the SCF_3-bearing molecules. To ensure fair comparisons between different reactor systems, we chose to convert the productivities into space-time yield (STY) ($g \cdot L^{-1} \cdot h^{-1}$). This normalization factorizes the reactor volume, allowing for a more equitable assessment of performance across varying reactor sizes.

For each substrate, a total of 18-36 experiments were conducted within an 8-16-hour timeframe. This comprised 8 initialization experiments followed by refinement experiments for each optimization campaign until a sufficient yield/hypervolume was achieved. Notably, substantial yield improvements were observed compared to their respective model counterparts in batch reactions. The platform demonstrated also a remarkable increase in productivities, ranging from 70 to 100 times higher. Next,

the reaction conditions selected by the AI model were successfully employed for scale-up to a 5 mmol scale. In all cases, the isolated yields obtained during the scale-up process closely matched the NMR yields observed with the AI-found reaction conditions.

Upon further analysis of the AI-discovered reaction conditions, several interesting observations arise. The AI algorithm refines the reaction conditions to achieve optimal reactivity and selectivity for each specific substrate. Notably, the Bayesian Optimization algorithm identifies experimental conditions that deviate significantly from the standard conditions reported by Koenig et al. One remarkable finding is the substantial differences in reaction/residence time and light intensity, which are parameters that experienced chemists often overlook initially. To illustrate this, we can compare the results obtained for trifluoromethylthiolated Sclareolide (**5**) and Ambroxide (**6**). It becomes evident that the catalyst loading and light intensity are notably lower for Ambroxide. This can be rationalized by the fact that Ambroxide can undergo an additional reaction with another equivalent of Phth-SCF_3, resulting in a double functionalized final product. However, such a reaction is not possible with Sclareolide, as the α-to-O carbon position is blocked by the carbonyl group. By reducing the catalyst loading and light intensity, the AI algorithm successfully enhances the yield and selectivity of the mono-functionalized product (**6**).

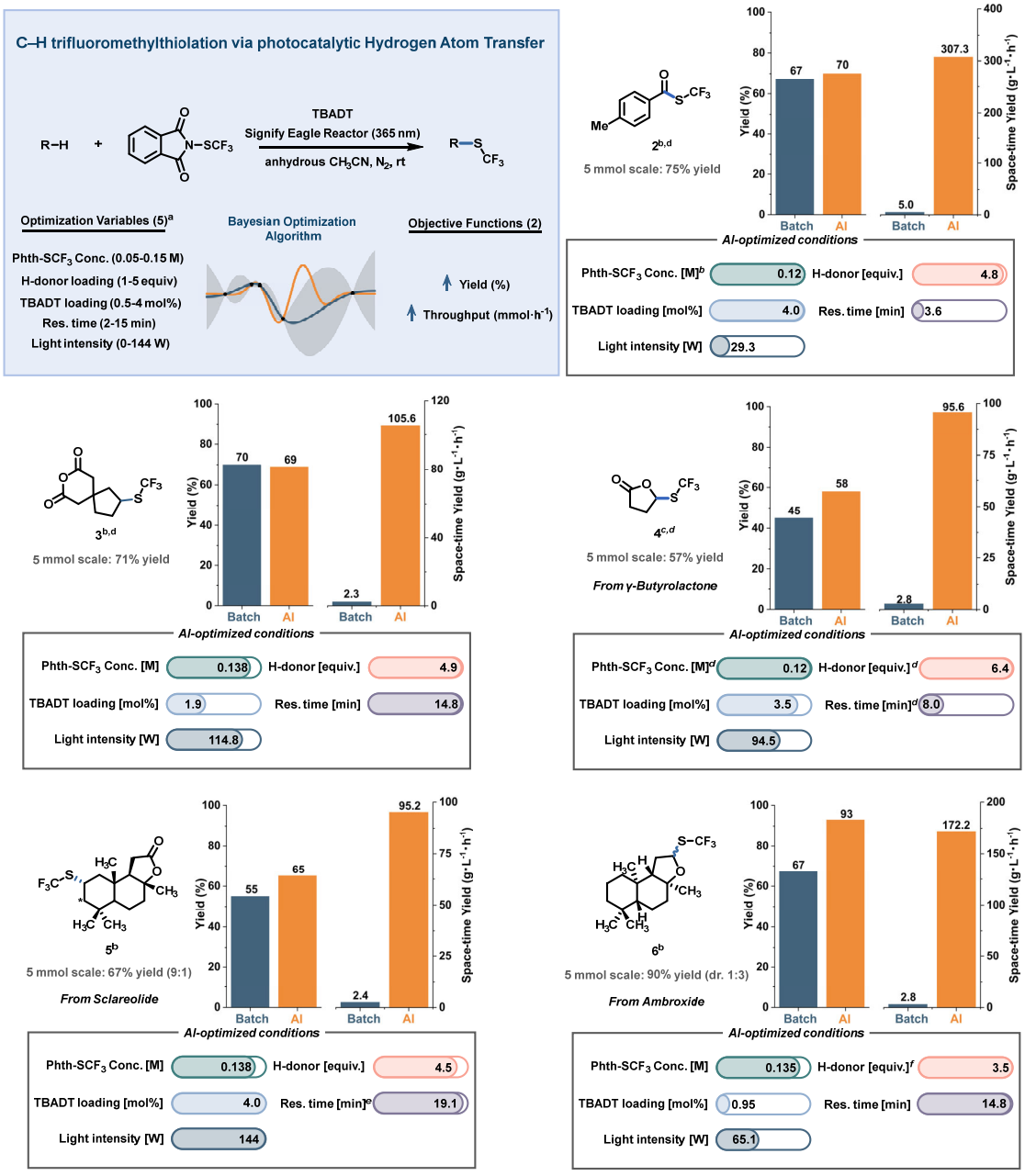

Figure 5. Substrate scope and associated summary for the C–H Trifluoromethylthiolation of enabled by Robochem. [a] Outer bounds of the chemical space explored for all experiments given, for more experimental details and exact chemical spaces explored for each experiment, see supplementary information part S5.3). [b] Batch conditions compared to those from literature: PhthSCF_3 Conc. (0.2 M), H-donor equiv. (1.5 equiv. for aldehydes, 2.5 equiv. for substrates with activated C(sp^3)–H bonds, 5 equiv. for substrates with unactivated C(sp^3)–H bonds), TBADT loading (2.5 mol%), resident time (16 h, 6 h for aldehydes), light (385 nm, 1.2 W), N_2, 25 °C. [c] Batch conditions compared to those carried out by us: Under literature conditions with home-made UFO reactor with 40 W 370 nm Kessil lamp, Yields determined by QNMR (See Supplementary Information Part S5.1). [d] Experiments run as single objective with the objective function to increase yield.

Case Study 3: Multi-Objective Optimization of Oxytrifluoromethylation of Alkenes using photocatalytic single electron transfer (SET)

Next, we directed our focus to the oxytrifluoromethylation of alkenes through a three-component process using photocatalytic single electron transfer (SET) with $Ru(bpy)_3(PF_6)_2$, as reported by Koike, Akita *et al*.[46] In the oxytrifluoromethylation campaign (Figure 6), we simultaneously optimized five reaction parameters: styrene concentration, CF_3 source loading, photocatalyst loading, residence time, and light intensity. Two objective functions (yield [%] and throughput [mmol h^{-1}]) were targeted for optimization. Similarly, for each substrate, a total of 14-25 experiments were conducted within a 3-10-hour timeframe. The optimization process utilized fast ^{19}F NMR analysis (2-minute per measurement) for molecules **7**, **8**, **9**, and **11**. However, molecule **10** required a longer optimization time of 19 hours due to the use of ^{1}H NMR for quantification. To ensure high accuracy, a 16-minute analysis window was allocated per reaction. As previously described, our experimental procedure involved 6 initialization experiments, followed by refinement experiments for each optimization campaign.

Similar to previous experiments, the photochemistry was conducted in the Signify Eagle Reactor, utilizing PFA tubing with a 0.8 mm I.D. However, for this campaign, chip-on-board blue LEDs with a tunable light intensity ranging from 0 to 188 W were employed to match the absorption maximum of the $Ru(bpy)_3$ photocatalyst. It is important to note that due to the short residence times, as low as 10 s, the internal volume of the photoreactor had to be reduced to 0.26 mL. This adjustment was necessary as the syringe pumps were unable to handle the high flow rates required with a larger internal volume.

The RoboChem platform successfully performed reaction optimization, resulting in conditions that produced outcomes closely aligned with the model batch reactions. Notably, a significant increase in space-time-yield, up to 565-fold, was achieved, demonstrating substantial potential for scale-up in the flow reactor. During the scale-up process, a slight improvement in yield was observed compared to the optimization carried out on the platform. This can be attributed to the fact that, for scale-up, the internal volume of the reactor was multiplied by a factor of 6, while the residence time remained the same. Consequently, an associated 6-fold increase in flow rate was necessary to maintain the desired residence time, leading to improved mass transfer facilitated by a higher Reynolds number. This phenomenon accounts for the observed increase in yield compared to the platform conditions.

The results indicate a significant dependence of this chemistry on the sustained power applied during the reaction, with higher wattage or longer residence time conditions resulting in noticeably lower yields. Remarkably, one reaction condition exhibited optimal performance at the lowest "turned on" power output, specifically molecule **9** at 2 W optical output. This outcome, which defies conventional expert intuition, further highlights the challenging nature of prediction in this context. Additionally, during the optimization of molecule **9**, various CF_3 sources were screened, including trifluoromethyl thianthrenium triflate[47] and Umemoto's reagent, serving as a test case to evaluate discrete variables. Notably, the algorithm determined that Umemoto's reagent was the optimal choice for this transformation.

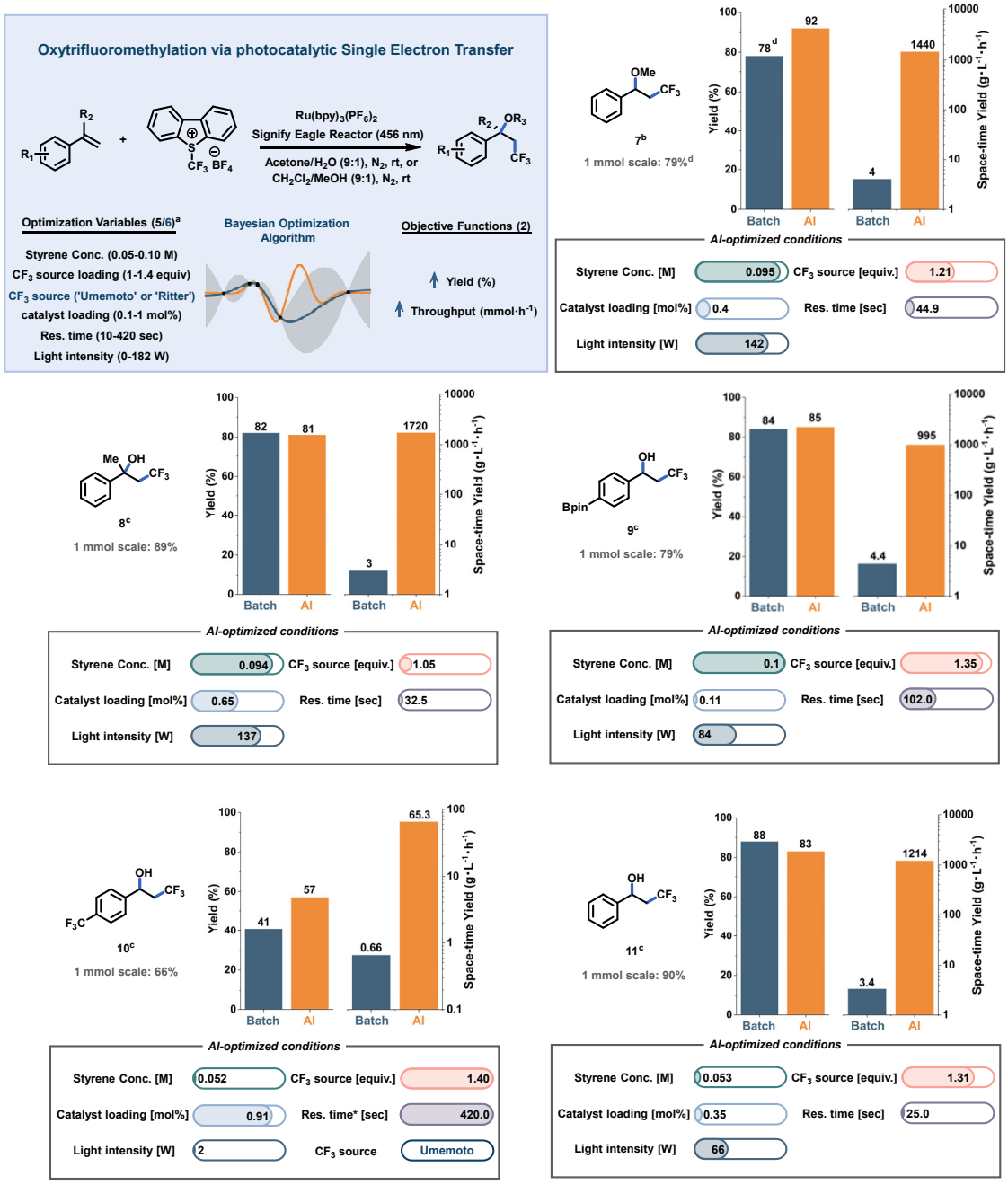

Figure 6. Substrate scope and associated summary for the three-component oxytrifluoromethylation of alkenes of enabled by Robochem. [a] Outer bounds of the chemical space explored for all experiments given, for more experimental details and exact chemical spaces explored for each experiment, see supplementary information part S6.2). [b] Batch conditions compared to those from literature: Styrene conc. (0.05M), *fac*-Ir(ppy)$_3$ loading (0.5 mol%), 'Umemoto' loading (1.05 equiv), $CHCl_2$:MeOH (9:1), 425 nm Blue LED (3 W), 2.5 h. [c] Batch conditions compared to those from literature: Styrene conc. (0.05M), *fac*-Ir(ppy)$_3$ loading (0.5 mol%), 'Umemoto' loading (1.1 equiv), Acetone:H_2O (9:1), 425 nm Blue LED (3 W), 2-4 h. [d] Yield >95% by QNMR.

Case Study 4: Multi-Objective Optimization of Aryl Trifluoromethylation

To provide a final example, our objective was to optimize the visible-light photocatalytic trifluoromethylation of highly functionalized heteroarenes developed by our group and researchers from Janssen pharmaceuticals (Figure 7).[48] In our original report, the reaction was carried out in a commercially-available Vapourtec UV-150 flow reactor. In the Signify Eagle Reactor equipped with blue LEDs, we scanned a search space consisting of 5 reaction parameters (Heteroarene concentration, CF_3SO_2Na loading, oxidant loading, residence time, and light intensity), targeting two objective functions (yield [%] and throughput [$mmolh^{-1}$]) for optimization. Notably, during the optimization of caffeine trifluoromethylation, we also incorporated a categorical variable to screen for the appropriate photocatalyst. This highlights the unique capability of the RoboChem platform to evaluate and optimize both discrete and continuous variables. For each substrate, a total of 17-35 experiments (including 6 initialization steps) were conducted within an 11-24-hour timeframe.

In this specific example, the RoboChem platform focused on optimizing a diverse range of densely-functionalized substrates that hold significant interest in drug discovery programs (Figure 7). Despite the original work being conducted in a flow system, we observed a substantial enhancement in both yield and productivity. This improvement can be attributed to the platform's ability to optimize each substrate individually, coupled with the utilization of a more potent light source in the Signify Eagle reactor.

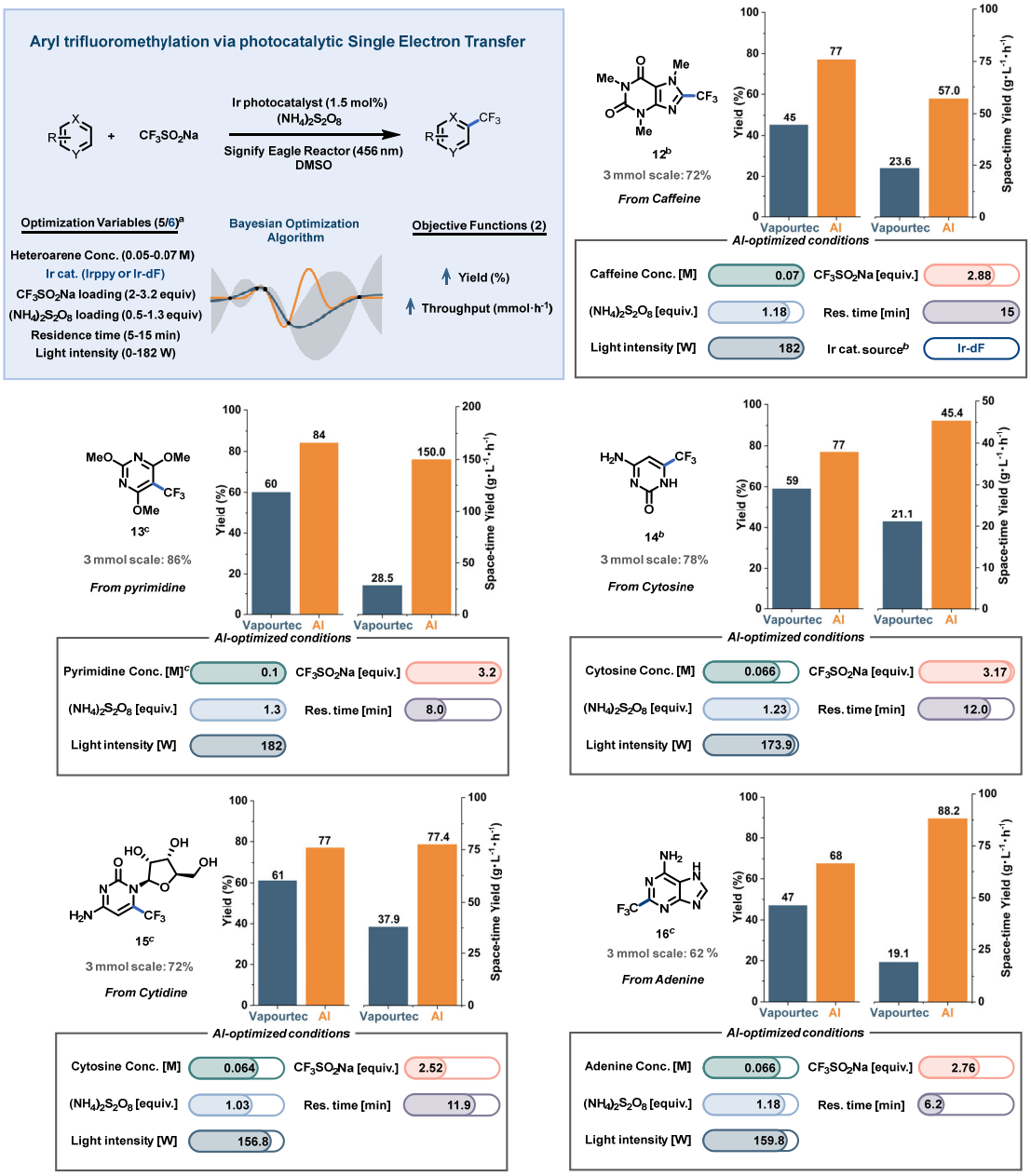

Figure 7. Substrate scope and associated summary for aryl trifluoromethylation via Single Electron Transfer enabled by Robochem. [a] Outer bounds of the chemical space explored for all experiments given, for more experimental details and exact chemical spaces explored for each experiment, see supplementary information part S7.3). [b] Comparisons made directly from literature under these conditions: Heteroarene Conc. (0.1 M), (Ir-dF = [Ir{dF(CF$_3$)ppy}$_2$](dtbpy)]PF$_6$) (1 mol%), CF_3SO_2Na equiv. (3 equiv.), $(NH_4)_2S_2O_8$ equiv. (1 equiv.), residence time (30 min), Vapourtec reactor (450 nm, 60 W). [c] Compared experiment was carried out by us: Under literature conditions with Vapourtec Photoreactor (456 nm, 60 W), Yields determined by QNMR (See Supplementary Information Part S7.2).

Conclusion and outlook

In conclusion, we have successfully developed a versatile and comprehensive robotic platform for the self-optimization, intensification, and scale-up of photochemistry in flow. Through the automated optimization and subsequent scale-up of 16 different photocatalytic reactions, our platform has showcased its capabilities.

We observed notable improvements in yield across all reactions, surpassing the performance under reported model reaction conditions. Additionally, we achieved remarkable increases in productivity, with a more than 500-fold enhancement compared to batch reactions and a 5-fold improvement over flow reactions.

RoboChem has proven to be a catalyst in expediting the optimization process in our laboratory. Its hands-off approach enhances safety while freeing our researchers to dedicate more time to the creative aspects of chemistry, rather than being burdened with the tedium of reaction optimization and intensification.

We are particularly excited about the modularity of the system and foresee its integration with different types of flow reactors and process analytical technologies in the future. Furthermore, we are eager to conduct further investigations into how different machine-learning settings and models impact the rate of reaction optimization.

Based on our observations throughout this study, we are confident that our RoboChem platform has immense potential for automating the evaluation of complete reaction scopes. This capability alone offers tremendous value in developing conditions for scalable compound libraries, such as expanding the size of building block catalogues.

Moreover, by individually optimizing reaction parameters and generating datasets that include both optimal and suboptimal conditions, we have uncovered intricate relationships between the targeted reaction parameters, the substrate structures, and the objective functions. The ability to automatically generate rich datasets, obtained within a highly reproducible reactor environment, paves the way for the future digitization of synthetic chemistry.

Acknowledgements

We gratefully acknowledge the generous funding provided by the Dutch Research Council (NWO) under the Open Technology Program (Multi-Modal Photochemistry, No 18433, A.S., D.P. & T.N). Z.W. was supported by the China Scholarship Council (CSC, No. 201808440313). Furthermore, we extend our gratitude for the generous funding received from the European Union H2020 research and innovation program under the European Research Council program (FlowHAT, No 101044355, T.N.), Marie S. Curie Grant Agreement (PhotoReAct, No 956324, P.T. & T.N.) and MSCA Individual Fellowship program (SELECTFLOW, No. 101061835, J.S.O.).

References

(1) Blakemore, D. C.; Castro, L.; Churcher, I.; Rees, D. C.; Thomas, A. W.; Wilson, D. M.; Wood, A. Organic Synthesis Provides Opportunities to Transform Drug Discovery. *Nat. Chem.* **2018**, *10* (4), 383–394. https://doi.org/10.1038/s41557-018-0021-z.

(2) Campos, K. R.; Coleman, P. J.; Alvarez, J. C.; Dreher, S. D.; Garbaccio, R. M.; Terrett, N. K.; Tillyer, R. D.; Truppo, M. D.; Parmee, E. R. The Importance of Synthetic Chemistry in the Pharmaceutical Industry. *Science (80-.).* **2019**, *363* (6424). https://doi.org/10.1126/science.aat0805.

(3) Lendrem, D. W.; Lendrem, B. C.; Woods, D.; Rowland-Jones, R.; Burke, M.; Chatfield, M.; Isaacs, J. D.; Owen, M. R. Lost in Space: Design of Experiments and Scientific Exploration in a Hogarth Universe. *Drug Discov. Today* **2015**, *20* (11), 1365–1371. https://doi.org/10.1016/j.drudis.2015.09.015.

(4) 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* (6), 3089–3126. https://doi.org/10.1021/acs.chemrev.2c00798.

(5) Torres, J. A. G.; Lau, S. H.; Anchuri, P.; Stevens, J. M.; Tabora, J. E.; Li, J.; Borovika, A.; Adams, R. P.; Doyle, A. G. A Multi-Objective Active Learning Platform and Web App for Reaction Optimization. *J. Am. Chem. Soc.* **2022**, *144* (43), 19999–20007. https://doi.org/10.1021/jacs.2c08592.

(6) Abolhasani, M.; Kumacheva, E. The Rise of Self-Driving Labs in Chemical and Materials Sciences. *Nat. Synth.* **2023**. https://doi.org/10.1038/s44160-022-00231-0.

(7) Mennen, S. M.; Alhambra, C.; Allen, C. L.; Barberis, M.; Berritt, S.; Brandt, T. A.; Campbell, A. D.; Castañón, J.; Cherney, A. H.; Christensen, M.; Damon, D. B.; Eugenio De Diego, J.; García-Cerrada, S.; García-Losada, P.; Haro, R.; Janey, J.; Leitch, D. C.; Li, L.; Liu, F.; Lobben, P. C.; Macmillan, D. W. C.; Magano, J.; McInturff, E.; Monfette, S.; Post, R. J.; Schultz, D.; Sitter, B. J.; Stevens, J. M.; Strambeanu, I. I.; Twilton, J.; Wang, K.; Zajac, M. A. The Evolution of High-Throughput Experimentation in Pharmaceutical Development and Perspectives on the Future. *Org. Process Res. Dev.* **2019**, *23* (6), 1213–1242. https://doi.org/10.1021/acs.oprd.9b00140.

(8) Krska, S. W.; DiRocco, D. A.; Dreher, S. D.; Shevlin, M. The Evolution of Chemical High-Throughput Experimentation to Address Challenging Problems in Pharmaceutical Synthesis. *Acc. Chem. Res.* **2017**, *50* (12), 2976–2985. https://doi.org/10.1021/acs.accounts.7b00428.

(9) Grainger, R.; Heightman, T. D.; Ley, S. V.; Lima, F.; Johnson, C. N. Enabling Synthesis in Fragment-Based Drug Discovery by Reactivity Mapping: Photoredox-Mediated Cross-Dehydrogenative Heteroarylation of Cyclic Amines. *Chem. Sci.* **2019**, *10* (8), 2264–2271. https://doi.org/10.1039/C8SC04789H.

(10) Buitrago Santanilla, A.; Regalado, E. L.; Pereira, T.; Shevlin, M.; Bateman, K.; Campeau, L.-C.; Schneeweis, J.; Berritt, S.; Shi, Z.-C.; Nantermet, P.; Liu, Y.; Helmy, R.; Welch, C. J.; Vachal, P.; Davies, I. W.; Cernak, T.; Dreher, S. D. Nanomole-Scale High-Throughput Chemistry for the Synthesis of Complex Molecules. *Science (80-.).* **2015**, *347* (6217),

49–53. https://doi.org/10.1126/science.1259203.

(11) Lin, S.; Dikler, S.; Blincoe, W. D.; Ferguson, R. D.; Sheridan, R. P.; Peng, Z.; Conway, D. V.; Zawatzky, K.; Wang, H.; Cernak, T.; Davies, I. W.; DiRocco, D. A.; Sheng, H.; Welch, C. J.; Dreher, S. D. Mapping the Dark Space of Chemical Reactions with Extended Nanomole Synthesis and MALDI-TOF MS. *Science (80-.).* **2018**, *361* (6402). https://doi.org/10.1126/science.aar6236.

(12) Halperin, S. D.; Kwon, D.; Holmes, M.; Regalado, E. L.; Campeau, L. C.; Dirocco, D. A.; Britton, R. Development of a Direct Photocatalytic C-H Fluorination for the Preparative Synthesis of Odanacatib. *Org. Lett.* **2015**, *17* (21), 5200–5203. https://doi.org/10.1021/acs.orglett.5b02532.

(13) Corcoran, E. B.; McMullen, J. P.; Lévesque, F.; Wismer, M. K.; Naber, J. R. Photon Equivalents as a Parameter for Scaling Photoredox Reactions in Flow: Translation of Photocatalytic C−N Cross-Coupling from Lab Scale to Multikilogram Scale. *Angew. Chemie Int. Ed.* **2020**, *59* (29), 11964–11968. https://doi.org/10.1002/anie.201915412.

(14) Santanilla, A. B.; Regalado, E. L.; Pereira, T.; Shevlin, M.; Bateman, K.; Campeau, L.; Schneeweis, J.; Berritt, S.; Shi, Z.; Nantermet, P.; Liu, Y.; Helmy, R.; Welch, C. J.; Vachal, P.; Davies, I. W.; Cernak, T.; Dreher, S. D. Nanomole-Scale High-Throughput Chemistry for the Synthesis of Complex Molecules. *Science (80-.).* **2015**, *347* (6217), 443–448.

(15) Plutschack, M. B.; Pieber, B.; Gilmore, K.; Seeberger, P. H. The Hitchhiker's Guide to Flow Chemistry. *Chem. Rev.* **2017**, *117* (18), 11796–11893. https://doi.org/10.1021/acs.chemrev.7b00183.

(16) Simon, K.; Sagmeister, P.; Munday, R.; Leslie, K.; Hone, C. A.; Kappe, C. O. Automated Flow and Real-Time Analytics Approach for Screening Functional Group Tolerance in Heterogeneous Catalytic Reactions. *Catal. Sci. Technol.* **2022**, *12* (6), 1799–1811. https://doi.org/10.1039/D2CY00059H.

(17) 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. Chemie Int. Ed.* **2023**, *62* (3). https://doi.org/10.1002/anie.202214511.

(18) 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. Chemie Int. Ed.* **2021**, *60* (15), 8139–8148. https://doi.org/10.1002/anie.202016007.

(19) 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 (80-.).* **2018**, *361* (6408), 1220–1225. https://doi.org/10.1126/science.aat0650.

(20) 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. *Org.*

Process Res. Dev. **2018**, *22* (4), 542–550. https://doi.org/10.1021/acs.oprd.8b00018.

(21) Buglioni, L.; Raymenants, F.; Slattery, A.; Zondag, S. D. A.; Noël, T. Technological Innovations in Photochemistry for Organic Synthesis: Flow Chemistry, High-Throughput Experimentation, Scale-up, and Photoelectrochemistry. *Chem. Rev.* **2022**, *122* (2), 2752–2906. https://doi.org/10.1021/acs.chemrev.1c00332.

(22) Cambié, D.; Bottecchia, C.; Straathof, N. J. W.; Hessel, V.; Noël, T. Applications of Continuous-Flow Photochemistry in Organic Synthesis, Material Science, and Water Treatment. *Chem. Rev.* **2016**, *116* (17), 10276–10341. https://doi.org/10.1021/acs.chemrev.5b00707.

(23) Freiesleben, J.; Keim, J.; Grutsch, M. Machine Learning and Design of Experiments: Alternative Approaches or Complementary Methodologies for Quality Improvement? *Qual. Reliab. Eng. Int.* **2020**, *36* (6), 1837–1848. https://doi.org/10.1002/qre.2579.

(24) Jordan, M. I.; Mitchell, T. M. Machine Learning: Trends, Perspectives, and Prospects. *Science (80-.).* **2015**, *349* (6245), 255–260. https://doi.org/10.1126/science.aaa8415.

(25) Houben, C.; Lapkin, A. A. Automatic Discovery and Optimization of Chemical Processes. *Curr. Opin. Chem. Eng.* **2015**, *9*, 1–7. https://doi.org/10.1016/j.coche.2015.07.001.

(26) Shields, B. J.; Stevens, J.; Li, J.; Parasram, M.; Damani, F.; Alvarado, J. I. M.; Janey, J. M.; Adams, R. P.; Doyle, A. G. Bayesian Reaction Optimization as a Tool for Chemical Synthesis. *Nature* **2021**, *590* (7844), 89–96. https://doi.org/10.1038/s41586-021-03213-y.

(27) Häse, F.; Roch, L. M.; Aspuru-Guzik, A. Chimera: Enabling Hierarchy Based Multi-Objective Optimization for Self-Driving Laboratories. *Chem. Sci.* **2018**, *9* (39), 7642–7655. https://doi.org/10.1039/C8SC02239A.

(28) González-Esguevillas, M.; Fernández, D. F.; Rincón, J. A.; Barberis, M.; de Frutos, O.; Mateos, C.; García-Cerrada, S.; Agejas, J.; MacMillan, D. W. C. Rapid Optimization of Photoredox Reactions for Continuous-Flow Systems Using Microscale Batch Technology. *ACS Cent. Sci.* **2021**, *7* (7), 1126–1134. https://doi.org/10.1021/acscentsci.1c00303.

(29) Avila, C.; Cassani, C.; Kogej, T.; Mazuela, J.; Sarda, S.; Clayton, A. D.; Kossenjans, M.; Green, C. P.; Bourne, R. A. Automated Stopped-Flow Library Synthesis for Rapid Optimisation and Machine Learning Directed Experimentation. *Chem. Sci.* **2022**, *13* (41), 12087–12099. https://doi.org/10.1039/D2SC03016K.

(30) 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* (5), 373–386. https://doi.org/10.1016/j.trechm.2021.02.005.

(31) Kandasamy, K.; Vysyaraju, K. R.; Neiswanger, W.; Paria, B.; Collins, C. R.; Schneider, J.; Poczos, B.; Xing, E. P. Tuning Hyperparameters without Grad Students: Scalable and Robust Bayesian Optimisation with Dragonfly. *J. Mach. Learn. Res.* **2020**, *21* (81), 1–27.

(32) Wan, T.; Wen, Z.; Laudadio, G.; Capaldo, L.; Lammers, R.; Rincón, J. A.; García-Losada,

P.; Mateos, C.; Frederick, M. O.; Broersma, R.; Noël, T. Accelerated and Scalable C(Sp 3)–H Amination via Decatungstate Photocatalysis Using a Flow Photoreactor Equipped with High-Intensity LEDs. *ACS Cent. Sci.* **2022**, *8* (1), 51–56. https://doi.org/10.1021/acscentsci.1c01109.

(33) Shahriari, B.; Swersky, K.; Wang, Z.; Adams, R. P.; de Freitas, N. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proc. IEEE* **2016**, *104* (1), 148–175. https://doi.org/10.1109/JPROC.2015.2494218.

(34) Paria, B.; Kandasamy, K.; Póczos, B. A Flexible Framework for Multi-Objective Bayesian Optimization Using Random Scalarizations. **2018**.

(35) 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* (6), 825–836. https://doi.org/10.1021/acscentsci.2c00207.

(36) Li, W.; Lu, L.; Xie, X.; Yang, M. A Novel Extension Algorithm for Optimized Latin Hypercube Sampling. *J. Stat. Comput. Simul.* **2017**, *87* (13), 2549–2559. https://doi.org/10.1080/00949655.2017.1340475.

(37) McKay, M. D.; Beckman, R. J.; Conover, W. J. Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics* **1979**, *21* (2), 239–245. https://doi.org/10.1080/00401706.1979.10489755.

(38) Clayton, A. D.; Manson, J. A.; Taylor, C. J.; Chamberlain, T. W.; Taylor, B. A.; Clemens, G.; Bourne, R. A. Algorithms for the Self-Optimisation of Chemical Reactions. *React. Chem. Eng.* **2019**, *4* (9), 1545–1554. https://doi.org/10.1039/C9RE00209J.

(39) Daulton, S.; Balandat, M.; Bakshy, E. Differentiable Expected Hypervolume Improvement for Parallel Multi-Objective Bayesian Optimization. **2020**. https://doi.org/10.48550/arXiv.2006.05078.

(40) Steiner, S.; Wolf, J.; Glatzel, S.; Andreou, A.; Granda, J. M.; Keenan, G.; Hinkley, T.; Aragon-Camarasa, G.; Kitson, P. J.; Angelone, D.; Cronin, L. Organic Synthesis in a Modular Robotic System Driven by a Chemical Programming Language. *Science (80-.).* **2019**, *363* (6423). https://doi.org/10.1126/science.aav2211.

(41) Ley, S. V.; Fitzpatrick, D. E.; Ingham, R. J.; Myers, R. M. Organic Synthesis: March of the Machines. *Angew. Chemie Int. Ed.* **2015**, *54* (11), 3449–3464. https://doi.org/10.1002/anie.201410744.

(42) Wilkinson, M. D.; Dumontier, M.; Aalbersberg, Ij. J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.-W.; da Silva Santos, L. B.; Bourne, P. E.; Bouwman, J.; Brookes, A. J.; Clark, T.; Crosas, M.; Dillo, I.; Dumon, O.; Edmunds, S.; Evelo, C. T.; Finkers, R.; Gonzalez-Beltran, A.; Gray, A. J. G.; Groth, P.; Goble, C.; Grethe, J. S.; Heringa, J.; 't Hoen, P. A. .; Hooft, R.; Kuhn, T.; Kok, R.; Kok, J.; Lusher, S. J.; Martone, M. E.; Mons, A.; Packer, A. L.; Persson, B.; Rocca-Serra, P.; Roos, M.; van Schaik, R.; Sansone, S.-A.; Schultes, E.; Sengstag, T.; Slater, T.; Strawn, G.; Swertz, M. A.; Thompson, M.; van der Lei, J.; van Mulligen, E.; Velterop, J.; Waagmeester, A.; Wittenburg, P.; Wolstencroft, K.; Zhao, J.; Mons, B. The FAIR Guiding Principles for

Scientific Data Management and Stewardship. *Sci. Data* **2016**, *3* (1), 160018. https://doi.org/10.1038/sdata.2016.18.

(43) Maloney, M. P.; Coley, C. W.; Genheden, S.; Carson, N.; Helquist, P.; Norrby, P.-O.; Wiest, O. Negative Data in Data Sets for Machine Learning Training. *J. Org. Chem.* **2023**, *88* (9), 5239–5241. https://doi.org/10.1021/acs.joc.3c00844.

(44) Wen, Z.; Maheshwari, A.; Sambiagio, C.; Deng, Y.; Laudadio, G.; Van Aken, K.; Sun, Y.; Gemoets, H. P. L.; Noël, T. Optimization of a Decatungstate-Catalyzed C(Sp 3)–H Alkylation Using a Continuous Oscillatory Millistructured Photoreactor . *Org. Process Res. Dev.* **2020**, *24* (10), 2356–2361. https://doi.org/10.1021/acs.oprd.0c00235.

(45) Schirmer, T. E.; Rolka, A. B.; Karl, T. A.; Holzhausen, F.; König, B. Photocatalytic C–H Trifluoromethylthiolation by the Decatungstate Anion. *Org. Lett.* **2021**, *23* (15), 5729–5733. https://doi.org/10.1021/acs.orglett.1c01870.

(46) Yasu, Y.; Koike, T.; Akita, M. Three-Component Oxytrifluoromethylation of Alkenes: Highly Efficient and Regioselective Difunctionalization of C□C Bonds Mediated by Photoredox Catalysts. *Angew. Chemie Int. Ed.* **2012**, *51* (38), 9567–9571. https://doi.org/10.1002/anie.201205071.

(47) Jia, H.; Häring, A. P.; Berger, F.; Zhang, L.; Ritter, T. Trifluoromethyl Thianthrenium Triflate: A Readily Available Trifluoromethylating Reagent with Formal CF 3 + , CF 3 • , and CF 3 – Reactivity. *J. Am. Chem. Soc.* **2021**, *143* (20), 7623–7628. https://doi.org/10.1021/jacs.1c02606.

(48) Abdiaj, I.; Bottecchia, C.; Alcazar, J.; Noël, T. Visible-Light-Induced Trifluoromethylation of Highly Functionalized Arenes and Heteroarenes in Continuous Flow. *Synthesis (Stuttg).* **2017**, *49* (22), 4978–4985. https://doi.org/10.1055/s-0036-1588527.