nature synthesis

Article

https://doi.org/10.1038/s44160-026-01053-0

A flexible and affordable self-driving laboratory for automated reaction optimization

Received: 7 August 2025

Accepted: 13 March 2026

Published online: 13 April 2026

Check for updates

Simone Pilon[1,8], Elia Savino[1,8], Oliver M. Bayley[1,2,8], Michael Vanzella[1], Miguel Claros[1], Petros Siasiaridis[1], Junsong Liu[1], Florian Lukas[1], Matteo Damian[3], Vasilis Tseliou[3], Niccolò Intini[1,4], Aidan Slattery[1], Jesus SanJosé-Orduna[1], Tim den Hartog[1,5,6], Ron A. H. Peters[2,7], Andrea F. G. Gargano[2], Francesco G. Mutti[3] & Timothy Noël[1]

Self-driving laboratories have the potential to revolutionize chemical discovery and optimization, yet their widespread adoption remains limited by high costs, complex infrastructure and limited accessibility. Here we introduce RoboChem-Flex, a low-cost, modular self-driving laboratory platform designed to democratize autonomous chemical experimentation. The system combines customizable, in-house-built hardware with a flexible Python-based software framework that integrates real-time device control and advanced Bayesian optimization strategies, including multi-objective and transfer learning workflows. RoboChem-Flex supports both fully autonomous closed-loop operation and human-in-the-loop configurations, enabling seamless integration with shared analytical equipment and minimizing entry barriers. We validate the versatility of the platform across six diverse case studies, including photocatalysis, biocatalysis, thermal cross-couplings and enantioselective catalysis, spanning both single- and multi-objective optimizations. Through these case studies, we demonstrate RoboChem-Flex's ability to navigate large, complex chemical spaces, autonomously identify scalable high-performance reaction conditions, and flexibly adapt to a variety of analytical set-ups. By providing an affordable, scalable and open platform, RoboChem-Flex offers a tangible step towards making self-driving laboratories accessible to resource-limited laboratories, fostering broader participation in automated chemical research.

The rapid rise of automation, artificial intelligence (AI) and self-driving laboratories (SDLs) has the potential to revolutionize synthetic organic chemistry. SDLs streamline essential tasks such as reagent handling, synthesis and data analysis, generating high-quality, reproducible results at unprecedented speeds[1,2]. By automating repetitive procedures, SDLs allow researchers to focus on higher-level scientific objectives, including ideation, data interpretation and hypothesis formulation. Moreover, SDLs hold transformative potential for tackling global challenges in energy, pharmaceuticals and materials science[3].

Despite their promise, the widespread adoption of SDLs faces substantial barriers. Primarily, designing, constructing and operating these platforms requires multidisciplinary expertise in hardware (for example, liquid handlers and robotic arms), software (for example, optimization algorithms), chemistry, analytical and data science. Furthermore, existing SDLs remain prohibitively expensive. High-end systems incorporating robotic arms, liquid handlers and pumps often exceed US$100,000, excluding analytical components. The first 'RoboChem' platform developed by our group cost approximately

A full list of affiliations appears at the end of the paper. e-mail: t.noel@uva.nl

US$50,000 for essential components alone, not including the most expensive component, a benchtop nuclear magnetic resonance (NMR) spectrometer[4]. More sophisticated set-ups can far exceed these figures, restricting SDL accessibility to a select group of well-funded institutions. As a result, these advancements risk amplifying the Matthew effect in science, where well-resourced research groups gain disproportionate advantages through a feedback loop of resource accumulation[5]. This further widens the disparities in research capabilities, innovation opportunities and scientific influence (Fig. 1a,b).

We took up the challenge to develop a low-cost SDL capable of performing experiments with minimal human supervision, making it accessible to researchers worldwide. Affordable automated systems do exist; however, these platforms sacrifice research potential by focusing on narrowly defined problems[6,7]. Moreover, synthetic organic chemistry lags behind other fields in digitalization, largely due to a lack of reliable, high-quality datasets that comprehensively cover diverse chemical substrates and reactions[8–10]. SDLs offer a solution by generating reproducible, well-documented datasets. However, digitalizing chemistry must not remain an exclusive privilege of well-funded laboratories. To democratize access, it is essential to develop scalable, cost-effective tools that empower researchers across all resource levels.

In this work, we present 'RoboChem-Flex', a low-cost yet powerful robotic platform designed to democratize automation in synthetic organic chemistry. Affordability and flexibility are achieved using three-dimensionally (3D)-printed or readily available subcomponents, which reduces costs considerably while allowing rapid customization and iterative development[11]. Communication between hardware components is orchestrated by our open-source OmniPlatypus package, which ensures seamless modularity and enables a plug-and-play architecture with minimal coding effort required from the user. At the software level, RoboChem-Flex integrates a highly modular Bayesian optimization agent, allowing users to customize AI-driven optimization workflows to meet specific experimental goals. The platform also supports integration with a range of inline analytical instruments, including NMR spectroscopy, ultrahigh-performance liquid chromatography (UHPLC)–mass spectrometry and Raman spectroscopy, enabling fully closed-loop reaction optimization. Recognizing, however, that inline analytics may represent a sizeable investment, we have also developed a cost-effective, 3D-printed liquid sampling unit. This module enables the robot to collect reaction samples, which can then be analysed using existing, often departmentally shared, analytical equipment. This human-in-the-loop approach provides a practical and affordable entry point for laboratories, reducing the overall system cost to approximately US$5,000. Thus, by equipping resource-limited research groups with tools on par with those in well-funded institutions, RoboChem-Flex aims to level the playing field and foster innovation at all scales (Fig. 1c).

Results and discussion

The hardware

The RoboChem-Flex self-driving platform uses a slug-flow regime, enabling the sequential investigation of reaction conditions with 5–50 µmol of material per run. Leveraging the inherent scalability of flow reactors, results obtained at the microscale can be reliably translated to the millimole scale[12]. The platform architecture is built around a minimal, highly modular set of hardware and software components that prioritize accessibility, versatility and full customizability. Each hardware module is self-assembled using standardized, low-cost and widely available materials, such as slotted aluminium profiles, linear ball bearings and trapezoidal lead screws. Arduino Uno microcontrollers manage low-level hardware control. Their widespread use in education makes them particularly accessible to multidisciplinary researchers without specialized engineering backgrounds[13]. All non-standard components, which are typically expensive or difficult to source, have been designed using open-source software and fabricated via 3D printing in polyethylene terephthalate glycol plastic. In addition, custom printed circuit boards have been implemented in several devices to improve electrical safety, minimize loose cabling and streamline assembly. These printed circuit boards can be readily manufactured at low cost using commercially available services based on the design files provided.

Two sampler modules are responsible for manipulating reagents and solvents within the system. They are built upon an affordable, commercially available three-axis Cartesian robot, which has been modified to enable liquid sampling, mixing and transferring operations. The samplers can be flexibly configured with vials of various sizes via 3D-printed vial holders, as well as multiple injection ports to direct slugs through different flow paths within the experimental set-up. Vials are sealed with readily available septum caps, and pressure compensation is achieved with a dedicated inert gas line. This limits sample degradation without requiring a costly glovebox setup. Two custom-built syringe pumps manage reagent sampling and slug movement through the system. These pumps offer fine-grained control and real-time feedback not typically available in commercial models. Integrated encoders and end-of-travel sensors prevent blockages and plunger misalignment, while electronic flow-switch valves are fully integrated within the device controller. Additional modules, designed using similar principles, allow automated control of light sources, gas valves, cooling and heating.

The software and machine learning framework

The RoboChem-Flex platform control software is developed entirely in Python, providing a robust, user-friendly environment that minimizes the specialized skillset required to fully operate and understand the system. Its operation is governed by a hierarchical, modular architecture. At the top level, a web-based graphical user interface offers an intuitive entry point for users without programming expertise. Through this interface, users can select experimental workflows, define reagents and conditions, and specify the chemical space to explore. These parameters are passed on to 'RoBrains', a custom Bayesian optimization engine based on BoTorch and developed for flexible, data-driven experimentation[14]. RoBrains enables autonomous platform operation while minimizing the number of required experiments by making informed, probabilistic decisions. The optimization engine supports a wide range of features, offering multiple surrogate models, transfer learning workflows, weighted multi-objective optimization and batching strategies which improve efficiency. In all campaigns presented in this study, RoBrains generated initial sets of reaction conditions using Latin Hypercube Sampling (LHS) across the defined parameter space. Then, as experimental results became available, RoBrains iteratively selected new conditions based on the specified acquisition function and optimization parameters. This iterative cycle continued until a satisfactory solution was identified, as indicated by hypervolume plateauing.

A second Python package, 'OmniPlatypus', manages real-time device control, orchestrating the execution of experiments and the acquisition of analytical data. It implements software objects for each physical device of the platform, which handle all low-level communication. Platforms are described within an accessible configuration file, listing devices, their physical interconnections and operational details. This design allows multiple hardware implementations to seamlessly coexist within the same codebase. Experimental workflows are composed of Unit Tasks, which describe mid-level atomic actions such as filling a syringe or transferring a reagent from a vial. This modular design enables rapid reconfiguration and high software reusability, as new experiments can be formulated by combining existing Unit Tasks.

In this work, we present six case studies that demonstrate the versatility of RoboChem-Flex across diverse chemical optimization challenges, including photocatalytic transformations, biocatalysis, thermal cross-couplings and enantioselective catalysis. An additional

https://doi.org/10.1038/s44160-026-01053-0

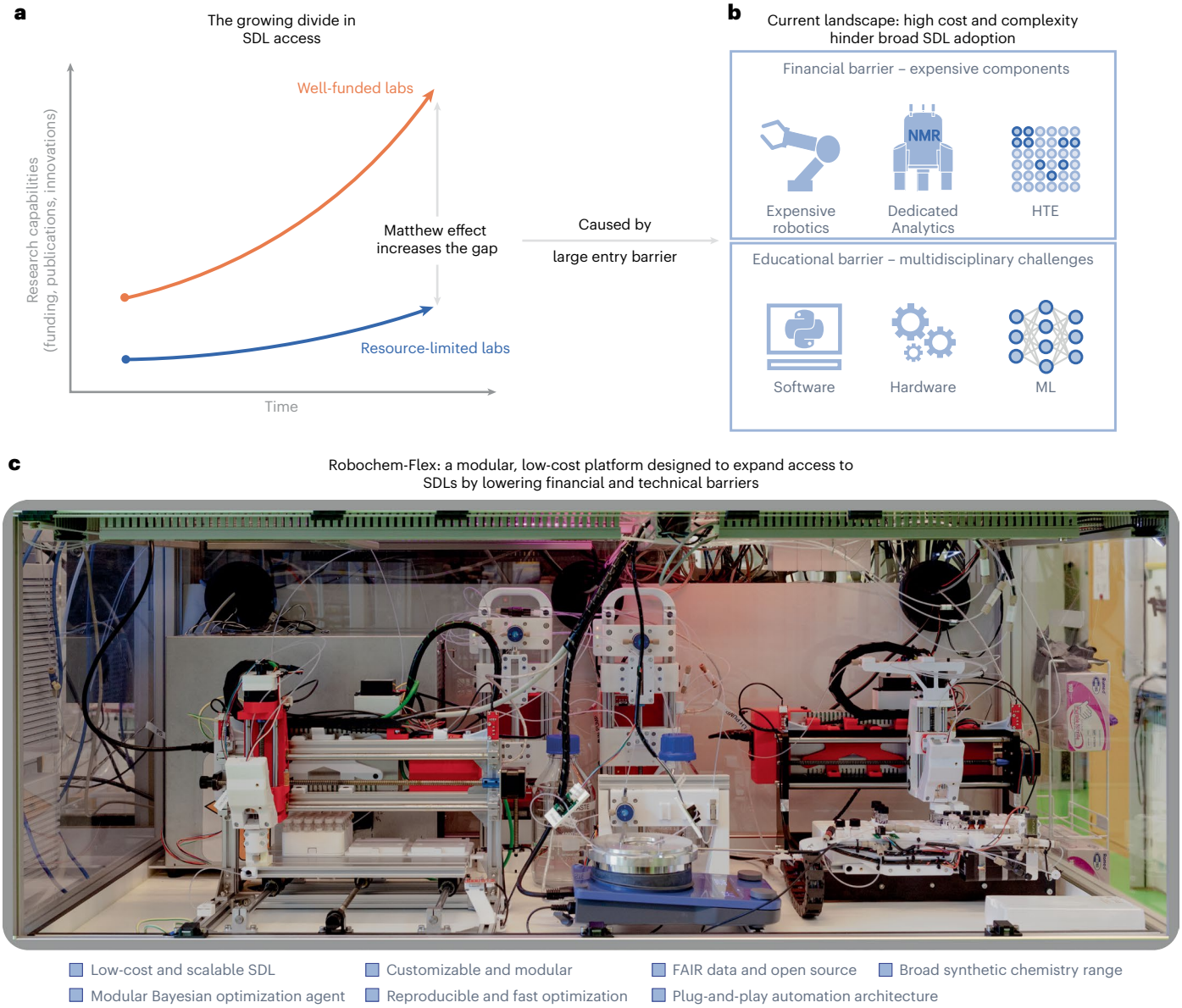


Fig. 1 | RoboChem-Flex: a low-cost SDL for closed-loop or human-in-the-loop, multi-objective optimization in synthetic organic chemistry. a–c, Shown are the challenges associated with the adoption of current SDLs: the growing divide in access to SDLs (**a**), the current high-cost/high-complexity boundary (**b**) and a picture of the RoboChem-Flex platform (**c**). FAIR, findable, accessible, interoperable, reusable.

benchmark case study is presented in Supplementary Section 3.1. Each case study demonstrates how algorithmic workflows and hardware modules can be specifically tailored to the problem at hand. In all cases, the real-world applicability of the optimized solutions was validated at preparative scale.

Optimization of pyrrole trifluoromethylation using adaptive weighted exploration and NMR spectroscopy analysis (case study 1)

Assessing closed-loop optimization using RoboChem-Flex, we configured the platform with a cost-effective, in-house-built Uflow photoreactor and an inline benchtop ^{19}F NMR spectrometer[15]. As a model reaction, we selected a photochemical trifluoromethylation, previously described by Stephenson et al.[16]. and used by MacMillan et al.[17] to benchmark the Penn batch reactor. A weighted multi-objective Bayesian optimizer was used targeting high yield and low residence time with a 9:1 preference, implementing an upper confidence bound (UCB) acquisition function. UCB balances predicted mean and model uncertainty to manage the trade-off between exploration and exploitation. This function was enhanced with an adaptive variance-driven mechanism: When three consecutive experiments produced similar yields, the optimizer switched from an exploitative to a more explorative regime by increasing the acquisition function β. Conversely, when yields were sparse, it favoured exploitation by decreasing β (Supplementary Section 2.1.3.2). This broadened its search across the chemical space, helping to better identify global maxima. While metrics such as productivity (yield/residence time) are conceptually similar, the weighted dual target approach allows prioritization of one factor, reflecting realistic priorities within a typical research environment. The optimization problem was defined over a high-dimensional mixed-variable space comprising five continuous features (reaction concentration, reagent stoichiometries and residence time), one discrete variable (light intensity), and two categorical variables (photocatalyst and oxidant). Following an initialization with 15 experiments via LHS, the optimization continued for 35 closed-loop iterations. After 13 optimization runs, the system converged on conditions delivering

https://doi.org/10.1038/s44160-026-01053-0

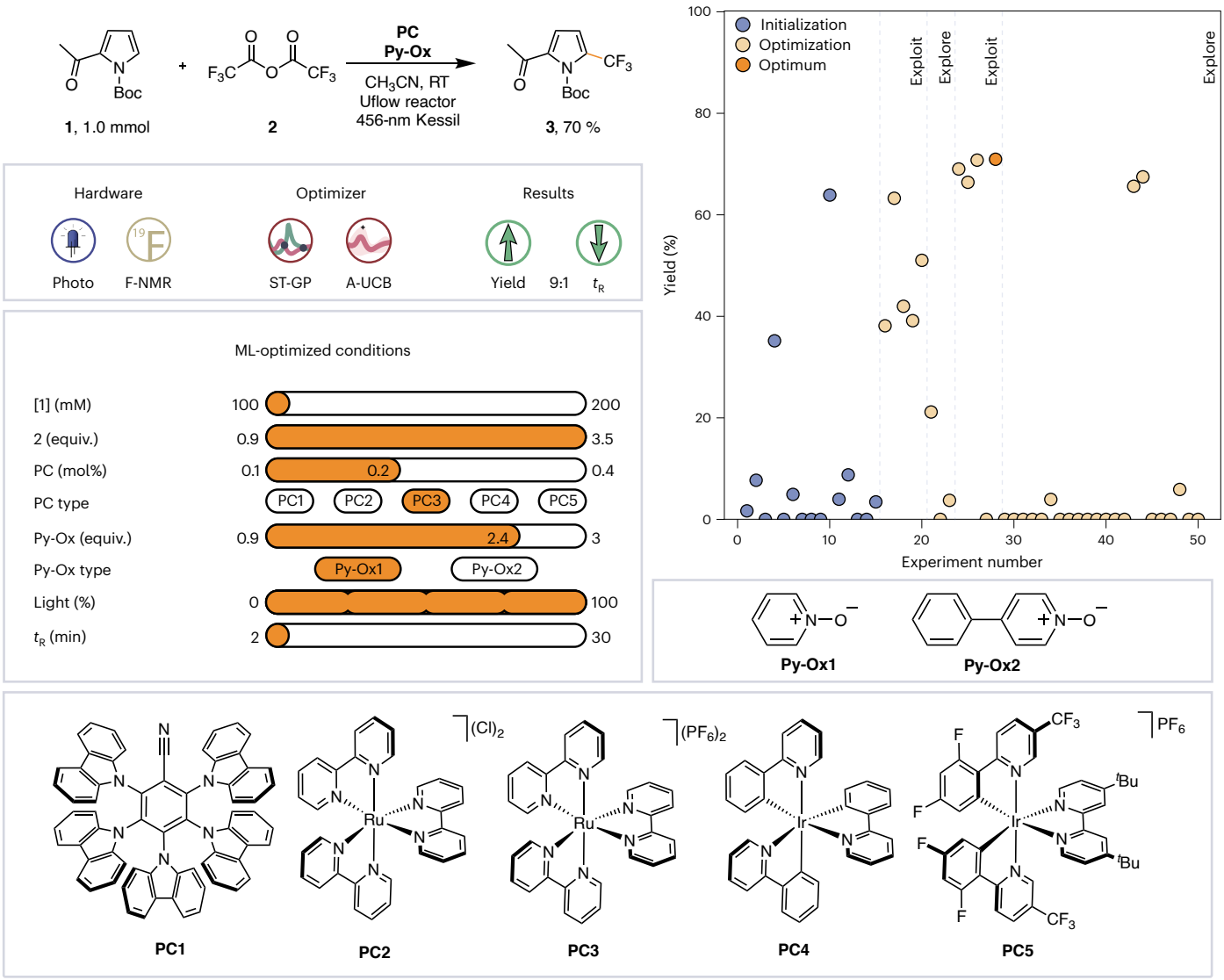


Fig. 2 | Optimization of photocatalytic trifluoromethylation of pyrrole 1 with trifluoroacetic anhydride 2 using an adaptive weighted optimization and ^{19}F NMR spectroscopy analysis. Top left: the figure presents the reaction scheme for the photocatalytic trifluoromethylation of pyrrole, including the isolated yield obtained under the optimized and scaled-up conditions. Top right: the optimization campaign is summarized (top right), highlighting the evolution of reaction performance across iterations. Middle left: the robotic platform and ML configuration, together with the conditions identified by the optimization algorithm, are shown. Bottom left and right: The categorical variables explored during the campaign are depicted as their corresponding chemical structures. PC, photocatalyst; Py-Ox, pyridine N-oxide, Boc, *tert*-butyloxycarbonyl; RT, room temperature; t_R, residence time; ST-GP, single-task Gaussian process; A-UCB, adaptive UCB.

a 71% yield with a residence time of 2 min. This triggered the adaptive behaviour of the optimizer, switching from an exploitation- to an exploration-focused bias, to validate global optimality. Over the next 21 runs, diverse regions of the space were sampled, confirming that deviations from the identified optimum consistently underperformed. This adaptive exploration strategy ensures robustness and prevents premature convergence to local maxima (Fig. 2).

To validate these results, a 1-mmol continuous-flow reaction was performed under the AI-refined conditions, yielding 70% isolated yield for **3**. This is in excellent agreement with the 71% inline NMR yield obtained on microscale (60–120 µmol) optimization. Compared with prior reports (Stephenson et al.: 71% in 5 h (ref. 18); MacMillan et al.: 64% in 3 min (ref. 17)), RoboChem-Flex identified conditions that gave comparable or improved yield while achieving reaction times of just 2 min. This highlights the ability of the platform to autonomously identify high-performing and scalable conditions for photocatalytic transformations (Extended Data Fig. 1).

Deoxygenative C–H functionalization via hypervolume optimization using HPLC analysis (case study 2)

To demonstrate the configurability of RoboChem-Flex, an inline high-performance liquid chromatography with ultraviolet-visible and mass spectrometry detection (HPLC)–UV/MS analysis module was introduced. This enabled us to tackle a more complex test case, for which we selected the photochemical C–H alkylation of pyridines via alcohol deoxygenation, originally developed by MacMillan et al.[19]. This transformation forms C(sp^3)–C(sp^2) bonds from alcohol precursors and poses a common synthetic challenge: overfunctionalization. The substrate used in this study contains two reactive C–H positions, yielding both mono- and di-alkylated products. The optimization goal was to maximize the yield for mono-substituted product **5a** while minimizing di-substitution. This requires a genuine trade-off, which is an ideal benchmark for a hypervolume-based optimization approach. Therefore, we selected a different Bayesian optimization agent. This agent uses a qEHVI (expected hypervolume improvement) acquisition

https://doi.org/10.1038/s44160-026-01053-0

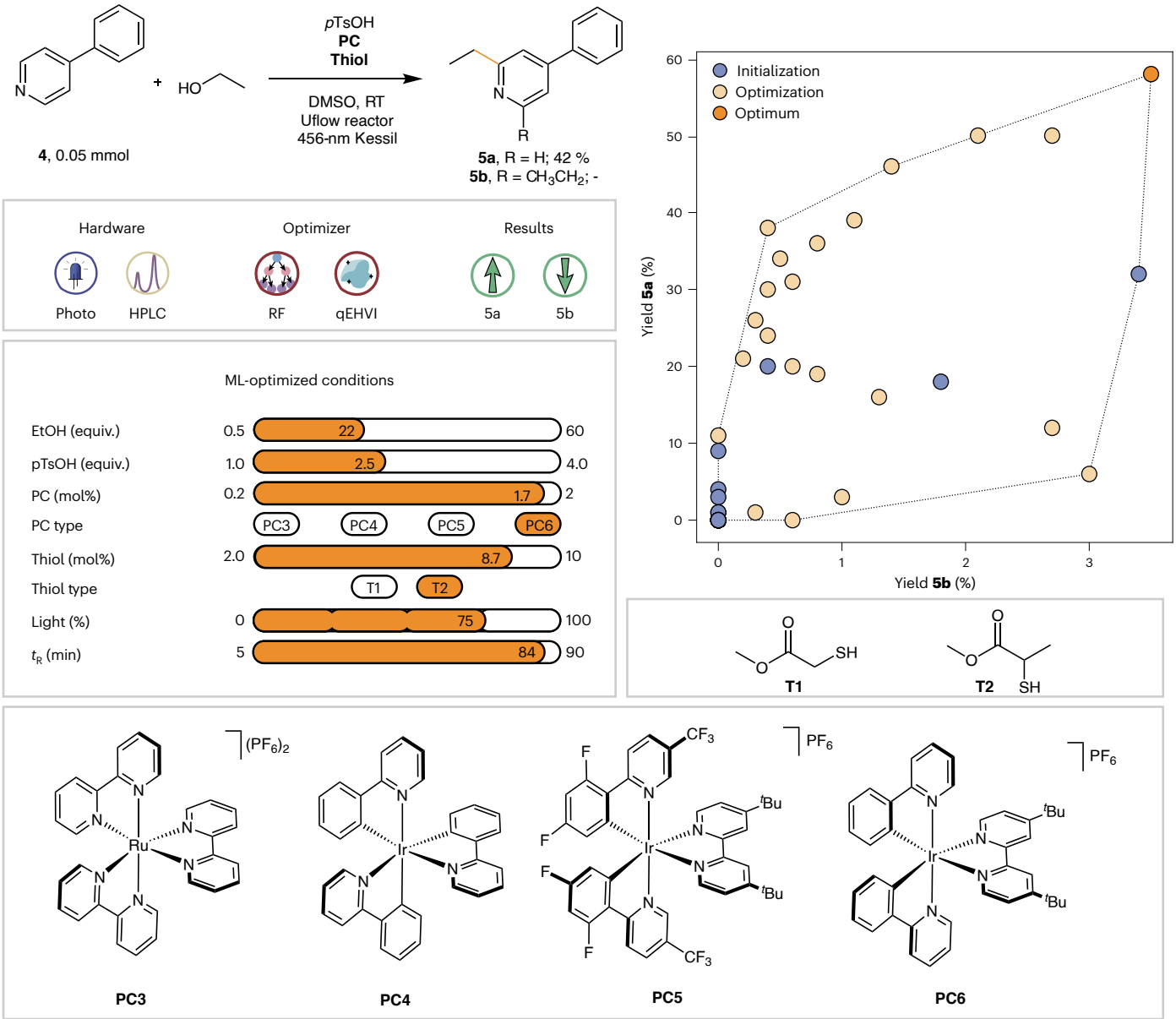


Fig. 3 | Optimization of deoxygenative C–H functionalization of 4 with ethanol via hypervolume optimization with HPLC analysis. Top left: the figure presents the reaction scheme for the deoxygenative C–H functionalization of substrate **4** with ethanol, including the isolated yields obtained under optimized conditions at scale-up. Top right: the results of the multi-objective optimization campaign are shown as a two-dimensional objective space, where the two reaction yields are plotted against each other to visualize trade-offs and identify Pareto-optimal conditions. Middle left: the robotic platform and ML configuration, together with the conditions identified by the optimization algorithm, are shown. Bottom left and right: the categorical variables explored during the optimization are depicted as their corresponding chemical structures. *p*Ts, *para*-toluenesulfonyl; PC, photocatalyst; RT, room temperature; PC, photocatalyst; ^{t}Bu, *tert*-butyl (1,1-dimethylethyl); t_R, residence time; RF, random forest.

function paired with a random forest surrogate model. While the UCB function used in the first case study offers strong control over the exploration–exploitation balance, the qEHVI acquisition function, enables more efficient multi-objective optimization in the presence of outcome trade-offs. The use of a random forest surrogate also allows more robust handling of categorical variables, demonstrating the modularity of the RoboChem-Flex Bayesian framework and its compatibility with non-Gaussian process-based strategies[20].

The chemical search space composed of five continuous features (reagent stoichiometries and residence time), one discrete (light intensity) and two categorical variables (photocatalyst and thiol co-catalyst). The optimizer was initialized with 19 LHS experiments, followed by 38 closed-loop iterations. After 31 runs, the qEHVI acquisition function converged on conditions delivering the mono-alkylated product in 58% yield with 94% selectivity (mono:di) (Fig. 3). The reaction was then scaled up 40-fold (1 mmol), affording 42% isolated yield with 94% selectivity. Although this specific substrate–alcohol combination was not studied in the original report, methanol alkylation under literature conditions gave a 65% yield and 84% selectivity. RoboChem-Flex delivered comparable selectivity and strong performance on an unreported substrate, demonstrating its capacity to tackle analytically demanding, multi-objective optimizations in new experimental configurations (Extended Data Fig. 2).

Noisy hypervolume optimization of photocatalytic isotope labelling using Raman spectroscopy (case study 3)

In-line analytical instruments typically represent the largest cost component of SDLs, yet they are essential for closed-loop automation.

https://doi.org/10.1038/s44160-026-01053-0

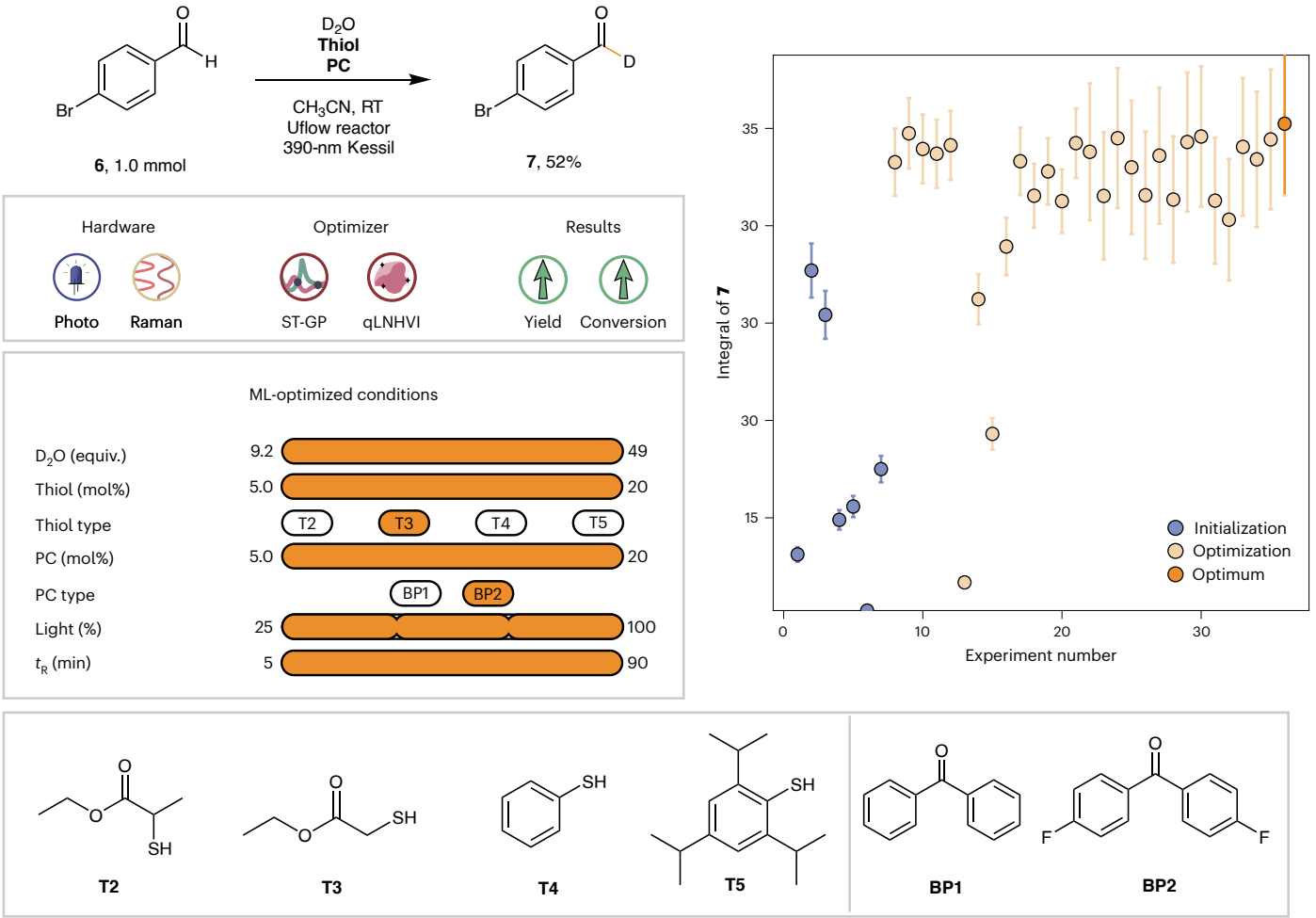
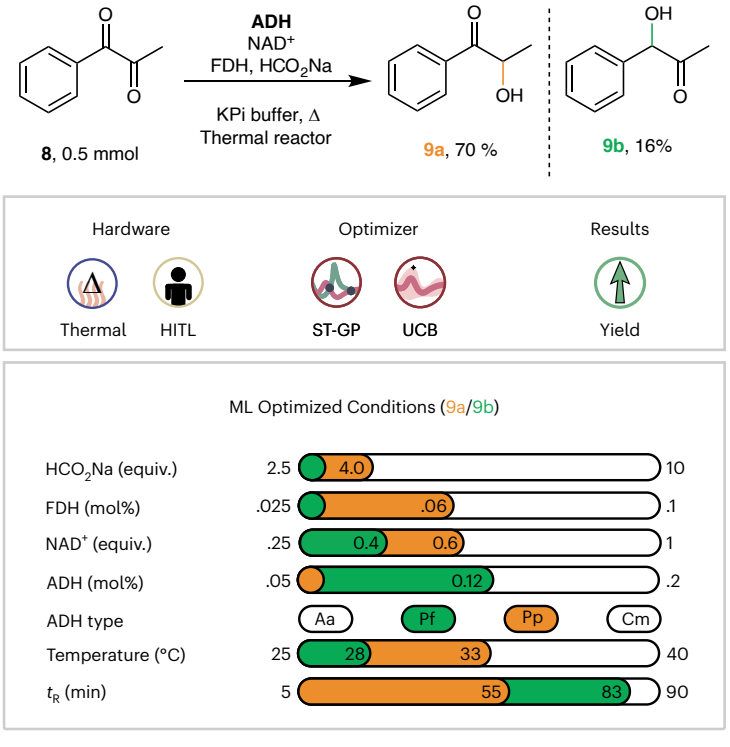


Fig. 4 | Noisy hypervolume optimization of photocatalytic isotope labelling of 6 using Raman spectroscopy. Top left: the figure presents the reaction scheme for the photocatalytic isotope labelling of **6**, including isolated yield obtained under optimized conditions at scale-up. Top right: the optimization campaign is summarized, highlighting the evolution of the reaction performance across iteration. In this plot, each data point corresponds to a single independent experiment ($n = 1$). Central values represent the measured Raman integral over the range (1,675–1,681 cm^{-1}) normalized to the integral in the range (2,900–2,970 cm^{-1}). The *y*-axis is unitless. Error bars indicate uncertainty arising from the out-of-sample prediction error of the calibration curve estimated as described in Supplementary Section 3.4.3. Middle left: the robotic platform and ML configuration, together with the conditions identified by the optimization algorithm, are shown. Bottom left and right: the categorical variables explored during the optimization are depicted as their corresponding chemical structures. PC, photocatalyst; RT, room temperature; BP, benzophenone; t_R, residence time; ST-GP, single-task Gaussian process.

Addressing this cost barrier while adhering to the ethos of RoboChem-Flex, we developed a low-cost Raman spectroscopy setup using commodity components (Supplementary Section 1.3.6), reducing costs to a fraction of commercial alternatives. Naturally, these savings come with performance trade-offs, including lower spectral resolution and reduced signal-to-noise ratios. We explicitly modelled uncertainty measurements in the optimization process following the approach of Lapkin et al. to compensate for these limitations[21]. Calibration curves were used to quantify the concentration-dependent error for both starting material and product signals (Supplementary Section 3.4.3). These error estimates were subsequently incorporated into the surrogate model as heteroskedastic noise (noise that varies across experimental conditions). In this study, the Bayesian agent used the qLogNEHVI (logarithmic noisy expected hypervolume improvement) acquisition function, which, unlike standard EHVI, combines multiple surrogate predictions based on the model variance. This makes it more robust to measurement error, prioritizing points that are genuinely informative rather than those that appear optimal due to noise[22]. As a benchmark for this configuration, we targeted a photocatalytic hydrogen-deuterium exchange on 4-bromo-benzaldehyde. This transformation is poorly suited to low-field NMR spectroscopy or HPLC detection, while Raman analysis is a viable option. Rather than relying on established high-yielding decatungstate-based protocols, our goal was to identify optimal, metal-free conditions, which are far less explored[23].

A feature space consisting of five continuous parameters (reagent stoichiometries and residence time), one discrete variable (light intensity) and two categorical variables (photocatalyst and thiol co-catalyst) was defined. The optimization engine was initialized with 7 LHS experiments, followed by 30 optimization iterations. The noise-aware workflow converged on conditions giving 70% ± 10% yield with 71% deuterium incorporation (Fig. 4). These conditions were used in an 8-fold scale-up experiment (1 mmol), which allowed isolation of the deuterated product in 52% yield with 64% deuterium incorporation (82% recovery). This outcome surpasses prior reports of 0–38% deuterium incorporation for comparable organocatalytic systems[24]. Although implementation of a stronger light source or longer reaction times may lead to further improvements, as suggested by the identified optimum being on the upper limit of the explorable space, such efforts are, however, outside the scope of this study (Extended Data Fig. 3). We rather aim to highlight how custom low-cost analytics, combined with noise-enabled Bayesian optimization, can

https://doi.org/10.1038/s44160-026-01053-0



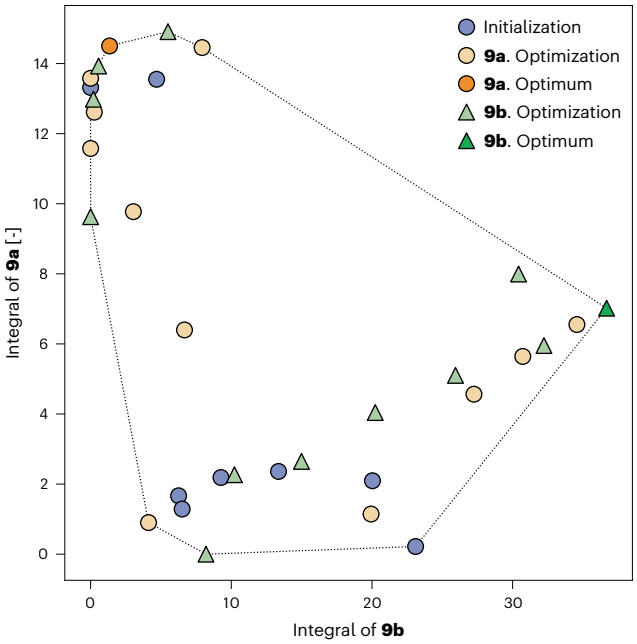


Fig. 5 | Selective enzymatic reduction of diketone 8 using a human-in-the-loop and dual acquisition function batching approach. Top left: the figure presents the reaction scheme for the enzymatic reduction of **8**, including the isolated yield obtained under the optimized and scaled-up conditions. Top right: the results of the multi-objective optimization campaign are shown as a two-dimensional objective space, where the integrals of compounds **9a** and **9b** are plotted against each other to visualize trade-offs and identify Pareto-optimal conditions. The *x* and *y*-axes are unitless. Bottom left: the robotic platform and ML configuration, together with the conditions identified by the optimization algorithm, are shown. ADH, NAD-dependent ADH; KPi, potassium phosphate buffer; t_R, residence time; HITL, human-in-the-loop; ST-GP, single-task Gaussian process. ADHs tested originated from *Aromatoleum aromaticum* (Aa)[41], *Paracoccus pantotrophus* (Pp)[42], *Pichia finlandica* (Pf)[27] and *Candida maris* (Cm)[26].

autonomously deliver competitive results in the optimization of a meaningful chemical transformation.

Selective enzymatic reduction of a diketone using human-in-the-loop and dual acquisition batching (case study 4)

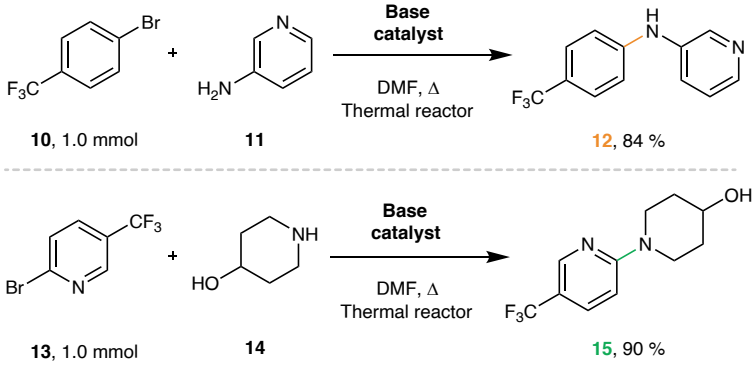
Having demonstrated analytical flexibility in photochemical reaction optimizations using in-line NMR spectroscopy, UHPLC and Raman analysis, we next sought to validate the reaction versatility of the platform in a biocatalytic transformation. Biocatalysis plays an increasingly important role in pharmaceutical synthesis, particularly for producing enantiopure compounds from achiral or prochiral substrates[25]. In this case study, we targeted the biocatalytic reduction of diketone **8** to chiral secondary alcohol **9a** using nicotinamide adenine dinucleotide (NAD)-dependent alcohol dehydrogenases (ADHs)[26]. This reaction presents an added challenge as substrate **8** contains two similar carbonyl groups, leading to the formation of both **9a** and regioisomeric side product **9b**, as an inseparable mixture. In addition, the reaction protocol uses in situ recycling of the NAD^+ coenzyme via a formate dehydrogenase enzyme (FDH), which reduces coenzyme loading but increases the dimensionality of the optimization space.

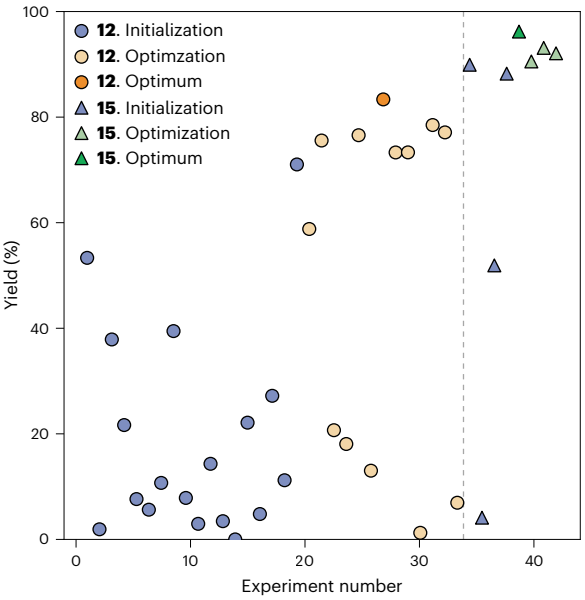
We also used RoboChem-Flex's human-in-the-loop workflow. Utilizing the fraction collector of the platform, each reaction slug was quenched and stored for deferred analysis. Quenching was performed by preloading the receiver vials with sodium dodecyl sulfate, a surfactant that denatures the ADH enzymes, halting the reaction without affecting the analytes. During early experiments, partial clogging of the tubing was occasionally observed due to protein precipitation; the issue and its resolution are discussed in Supplementary Section 3.5.3. The stored samples were subsequently analysed by UHPLC, and the raw integrals corresponding to the desired product **9a** were directly fed back into the Bayesian optimization agent, tasked with maximizing the formation of **9a**. This approach further demonstrates RoboChem-Flex's ability to operate effectively without authentic standards, making it suitable for workflows involving unknown or novel impurities.

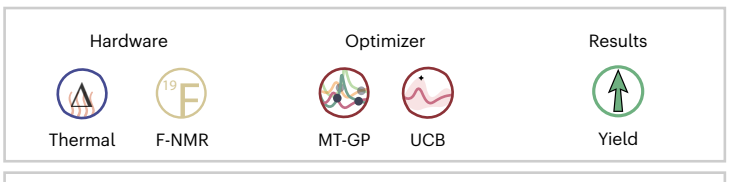
The optimization campaign began with eight initial experiments using LHS, followed by iterative batch optimization in sets of four experiments. The optimization space consisted of six continuous parameters (reagent stoichiometries, temperature and residence time) and one categorical (ADH enzyme). Standard Bayesian acquisition functions, such as UCB, typically yield a single optimal point per iteration, and batch selection is often performed using quasi-Monte Carlo strategies. However, the marginal information gain of additional batch points diminishes due to overlap in sampled regions[22,27]. To mitigate this, we implemented a hybrid batching strategy in which half of each batch was selected using the qUCB acquisition function, while the other half was selected using Monte Carlo integrated negative posterior variance (qINPV)[28]. The latter prioritizes regions of high model uncertainty to improve exploration without sacrificing optimization efficiency (Fig. 5).

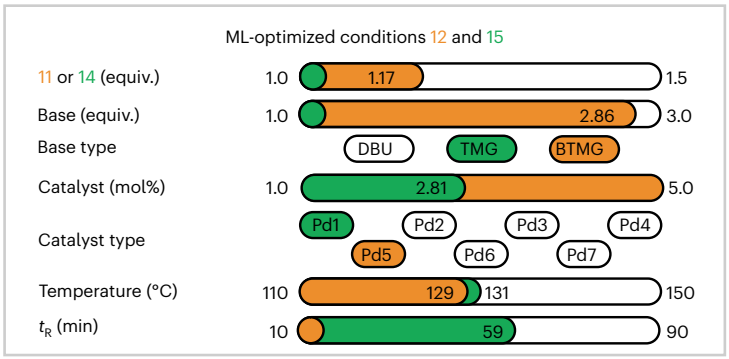
After two batches of optimization, conditions yielding the largest peak integral of product **9a** and the lowest peak integral of side-product **9b** were identified. Using these optimal conditions, a 200-fold scale-up experiment (1 mmol) was carried out, affording 58% yield of **9a** with >99% enantiomeric excess (e.e.) and only trace amounts of **9b**. Although impurity tracking is often overlooked in academic settings owing to time constraints, it is critically important in pharmaceutical and agrochemical development. Automated monitoring of impurities not only aids in suppressing their formation but also provides valuable mechanistic insight. We showcased this capability by performing a second optimization campaign, taking advantage of the data obtained

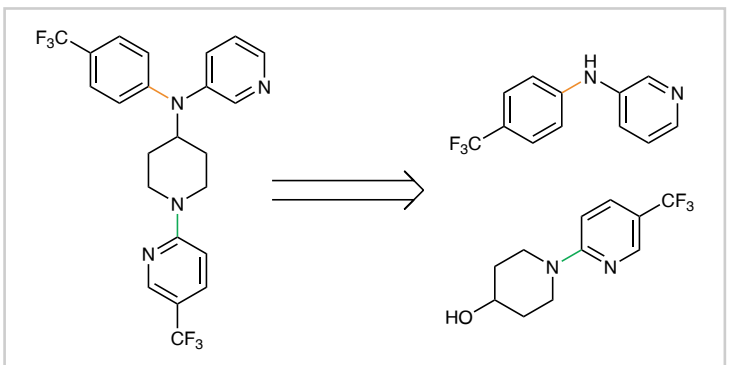
https://doi.org/10.1038/s44160-026-01053-0

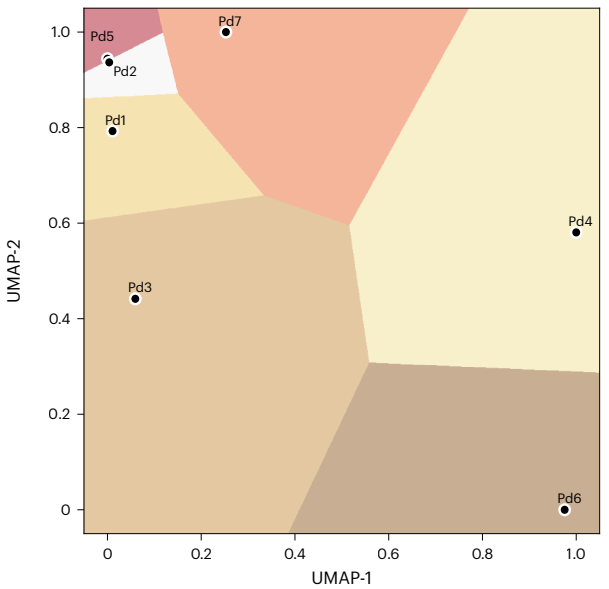












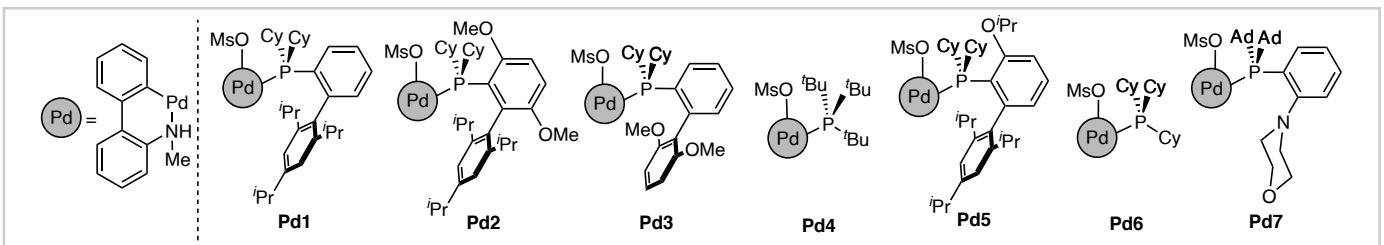


Fig. 6 | Optimization of Buchwald–Hartwig amination couplings via transfer learning and ligand featurization. Top left: the figure presents the reaction schemes for the Buchwald–Hartwig amination of substrates **10** and **13**, including their yields under optimized conditions at scale-up. Top right: the results of the optimization campaign showing reaction yield as a function of experiment number; dashed grey lines in the plot and reaction scheme represent two distinct chemical system and their corresponding optimization campaigns. Bottom left: two-dimensional UMAP projection (bottom left) of the ligand feature space, colourized by the nearest ligand as a Vornoi plot. Middle left: the robotic platform and ML configuration, together with the conditions identified by the optimization algorithm, are shown. Bottom left: compounds **12** and **15** are precursors to the pharmaceutically relevant molecule shown. Bottom: the ligands explored during the optimization are depicted as their corresponding chemical structures. DBU, 1,8-diazabicyclo[5,4,0]undec-7-ene; TMG, tetramethylguanidine; BTMG, 2-*tert*-butyl-1,1,3,3-tetramethylguanidine; Cy, cyclohexyl; Ms, mesylate; Ad, adamantane; ^{i}Pr, *iso*-propyl (propan-2-yl); ^{t}Bu, *tert*-butyl (1,1-dimethylethyl); t_R, residence time; MT-GP, multi-task Gaussian process.

https://doi.org/10.1038/s44160-026-01053-0

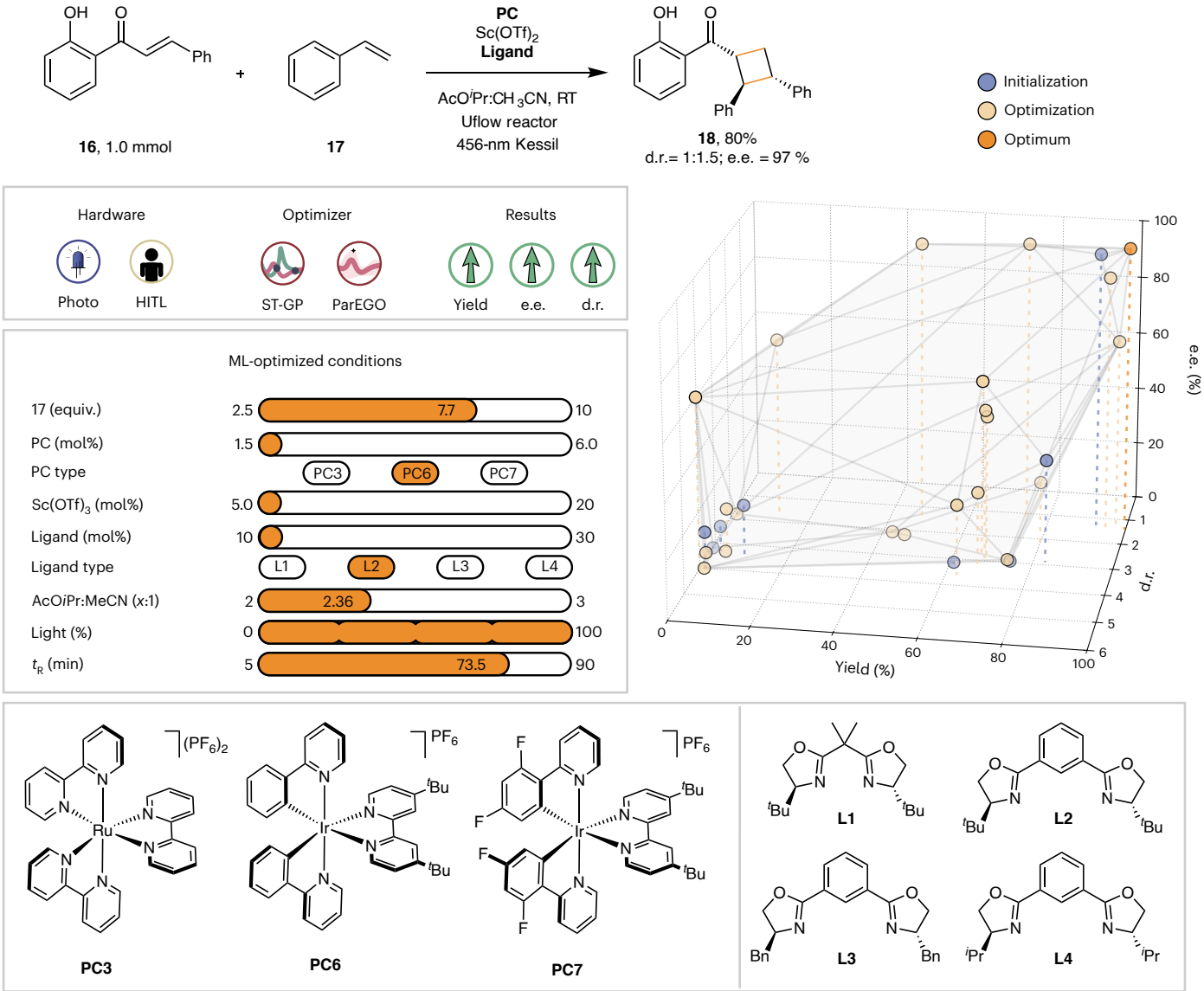


Fig. 7 | Multi-objective optimization of the enantioselective photocatalytic cycloaddition of 16 and 17 with chiral HPLC analysis. Top left: the figure presents the reaction scheme for the enantioselective photocatalytic cycloaddition of **16** and **17**, including the isolated yield obtained under the optimized and scaled-up conditions. Top right: the results of the multi-objective optimization campaign are shown as a two-dimensional objective space, where the e.e., d.r. and yield are plotted against each other to visualize trade-offs and identify Pareto-optimal conditions. Middle left: the robotic platform and ML configuration, together with the conditions identified by the optimization algorithm, are shown. Bottom: the categorical variables explored during the optimization are depicted as their corresponding chemical structures. PC, photocatalyst; Tf, triflyl; AcO^{i}Pr, isopropyl acetate; RT, room temperature; Ph, phenyl; ^{t}Bu, *tert*-butyl (1,1-dimethylethyl); ^{i}Pr, *iso*-propyl (propan-2-yl); Bn, benzyl; t_R, residence time; HITL, human-in-the-loop; ST-GP, single-task Gaussian process; ParEGO, Pareto efficient global optimization.

during optimization of **9a**, this time targeting the maximization of the side product **9b**. Within three iterations, new conditions were identified that increased the formation of **9b** (in 98% e.e.) considerably, to an extent which enabled its characterization and biological evaluation.

Optimization of Buchwald–Hartwig aminations via transfer learning and ligand featurization (case study 5)

Buchwald–Hartwig amination is a cornerstone of modern synthetic chemistry, providing a versatile coupling strategy for C–N bond formation[29]. Numerous catalyst systems have been developed to accommodate specific substrate classes and conditions, leading to high-throughput experimentation (HTE) becoming the preferred approach for catalyst screening and reaction optimization[30]. The high costs associated with HTE have driven efforts to apply machine learning (ML) models to study this reaction and reduce the number of experiments required for optimization[31].

In this case study, we developed a hybrid strategy that combines automated experimentation with ligand featurization to efficiently screen catalysts while minimizing the experimental workload. Based on the library of monodentate organophosphorus ligands compiled by Aspuru-Guzik et al., we selected seven commercially available Buchwald precatalysts[32]. Each ligand in the library was characterized by a set of density functional theory-derived descriptors, which we condensed into a two-dimensional chemical space using the Uniform Manifold Approximation and Projection (UMAP) algorithm[33,34]. This allowed the AI agent to learn which ligand features most effectively differentiate catalyst performance within the chemical space. To further showcase the capabilities of the RoboChem-Flex ML engine, we pursued the sequential optimization of two coupling reactions (leading to products **12** and **15**) using a transfer learning framework. Notably, both products feature as key building blocks of a pharmaceutically relevant molecule[35].

An extensive optimization campaign was first conducted for the synthesis of product **12** within a parameter space of seven continuous (reagent stoichiometries, temperature, residence time and two UMAP dimensions for catalyst selection) and one categorical (base) variable. The identified optimum led to 84% isolated yield at 1-mmol scale (14-fold increase), in excellent agreement with the ML-refined optimization outcome. The resulting dataset was then transferred to a multi-task Gaussian process model[36,37] (Supplementary Section 2.1.3.6), which was initialized with just four new experiments using substrate **13** for the synthesis of product **15**. The extensive experimental data collected during the optimization of **12** provided the multi-task model with a rich understanding of the accessible chemical space, while avoiding overfitting to the initial reaction system. As a result, the second optimization did not begin from scratch; instead, the model leveraged cross-task relationships to rapidly identify conditions likely to succeed for the new substrate (Fig. 6). Consequently, the model converged on optimal conditions after only four additional closed-loop iterations, affording product **15** in 90% isolated yield at 1-mmol scale. This is particularly noteworthy given that the two products were derived from either primary or secondary amines, and the algorithm successfully identified the most suitable ligand for each transformation[38].

This case study demonstrates how transfer learning and ligand featurization can be seamlessly integrated into the RoboChem-Flex ecosystem to accelerate catalyst selection and reduce experimental burden in complex cross-coupling optimizations.

Multi-objective optimization of an enantioselective photocatalytic [2 + 2] cycloaddition using chiral HPLC (case study 6)

As a final demonstration of RoboChem-Flex's capabilities, we conducted an experimental campaign targeting a large feature space with three simultaneous objectives. We selected the enantioselective [2 + 2] photocycloaddition protocol developed by Yoon et al. as an ideal test case, where yield, diastereomeric ratio (d.r.) and e.e. must all be maximized[39,40]. This case study also highlighted RoboChem-Flex's flexibility in integrating with shared analytical infrastructure: using a human-in-the-loop workflow, we performed chiral HPLC analysis offline, avoiding the need for a costly dedicated inline set-up.

The reaction was optimized by training three parallel surrogate models, each exposed to the full set of conditions but trained to predict a single objective. One candidate experiment per model was generated using a UCB acquisition function, focusing exclusively on maximising that specific model's objective. These models where then combined using the logarithmic noisy Pareto efficient global optimization (qLogNParEGO) strategy to generate three additional points that capture trade-offs across objectives, enabling efficient exploration of the multi-objective Pareto front (Supplementary Section 2.1.3.7). Because the models operated independently, similar or identical conditions could occasionally be proposed within the same batch. To mitigate this redundancy, a Euclidean distance filter was applied to eliminate closely related candidates.

This campaign explored the largest feature space in our study, consisting of six continuous parameters (reagent stoichiometries, solvent ratio and residence time), one discrete variable (light intensity) and two categorical variables (chiral ligand and photocatalyst), forming a combinatorially complex search space. The optimization began with eight LHS experiments, followed by four batches of iterative optimization. RoboChem-Flex identified conditions that delivered >99% yield, a 1:1.4 d.r. and >99% e.e. (Fig. 7). These conditions were successfully scaled up 34-fold (0.50 mmol), affording the crude product in 80% yield, a 1:1.5 d.r. and 97% e.e. These results compare favourably with literature benchmarks, which report 88% yield, a 1:2 d.r. and >99% e.e. achieved in 20 h under batch conditions[40] (Extended Data Fig. 4). Notably, the RoboChem-Flex campaign identified conditions with reduced loadings of photocatalyst, chiral ligand and styrene compared with the literature, without compromising reaction performance. This case study demonstrates RoboChem-Flex's ability to efficiently navigate large chemical spaces, optimize multiple objectives and successfully translate complex reaction conditions from batch to flow.

Conclusion

RoboChem-Flex provides an affordable and modular platform that lowers the barriers to autonomous chemical experimentation. By delivering high-performance capabilities at a fraction of the cost of traditional SDLs, it empowers resource-limited research groups to access automation previously restricted to well-funded institutions. Through its scalable, plug-and-play design and demonstrated versatility across diverse chemical challenges, RoboChem-Flex aims to foster broader innovation and ensure that scientific progress is driven by creativity and ambition, not financial constraints. Further extension of this platform aims to promote greater collaboration and adapt RoboChem-Flex to targeted needs of other groups. By fostering an open, interoperable network of SDLs, we aim to accelerate collective scientific discovery.

Methods

General procedure for the pyrrole trifluoro-methylation

Under an atmosphere of nitrogen, pyrrole **2** (1.00 mmol), $Ru(bpy)_3(PF_6)_2$ (bpy = 2,2′-bipyridine) (0.0184 mmol, 0.184 mol%) and pyridine *N*-oxide (2.41 mmol, 2.41 equiv.) were dissolved in CH_3CN (10.0 ml) to which trifluoroacetic anhydride (3.50 mmol, 3.50 equiv.) was added. The solution was loaded in a syringe and pumped through a UFlow reactor (3.95 ml) at 1.98 ml min^{-1} (residence time 2.00 min) while irradiating with a 456-nm Kessil lamp at 75% intensity. Upon completion, the solvent and volatiles were removed under reduced pressure. The resulting product was then purified via column chromatography (75:25 *n*-pentane:CH_2Cl_2) to give pyrrole **3** as an off-white oil (0.702 mmol, 70% yield).

General procedure for the deoxygenative C–H functionalization of pyridine

Under an atmosphere of nitrogen, 4-phenylpyridine **4** (0.50 mmol), p-TsOH (1.23 mmol, 2.45 equiv.) and $[Ir(dtbbpy)(ppy)_2]PF_6$] (dtbbpy =4,4′-di-*tert*-butyl-2,2′-dipyridyl; ppy = 2-phenylpyridine) (0.0086 mmol, 1.71 mol%) were dissolved in anhydrous dimethyl sulfoxide (DMSO) (4.36 ml), to which ethanol (0.64 ml, 11 mmol, 22 equiv.) was added giving a total volume of 5.0 ml. In a separate vial, ethyl 2-mercaptopropanoate (0.044 mmol, 8.7 mol%) was dissolved in anhydrous DMSO (5.0 ml). Each solution was transferred to a foil-wrapped syringe and co-pumped through a T-junction into a UFlow reactor (3.95 ml) at 0.047 ml min^{-1} (residence time 84.3 min) while irradiating with a 456-nm Kessil lamp at 75% intensity. Upon completion, the solvent and volatiles were removed under reduced pressure. The resulting product was then purified via column chromatography (100:0 to 80:20 *n*-pentane: EtOAc) to give functionalized pyridine **5a** as an off-white oil (0.210 mmol, 42% yield).

General procedure for photocatalytic isotope labelling

Under an atmosphere of nitrogen, 4-bromobenzaldehyde (1.00 mmol), benzophenone (0.200 mmol, 20 mol%) and ethyl thioglycolate (0.200 mmol, 20 mol%) were dissolved in anhydrous CH_3CN (3 ml), to which D_2O (1 ml, 50 mmol, 50 equiv.) was added. The solution was loaded into a syringe and pumped through a UFlow reactor (3.95 ml) at 0.044 ml min^{-1} (residence time 90.0 min) while irradiating with a 390-nm Kessil lamp at 100% intensity. Upon completion, the solvent and volatiles were removed under reduced pressure. The resulting product was then purified via column chromatography (100:0 to 95:5 *n*-pentane:diethyl ether) to give the benzaldehyde **7** as a white solid composed of the deuterated and non-deuterated products (0.152 g, 82% mass recovery, 64% deuterium incorporation by ^1H NMR spectroscopy).

General procedure for enzymatic reduction of distal carbonyls

1-Phenylpropane-1,2-dione (1.00 mmol) was dissolved in anhydrous DMSO (10 ml) and loaded in a 10-ml syringe. Separately sodium formate (4.03 mmol, 4.03 equiv.) enzyme Pp-ADH (0.50 µmol, 0.050 mol%), NADH (0.589 mmol, 0.589 equiv.) and FDH enzyme (0.63 µmol, 0.063 mol%) were dissolved in 50 mM phosphate buffer solution (90 ml; pH 8). The latter solution was loaded into two 50-ml syringes (45 ml each). The solutions were co-pumped through a T-junction a thermal reactor (3.55 ml) at 0.065 ml min^{-1} (residence time 54.6 min) while heating to 33 °C. The output stream was collected in a round-bottom flask containing 5% (m/V) sodium dodecyl sulfate aqueous solution (10 ml). Upon completion, the mixture was extracted twice with 80 ml EtOAc. The organic phases were then combined, washed with brine and dried over $MgSO_4$. The solvent and volatiles were removed under reduced pressure, and the resulting product was purified via column chromatography (100:0 to 80:20 *n*-pentane:EtOAc) to give the reduced product **9a** as a yellow oil (0.57 mmol 57%).

General procedure for enzymatic reduction of benzylic carbonyls

1-Phenylpropane-1,2-dione (0.50 mmol) was dissolved in anhydrous DMSO (5 ml) and loaded in a 10-ml syringe. Separately sodium formate (1.25 mmol, 2.50 equiv.) enzyme Pf-ADH (0.61 µmol, 0.122 mol%), NADH (0.143 g, 0.22 mmol, 0.44 equiv.) and enzyme FDH (0.16 µmol, 0.032 mol%) were dissolved in 50 mM phosphate buffer solution (45 ml; pH 8) and loaded into a 50-ml syringe. The solutions were co-pumped through a T-junction a thermal reactor (3.55 ml) at 0.043 ml min^{-1} (residence time 83.0 min) while heating to 29 °C. The output stream was collected in a round-bottom flask containing 5% (m/V) sodium dodecyl sulfate aqueous solution (5 ml). Upon completion, the mixture was extracted twice with 40 ml EtOAc, and the organic phases were then combined, washed with brine and dried over $MgSO_4$. The solvent and volatiles were removed under reduced pressure, and the resulting product was purified via column chromatography (100:0 to 80:20 *n*-pentane:EtOAc) to give a 49:51 mixture (calculated by ^{1}H NMR spectroscopy) of regioisomers **9a**:**9b** as a yellow oil (combined 34%; **9b** 17%).

General procedure for the Buchwald–Hartwig amination of primary amines

Under an atmosphere of nitrogen, 1-bromo-4-(trifluoromethyl) benzene (1.00 mmol), pyridin-3-amine (1.17 mmol, 1.17 equiv.), 2-(*tert*-butyl)-1,1,3,3-tetramethylguanidine (2.86 mmol, 2.86 equiv.) and EPhos-Pd-G4 (0.05 mmol, 0.05 equiv.) were dissolved in degassed dimethylformamide (DMF) (10.0 ml). The solution was loaded into a syringe and pumped through a thermal reactor (3.05 ml) at 0.269 ml min^{-1} (residence time 11.4 min) while heating to 129 °C. Upon completion, the solvent and volatiles were removed under reduced pressure. The resulting product was then purified via column chromatography (100:0 to 80:20 *n*-pentane:EtOAc) to give pyridine **12** as a white solid (0.84 mmol, 84% yield).

General procedure for the Buchwald–Hartwig amination of secondary amines

Under an atmosphere of nitrogen, 2-bromo-5-(trifluoromethyl) pyridine (1.00 mmol), piperidin-4-ol (1.06 mmol, 1.06 equiv.), 1,1,3,3-tetramethylguanidine (1.00 mmol, 1.00 equiv.) and XPhos-Pd-G4 (0.028 mmol, 0.028 equiv.) were dissolved in degassed DMF (10.0 ml). The solution was loaded into a syringe and pumped through a thermal reactor (3.05 ml) at 0.052 ml min^{-1} (residence time 58.8 min) while heating to 131 °C. Upon completion, the solvent and volatiles were removed under reduced pressure. The resulting product was then purified via column chromatography (100:0 to 80:20 *n*-pentane:EtOAc) to give pyridine **15** as a white solid (0.90 mmol, 90% yield).

General procedure for the enantioselective [2 + 2] photocycloaddition

Under an atmosphere of nitrogen, (*E*)-1-(2-hydroxyphenyl)-3-phenylprop-2-en-1-one (0.50 mmol), styrene (3.84 mmol, 7.69 equiv.) and 2,6-bis((*S*)-4-(*tert*-butyl)-4,5-dihydrooxazol-2-yl)pyridine (0.05 mmol, 10 mol%) were dissolved in a 7:3 mixture of ^{i}PrOAc and MeCN (11.21 ml) and loaded into a syringe. Separately, $[Ir(dtbbpy)(ppy)_2]PF_6$ (0.007 mmol, 1.5 mol%) was dissolved in a 7:3 mixture of *i*PrOAc and MeCN (2.5 ml) and loaded in a 5-ml syringe. Finally, $Sc(OTf)_3$ (0.012 g, 0.025 mmol, 5 mol%) was dissolved in a 7:3 mixture ^{i}PrOAc and MeCN (2.5 ml) and loaded in a third syringe. The solutions were co-pumped through a T-junction to a photochemical reactor (3.95 ml) at 0.054 ml min^{-1} (residence time 73.5 min) while irradiating with a 456-nm Kessil lamp at 100% intensity. Upon completion, the crude mixture was analysed directly by HPLC to identify a yield of 80% and a d.r. of 1:1.5. The major diastereomer was then isolated via HPLC, reduced and redissolved in 1:1 heptane:methyl *tert*-butyl ether and subsequently analysed by chiral HPLC to identify an e.e of 96%.

Data availability

All data supporting the findings of this study are available within the Article and its Supplementary Information. Benchmarking simulation results, further statistical analysis, spectral data and optimization datasets are provided in the Supplementary Information. Additional datasets, including raw simulation outputs and raw optimization data, are available via GitHub at https://github.com/Noel-Research-Group/Robochem_Flex.

Code availability

All code used in this study is openly available via GitHub at https://github.com/Noel-Research-Group/Robochem_Flex. This includes, but is not limited to, ML and optimization code, graphical user interface software, device firmware and operational control code, 3D printing design files, schematics for hardware, and scripts used to analyse and generate all data figures presented in the Article and its Supplementary Information.

References

1. Abolhasani, M. & Kumacheva, E. The rise of self-driving labs in chemical and materials sciences. *Nat. Synth.* **2**, 483–492 (2023).
2. Bayley, O., Savino, E., Slattery, A. & Noël, T. Autonomous chemistry: navigating self-driving labs in chemical and material sciences. *Matter* **7**, 2382–2398 (2024).
3. Tom, G. et al. Self-driving laboratories for chemistry and materials science. *Chem. Rev.* **124**, 9633–9732 (2024).
4. Slattery, A. et al. Automated self-optimization, intensification, and scale-up of photocatalysis in flow. *Science* https://doi.org/10.1126/science.adj1817 (2024).
5. Bol, T., de Vaan, M. & van de Rijt, A. The Matthew effect in science funding. *Proc. Natl Acad. Sci. USA* **115**, 4887–4890 (2018).
6. Lo, S. et al. Review of low-cost self-driving laboratories in chemistry and materials science: the 'frugal twin' concept. *Digit. Discov.* **3**, 842–868 (2024).
7. Baird, S. G. & Sparks, T. D. What Is a minimal working example for a self-driving laboratory?. *Matter* **5**, 4170–4178 (2022).
8. Beker, W. et al. Machine learning may sometimes simply capture literature popularity trends: a case study of heterocyclic Suzuki–Miyaura coupling. *J. Am. Chem. Soc.* **144**, 4819–4827 (2022).
9. Maloney, M. P. et al. Negative data in data sets for machine learning training. *Org- Lett-* **25**, 2945–2947 (2023).
10. Strieth-Kalthoff, F. et al. Machine learning for chemical reactivity: the importance of failed experiments. *Angew. Chem. Int. Ed.* https://doi.org/10.1002/anie.202204647 (2022).
11. Capel, A. J., Rimington, R. P., Lewis, M. P. & Christie, S. D. R. 3D printing for chemical, pharmaceutical and biological applications. *Nat. Rev. Chem.* **2**, 422–436 (2018).

12. Capaldo, L., Wen, Z. & Noël, T. A field guide to flow chemistry for synthetic organic chemists. *Chem. Sci.* **14**, 4230–4247 (2023).
13. Prabhu, G. R. D. & Urban, P. L. Elevating chemistry research with a modern electronics toolkit. *Chem. Rev.* **120**, 9482–9553 (2020).
14. Balandat, M. et al. BoTorch: a framework for efficient Monte-Carlo Bayesian optimization. In *34th Conference on Neural Information Processing Systems (NeurIPS 2020)* https://proceedings.neurips.cc/paper/2020/file/f5b1b89d98b7286673128a5fb112cb9a-Paper.pdf (2020).
15. Masson, T. M., Zondag, S. D. A., Schuurmans, J. H. A. & Noël, T. Open-source 3D printed reactors for reproducible batch and continuous-flow photon-induced chemistry: design and characterization. *React. Chem. Eng.* **9**, 2218–2225 (2024).
16. Beatty, J. W., Douglas, J. J., Cole, K. P. & Stephenson, C. R. J. A scalable and operationally simple radical trifluoromethylation. *Nat. Commun.* **6**, 1–6 (2015).
17. Le, C. C. et al. A general small-scale reactor to enable standardization and acceleration of photocatalytic reactions. *ACS Cent. Sci.* **3**, 647–653 (2017).
18. Beatty, J. W. et al. Photochemical perfluoroalkylation with pyridine *N*-oxides: mechanistic insights and performance on a kilogram scale. *Chem* **1**, 456–472 (2016).
19. Jin, J. & MacMillan, D. W. C. Alcohols as alkylating agents in heteroarene C–H functionalization. *Nature* **525**, 87–90 (2015).
20. Ru, B., Alvi, A. S., Nguyen, V., Osborne, M. A. & Roberts, S. J. Bayesian optimisation over multiple continuous and categorical inputs. In *Proc. 37th International Conference on Machine Learning (ICML'20)* 8276–8285 (2020).
21. Zhang, J., Sugisawa, N., Felton, K. C., Fuse, S. & Lapkin, A. A. Multi-objective Bayesian optimisation using *q*-noisy expected hypervolume improvement (*q*NEHVI) for the Schotten–Baumann reaction. *React. Chem. Eng.* **9**, 706–712 (2024).
22. Daulton, S., Balandat, M. & Bakshy, E. Parallel Bayesian optimization of multiple noisy objectives with expected hypervolume improvement. In *Proc. 35th International Conference on Neural Information Processing Systems (NIPS '21)* 2187–2200 (Curran, 2021).
23. Dong, J. et al. Formyl-selective deuteration of aldehydes with D_2O via synergistic organic and photoredox catalysis. *Chem. Sci.* **11**, 1026–1031 (2020).
24. Kuang, Y. et al. Visible light driven deuteration of formyl C–H and hydridic C(sp^3)–H bonds in feedstock chemicals and pharmaceutical molecules. *Chem. Sci.* **11**, 8912–8918 (2020).
25. Bayer, T., Wu, S., Snajdrova, R., Baldenius, K. & Bornscheuer, U. T. An update: enzymatic synthesis for industrial applications. *Angew. Chem. Int. Ed.* **64**, e202505976 (2025).
26. Damian, M. & Mutti, F. G. Two *anti*-prelog NAD-dependent alcohol dehydrogenases with broad substrate scope and excellent enantioselectivity. *Eur. J. Org. Chem.* https://doi.org/10.1002/ejoc.202300734 (2023).
27. Srinivas, N., Krause, A., Kakade, S. M. & Seeger, M. W. Information-theoretic regret bounds for gaussian process optimization in the bandit setting. *IEEE Trans. Inf. Theory* **58**, 3250–3265 (2012).
28. Gonzalez, J., Dai, Z., Hennig, P. & Lawrence, N. Batch Bayesian optimization via local penalization. In *Proc. 19th International Conference on Artificial Intelligence and Statistics (AISTATS) 2016* 648–657 (2016).
29. Ruiz-Castillo, P. & Buchwald, S. L. Applications of palladium-catalyzed C–N cross-coupling reactions. *Chem. Rev.* **116**, 12564–12649 (2016).
30. Isbrandt, E. S., Sullivan, R. J. & Newman, S. G. High throughput strategies for the discovery and optimization of catalytic reactions. *Angew. Chem. Int. Ed.* **58**, 7180–7191 (2019).
31. Ahneman, D. T., Estrada, J. G., Lin, S., Dreher, S. D. & Doyle, A. G. Predicting reaction performance in C–N cross-coupling using machine learning. *Science* **360**, 186–190 (2018).
32. Gensch, T. et al. A comprehensive discovery platform for organophosphorus ligands for catalysis. *J. Am. Chem. Soc.* **144**, 1205–1217 (2022).
33. McInnes, L., Healy, J., Saul, N. & Großberger, L. UMAP: uniform manifold approximation and projection. *J. Open Source Softw.* **3**, 861 (2018).
34. Kariofillis, S. K. et al. Using data science to guide aryl bromide substrate scope analysis in a ni/photoredox-catalyzed cross-coupling with acetals as alcohol-derived radical sources. *J. Am. Chem. Soc.* **144**, 1045–1055 (2022).
35. Keenan, M. et al. Two analogues of fenarimol show curative activity in an experimental model of chagas disease. *J. Med. Chem.* **56**, 10158–10170 (2013).
36. Bonilla, E. V., Chai, M. A. & Williams, C. K. I. Multi-task Gaussian process prediction. In *Proc. 21st International Conference on Neural Information Processing Systems (NIPS'07)* 153–160 (Curran, 2007).
37. Swersky, K., Snoek, J. & Adams, R. P. Multi-task Bayesian optimization. In *Proc. 26th International Conference on Neural Information Processing Systems (NIPS'13)* https://papers.nips.cc/paper_files/paper/2013/file/f33ba15effa5c10e873bf3842afb46a6-Paper.pdf (2013).
38. Maiti, D., Fors, B. P., Henderson, J. L., Nakamura, Y. & Buchwald, S. L. Palladium-catalyzed coupling of functionalized primary and secondary amines with aryl and heteroaryl halides: two ligands suffice in most cases. *Chem. Sci.* **2**, 57–68 (2010).
39. Daub, M. E. et al. Enantioselective [2+2] cycloadditions of cinnamate esters: generalizing Lewis acid catalysis of triplet energy transfer. *J. Am. Chem. Soc.* **141**, 9543–9547 (2019).
40. Miller, Z. D., Lee, B. J. & Yoon, T. P. Enantioselective crossed photocycloadditions of styrenic olefins by Lewis acid catalyzed triplet sensitization. *Angew. Chem. Int. Ed.* **56**, 11891–11895 (2017).
41. Mutti, F. G., Knaus, T., Scrutton, N. S., Breuer, M. & Turner, N. J. Conversion of alcohols to enantiopure amines through dual-enzyme hydrogen-borrowing cascades. *Science* **349**, 1525–1529 (2015).
42. Corrado, M. L., Knaus, T. & Mutti, F. G. Regio- and stereoselective multi-enzymatic aminohydroxylation of β-methylstyrene using dioxygen, ammonia and formate. *Green Chem.* **21**, 6246–6251 (2019).

Acknowledgements

We acknowledge the generous funding provided by the Dutch Research Council (NWO) under the Open Technology Program (Multi-Modal Photochemistry, grant number 18433, A.S. and T.N) and the Talent Program VICI (SynthBot, grant number 20453, T.N). T.d.H. and T.N. also acknowledge funding from Taskforce for Applied Research SIA, part of NWO (Light-Up project). Furthermore, we extend our gratitude for the generous funding received from the European Union H2020 research and innovation programme under the European Research Council program (FlowHAT, grant number 101044355, F.L., M.C. and T.N.), the European Innovation Council project (reaCtor, grant number 101099405, E.S. and T.N.), HORIZON-CL5-2021-D2-01 (Photo2Fuel, grant number 101069357, M.C. and T.N.), HORIZON-HLTH-2021-IND-07 (SusPharma, grant number 101057430, T.N.), MSCA Doctoral Networks (GreenDigiPharma, grant number 101073089, T.N.) and MSCA Individual Fellowship program (SELECTFLOW, grant number 101061835, J.S.-O.). N.I. thanks the PON REACT-EU programme, Olon S.p.A. and the Erasmus+ Traineeship programme for the financial support. This research received funding from the Molecular and Material Design Technology Hub of the Faculty of Science of the University of Amsterdam (A.F.G.G. and T.N.). F.G.M. thanks the Dutch Research Council (NWO) for an XS grant (grant number OCENW.XS24.3.183), and the EU Horizon Europe MSCA for YcascadE (grant number 101153173) and SuBStrAte (grant number 101208425) grants.

Author contributions

S.P., E.S., O.M.B. and T.N. jointly conceived the project. S.P. led the development of the experimental hardware, firmware and control software; E.S. led the development of the ML framework, software infrastructure and benchmarking methodologies; O.M.B. led analytical method development and analytical software. S.P., E.S. and O.M.B. all contributed to the software and hardware development and to the design and execution of the optimization campaigns; M.V., M.C., P.S., J.L. and F.L. designed and executed the preparatory chemical reaction studies, as well as the scale-up and validation experiments; M.D. and V.T. designed and supported the biocatalytic aspects of the study, including enzyme selection and production; N.I. contributed to early stage hardware testing; S.P., O.M.B., M.V., M.C., F.L., A.S., J.S.-O., T.d.H. and T.N. jointly selected chemical systems to be tested on the platform; T.d.H., R.A.H.P., A.F.G.G., F.G.M. and T.N. provided supervision and scientific guidance; S.P., E.S., O.M.B. and T.N. drafted the manuscript with input from all authors; T.N. secured funding; all authors reviewed and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s44160-026-01053-0.

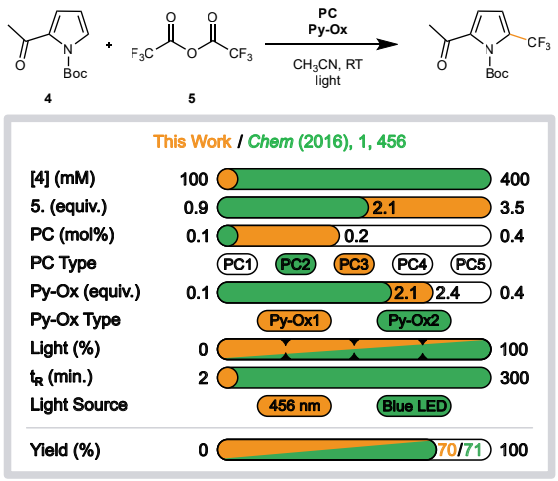
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44160-026-01053-0.

Correspondence and requests for materials should be addressed to Timothy Noël.

Peer review information *Nature Synthesis* thanks Adrian Amador and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Peter Seavill, in collaboration with the *Nature Synthesis* team.

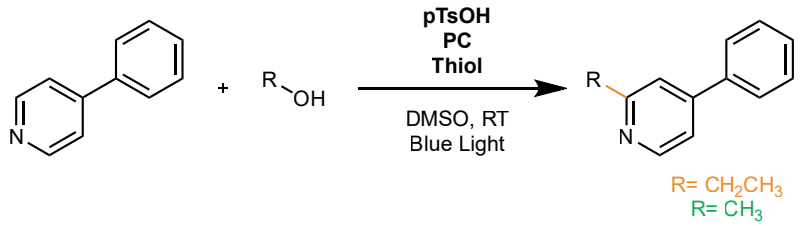
[1]Flow Chemistry Group, Van't Hoff Institute for Molecular Sciences, University of Amsterdam, Amsterdam, the Netherlands. [2]Analytical Chemistry Group, Van't Hoff Institute for Molecular Sciences, University of Amsterdam, Amsterdam, the Netherlands. [3]Biocatalysis Group, Van't Hoff Institute for Molecular Sciences, University of Amsterdam, Amsterdam, the Netherlands. [4]Dipartimento di Chimica, Università degli Studi di Milano, Milan, Italy. [5]Zuyd University of Applied Sciences, Heerlen, the Netherlands. [6]Netherlands Organisation for Applied Scientific Research (TNO), Eindhoven, the Netherlands. [7]Group Innovation and Sustainability, Testing, Analytics and Physics Group, Covestro B.V, Waalwijk, the Netherlands. [8]These authors contributed equally: Simone Pilon, Elia Savino, Oliver M. Bayley. ✉e-mail: t.noel@uva.nl

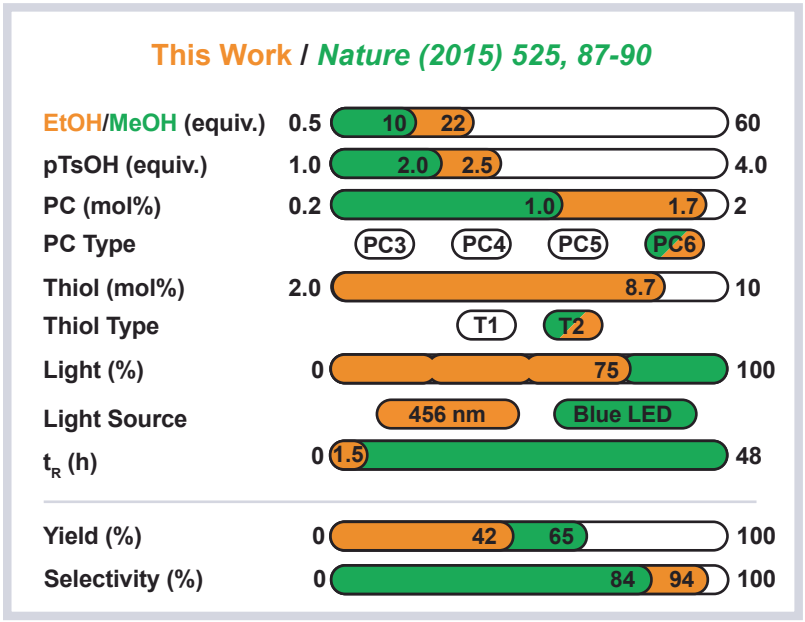
https://doi.org/10.1038/s44160-026-01053-0



Extended Data Fig. 1 | Comparison of the conditions for the trifluoroalkylation presented in Case Study 1. Comparison of the conditions for the trifluoroalkylation of pyrroles presented in Case Study 1 compared to the orginial report by Stevenson and co-workers in Chem (2016), **1**, 456; **PC1**: 4CzIPN, **PC2**: $Ru(bpy)_3Cl$, **PC3**: $Ru(bpy)_3PF_6$, **PC4**: $Ir(bpy)_3$, **PC5**: $[Ir[dF(CF_3)ppy]_2(dtbpy)]PF_6$; **Py-Ox1**: pyridine N-oxide, **Py-Ox2**: 4-phenylpyridine N-oxide. Bpy, 2,2′-bipiridine; ppy, 2-phenylpyridine; dtbpy, 4,4-di-*tert*-butyl-2,2-dipyridyl.

https://doi.org/10.1038/s44160-026-01053-0



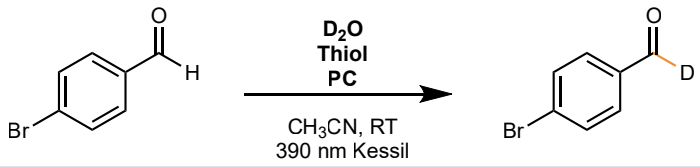


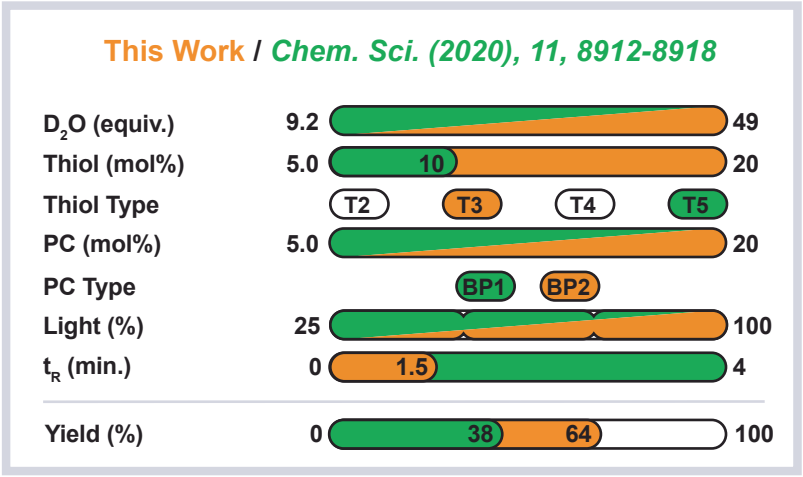
PC3: $Ru(ppy)_3(PF_6)_2$
PC4: $Ir(ppy)_3$
PC5: $Ir[dF(CF_3)ppy]_2(dtbbpy)PF_6$
PC6: $Ir(ppy)_2(dtbbpy)PF_6$

T1: methyl 2-mercaptoacetate
T2: ethyl 2-mercaptopropionate

Extended Data Fig. 2 | Comparison of the conditions for the pyridine alkylation presented in Case Study 2. Comparison of the conditions for the pyridine alkylation presented in Case Study 2 compared to the orginial report by MacMillan and co-workers in Nature (2015), **525**, 87; **PC3**: $Ru(bpy)_3PF_6$, **PC4:** $Ir(bpy)_3$, **PC5**: $[Ir[dF(CF_3)ppy]_2(dtbpy)]]PF_6$, **PC6**:$[Ir(dtbpy)(ppy)_2]PF_6$; **T1**: methyl thioglycolate, **T2**: ethyl 2-mercaptopropanoate. Bpy, 2,2′-bipiridine; ppy, 2-phenylpyridine; dtbpy, 4,4-di-*tert*-butyl-2,2-dipyridyl.

https://doi.org/10.1038/s44160-026-01053-0



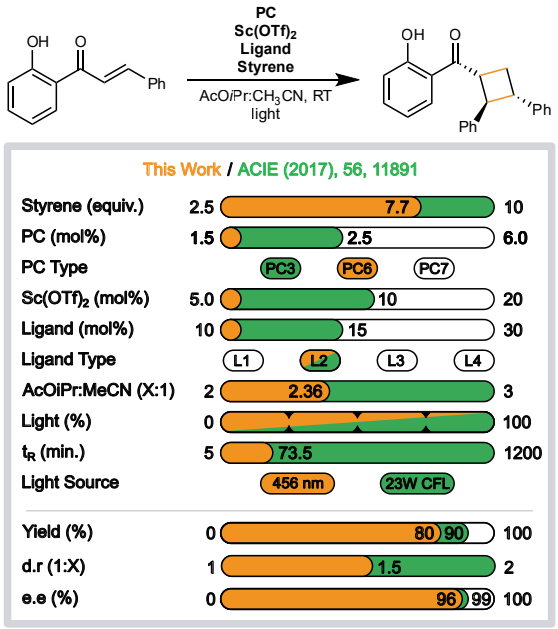


T2: ethyl 2-mercaptopropionate
T3: Ethyl Thioglycolate
T4: Thiophenol
T5: 2,4,6-triisopropylbenzenethiol

BP1: Benzophenone
BP2: 4,4'-Difluorobenzophenone

Extended Data Fig. 3 | Comparison of the conditions for the photocatalytic deuteration presented in Case Study 3. Comparison of the conditions for the photocatalytic deuteration presented in Case Study 3 compared to the original report by Wu and co-workers in Chem Sci (2020), 11, 8912; **T2**: ethyl 2-mercaptopropanoate, **T3**: ethyl thioglycolate, **T4**: thiophenol, **T5**: 2,4,6-triisopropylbenzenethiol. **BP1**: benzophenone, **BP2**: 4,4-difluorobenzophenone.

https://doi.org/10.1038/s44160-026-01053-0



Extended Data Fig. 4 | Comparison of the conditions for the enantioselective photoaddition presented in Case Study 6. Comparison of the conditions for the enantioselective photoaddition presented in Case Study 6 compared to the original report by Yoon and co-workers in ACIE (2017), **56**, 11891; **PC3**: $Ru(bpy)_3PF_6$, **PC6**:$[Ir(dtbpy)(ppy)_2]PF_6$, **PC7**: $[Ir(dFFppy)_2(dtbpy)]PF_6$; **L1**: ^{t}Bu-BOX, **L2**: ^{t}Bu-Py-BOX, **L3**: Bn-Py-BOX, **L4**: ^{i}Pr-Py-BOX. Bpy, 2,2′-bipiridine; ppy, 2-phenylpyridine; dtbpy, 4,4-di-*tert*-butyl-2,2-dipyridyl; ^{t}Bu, *tert*-butyl; Py, pyridyl; Bn, benzyl; ^{i}Pr, *iso*-propyl; BOX, bis(oxazoline); Ac, acetate; CFL, compact fluorescent lamp.