Data-science driven autonomous process optimization

Melodie Christensen[1,2], Lars P.E. Yunker[1], Folarin Adedeji[2], Florian Häse[3,4,5,7,9], Loïc M. Roch[3,4,5,9], Tobias Gensch[6], Gabriel dos Passos Gomes[4,5,7], Tara Zepel[1], Matthew S. Sigman*[6], Alán Aspuru-Guzik*[3,4,5,7,8] and Jason E. Hein[1*,9]

[1]Department of Chemistry, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada.
[2]Department of Process Research and Development, Merck & Co., Inc., Rahway, NJ 07065, United States.
[3]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, United States.
[4]Department of Chemistry, University of Toronto, Toronto, Ontario M5S 3H6, Canada.
[5]Department of Computer Science, University of Toronto, Toronto, Ontario M5T 3A1, Canada.
[6]Department of Chemistry, University of Utah, Salt Lake City, Utah 84112, United States.
[7]Vector Institute for Artificial Intelligence, Toronto, Ontario M5S 1M1, Canada.
[8]Canadian Institute for Advanced Research, Toronto, Ontario M5G 1Z8, Canada.
[9]ChemOS Sàrl, Lausanne, Vaud 1006, Switzerland.

*Corresponding authors: Matthew S. Sigman sigman@chem.utah.edu, Alán Aspuru-Guzik aspuru@utoronto.ca, Jason E. Hein jhein@chem.ubc.ca

Abstract

Autonomous process optimization involves the human intervention-free exploration of a range of pre-defined process parameters in order to improve responses such as reaction yield and product selectivity. Utilizing off-the-shelf components, we developed a closed-loop system capable of carrying out parallel autonomous process optimization experiments in batch with significantly reduced cycle times. Upon implementation of our system in the autonomous optimization of a palladium-catalyzed stereoselective Suzuki-Miyaura coupling, we found that the definition of a set of meaningful, broad, and unbiased process parameters was the most critical aspect of a successful optimization. In addition, we found that categorical parameters such as phosphine ligand were vital to determining the reaction outcome. To date, categorical parameter selection has relied on chemical intuition, potentially introducing an element of bias into the experimental design. In seeking a systematic method for the selection of a diverse set of phosphine ligands fully representative of the chemical space, we developed a strategy that leveraged computed molecular descriptor clustering analysis. This strategy allowed for the successful autonomous optimization of a stereoselective Suzuki-Miyaura coupling between a vinyl sulfonate and an arylboronic acid to selectively generate the *E*-product isomer in high yield.

Introduction

Recent advancements in computer science and automation technologies have led to the emergence of autonomous chemistry systems built on the premise of human intervention-free hypothesis generation and testing.[1] Autonomous workflows typically involve three key components: (1) a machine learning (ML) algorithm for hypothesis generation, (2) a robotic system for experimental execution, and (3) an online analytics platform for performance evaluation. The level of human intervention varies based on the degree of hardware and software integration among the key components. In a fully autonomous system, a "closed-loop" is achieved, where the scientist can define the search space, and then hit "go". Closed-loop applications have ranged from biologically active compound discovery[2] to materials development,[3] novel reaction scouting[4] and process optimization.[5]

Autonomous process optimization involves the human intervention-free exploration of a range of pre-defined process parameters to improve responses such as reaction yield, product selectivity and catalyst turnover number. The definition of a set of meaningful, broad, and unbiased process parameters is arguably the most critical aspect of a successful optimization. Work to date has focused on the multivariate optimization of continuous parameters such as temperature, stoichiometry and time; however, vital categorical parameters such as reagent, solvent or catalyst have rarely been incorporated. In fact, leading examples involving continuous and categorical parameter combinations have been limited to fewer than eight catalysts or ten solvents.[6] Furthermore, in these examples, categorical parameter selection was guided through chemical intuition, potentially introducing an element of bias into the experimental design. Thus, in our view, categorical parameter selection in the context of autonomous process optimization remains an unsolved challenge. We envisioned developing a more systematic method for the selection of a broad and diverse set of categorical parameters to fully represent the chemical space, driving more effective optimization campaigns.

Successful optimization also hinges on the identification of suitable automation equipment capable of effective experimental execution and analysis. Current research focuses heavily on custom-built continuous and segmented flow-reactor systems outfitted with online analytics for experimental execution. While these state-of-the-art systems have enabled the rapid multivariate optimization of several processes, examples have still been limited to the sequential execution of fast, homogeneous reactions amenable to flow reactors.[7] In this work, we expand the autonomous process optimization toolkit to include a broader set of reaction methodologies by integrating off-the-shelf robotic systems with online analytics to carry out parallel reaction loops in 96-well plates.

The final aspect of successful optimization is the selection of an effective ML algorithm. Recently, Bayesian optimization algorithms have gained traction in the in autonomous chemistry realm, leading to a number of successful optimization campaigns.[8] The advantage of Bayesian optimization is the implicit formulation of expected outcomes for all possible reaction conditions which are conditioned on all previous observations and constantly refined as soon as new observations become available. One example is the Phoenics[9] algorithm, a Bayesian optimizer developed for the global optimization of chemistry problems, with the capability to propose multiple parameter points in parallel. Another example is the Gryffin[10] algorithm, developed specifically for the identification of optimal choices for categorical parameters. Given our interest in the autonomous optimization of categorical and continuous parameters in tandem through the execution of parallel reaction loops, we combined the Phoenics and Gryffin optimization strategies to satisfy both requirements.

We identified a stereoselective Suzuki-Miyaura cross-coupling reaction that would benefit from a tandem categorical and continuous parameter optimization (Figure 1).[11] Typically, Suzuki-Miyaura cross-couplings of vinyl halides or sulfonates proceed with retention of the olefin bond geometry, but in this example, vinyl sulfonate **1-E** undergoes stereoinversion under ligand-free and electron-rich dialkylbiaryl phosphine-mediated palladium catalysis to generate product **2-Z**. In contrast, stereoinversion is partially suppressed using ferrocenyl bisphosphine-mediated palladium catalysis, facilitating the selective formation of product **2-E**. Thus, stereoselectivity in this system appears to be influenced by both the phosphine ligand selection and stoichiometry.[12] Importantly, traditional phosphines preferred in Suzuki couplings, such as dialkylbiaryl phosphine ligands,[13] appear to facilitate the undesired steroinversion pathway. Finally, with reaction completion times in the order of two hours, this system is not amenable to flow reactors due to impractically long residence times.
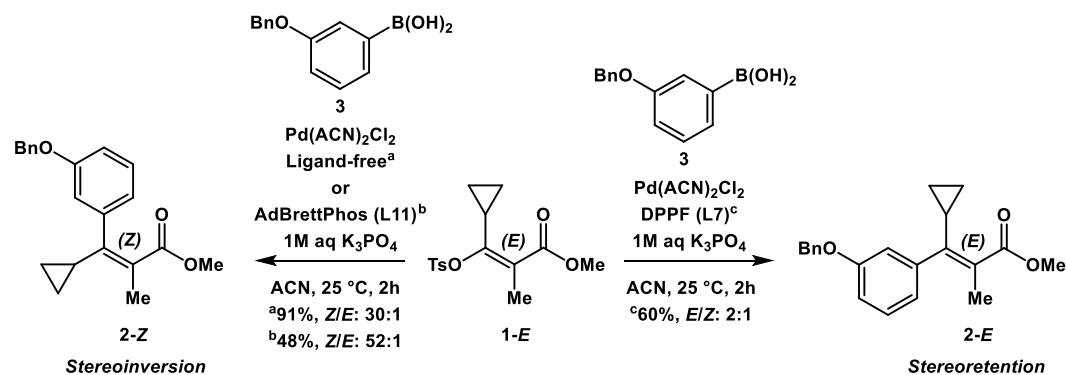
Figure 1. Phosphine ligand and reaction condition influence on a palladium-catalyzed stereoselective Suzuki-Miyaura coupling to generate the stereoinversion (2-*Z*) or stereoretention product (2-*E*).
[a] Conditions: 10 µmol **1-*E***, 1 µmol 1,3,5-trimethoxybenzene, 20 µmol **3**, 0.4 µmol Pd(ACN)$_2$Cl$_2$, 30 µmol K$_3$PO$_4$ (0.5M aq) in ACN (0.05M), 2 h at 25 °C.
[b,c] Conditions: 10 µmol **1-*E***, 1 µmol 1,3,5-trimethoxybenzene, 11 µmol **3**, 0.2 µmol Pd(ACN)$_2$Cl$_2$, 0.4 µmol **L**, 30 µmol K$_3$PO$_4$ (0.5M aq) in ACN (0.05M), 2 h at 25 °C. Ligand structures are provided in Figure 3.

Our goal was to improve the yield of stereoretention product ***2-E*** through an autonomous optimization campaign by exploring a selection of phosphines and continuous process parameters in tandem. We employed a Chemspeed SWING robotic system for the experimental execution of parallel reaction loops in batch and employed the Phoenics and Gryffin algorithms for the proposal of parallel combinations continuous and categorical process parameter selections. Recognizing the impact of phosphine selection on the optimization outcome, we employed a variety of categorical parameter selection strategies, including chemical intuition and computed molecular descriptor clustering of 365 commercially available phosphines.[14] Here, we discuss the advantages and limitations of each phosphine selection strategy and their impacts on this challenging optimization problem.

Results

Establishing a closed-loop system

The establishment of a closed-loop system required the integration of three main components: (1) ChemOS,[15] the experimental scheduler for coordination of experiments proposed by the ML algorithms (Phoenics and Gryffin), (2) Chemspeed SWING, the robotic system for automated experimental set-up, and (3) Agilent 1100, the HPLC-UV system for measurement of the experimental outcomes (Figure 2). The only hardware customization required was the integration of the Agilent HPLC-UV system with the Chemspeed SWING robotic platform. This integration was accomplished through the installation of an HPLC valve on the Chemspeed robot deck and incorporation of relay switches for triggering chromatographic resolution and photodiode array detection.[16] The next step was to establish automated data flow from the experimental scheduler to the robot, and from the online analytical system back to the experimental scheduler for the ML algorithm to interpret results and propose subsequent experiments. In lieu of developing a complex Application Programming Interface (API) among the three software components (ChemOS, Chemspeed Autosuite, and Agilent Chemstation), we opted to develop a lightweight Python framework for data transfer between these components. The script translated ChemOS parameter suggestions into stock mixture dispense volumes, calculated product assay yields from HPLC peak area ratios to an internal standard and reported experimental measurements back to ChemOS.
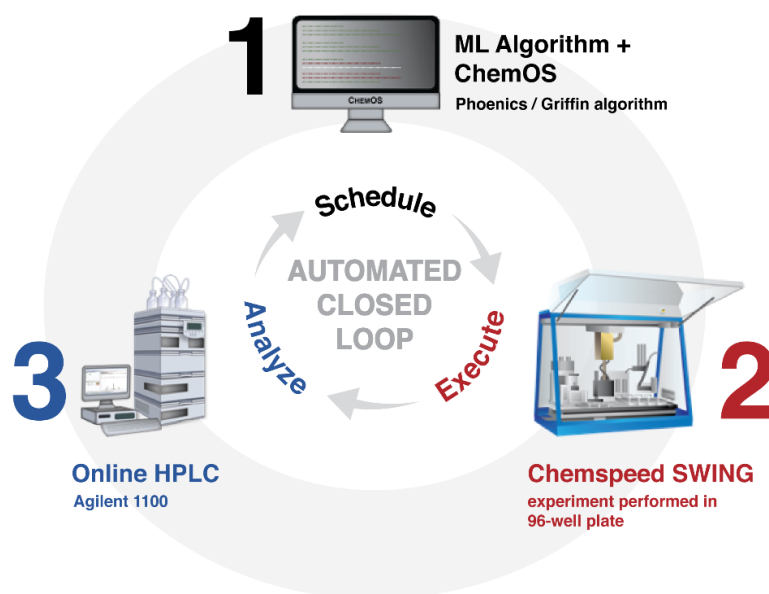
Figure 2. Closed-loop system for autonomous optimization in batch. The three main components to enable this closed-loop include (1) ChemOS to coordinate experiments and data-driven approaches, (2) Chemspeed SWING for automated experimental setup, and (3) Agilent 1100 to characterize the experimental outcomes.

Defining the process parameters and optimization objectives

A set of categorical and continuous parameters was selected for potential impact on the reaction outcome, including phosphine ligand, phosphine to palladium ratio, palladium loading, arylboronic acid equivalents, and reaction temperature (Table 1). Upper and lower bound selection for each continuous parameter was guided by chemical intuition but also designed to be sufficiently broad to explore a diverse set of responses. Ligand set selections varied from 12 up to 23 ligands, depending on the selection strategy (see below). The ML algorithms (Phoenics and Gryffin) were configured to maximize the yield of the *E*-product, minimize the yield of the *Z*-product, minimize the palladium loading, and minimize the arylboronic acid equivalents, in that order (Table 1). Each objective was configured with a 10 % relative threshold that would only consider the next objective once that threshold had been achieved. This multi-objective optimization, or Pareto optimization design, was made possible through the implementation of the scalarizing function Chimera.[17]

Table 1. Process parameters and optimization objectives

Parameter	Type	Range	Unit
Phosphine ligand (P ligand)	Categorical	12 – 23	Number of ligands
Phosphine to palladium ratio (P/Pd)	Continuous	0.5 – 4.0	ratio
Palladium loading (Pd mol%)	Continuous	1.0 – 5.0	mol %
Arylboronic acid equivalents (ArBA equiv)*	Continuous	1.0 – 2.0	equivalents
Reaction temperature (Rxn temp)	Continuous	10 – 40	°C
Response	**Objective**	**Priority**	**Unit**
E-product assay yield (*E*-PR AY)	Maximize	First	mol %
Z-product assay yield (*Z*-PR AY)	Minimize	Second	mol %
Palladium loading (Pd mol%)	Minimize	Third	mol %

4

Arylboronic acid equivalents (ArBA equiv)*   Minimize   Fourth   equivalents

*Note: The arylboronic acid equivalents parameter and response was removed from the experimental design upon expanding the search space from 12 to 23 ligands.

Designing the automation workflow

Given the broad search space, a reasonable number of reactions for an autonomous optimization campaign was determined to reside between 120 and 192, depending on the number of ligands under evaluation. Full conversion in this Suzuki-Miyaura coupling is typically achieved within two hours under the median values of the continuous parameter ranges defined above. Thus, if the maximum number of 192 consecutive reactions were carried out sequentially, the optimization campaign would take 16 days to complete. This lengthy duration would serve to obviate the benefits of autonomous optimization, and correspondingly, necessitated parallelization of the reactions. Therefore, the reactions were parallelized into loops of eight, allowing for a 192-reaction campaign to be completed within 24 loops over a more reasonable time period of four days. Conditions for the first loop of eight reactions were selected randomly by the ML algorithm. Subsequent conditions were determined autonomously by Phoenics and Gryffin as analytical results were returned, following a data-driven strategy. ChemOS, the experimental scheduler, parallelized the suggested reactions for the Chemspeed SWING system to dispense and initiate in 15-minute intervals. Analytical samples were also aliquoted and acquired in 15-minute intervals, at the endpoint of each reaction.

Trial runs and experimentally-derived constraints for the optimization strategy

Trial runs unveiled two instrument constraints that necessitated further enhancements of the optimization strategy. The first constraint involved the need to fix the reaction temperature across each loop of eight reactions given that these reactions were designed to be carried out within the same reactor block and timeframe. In order to accommodate this constraint, the capabilities of the Phoenics and Gryffin optimization strategies were extended to facilitate optimization with process constraints following the idea of a previously introduced basic process-constrained Bayesian optimization algorithm (pc-BO(basic)).[18] This extension allowed the suggestion of a total of eight different experiments where the temperature was fixed across one loop. The second instrument constraint was the inability of the Chemspeed robot to execute sub-microliter dispenses accurately; thus, the Python script was augmented to round calculated dispense volumes to the microliter level and update the suggested parameters with the executed parameters prior to returning analytical results to ChemOS. These enhancements allowed for the successful application of an algorithm in a real-world experimental setting.

Autonomous optimization with ligands selected through chemical intuition

Initial ligand selection was carried out through chemical intuition around phosphines with the potential to accelerate palladium-catalyzed cross-couplings[19] (here, chemical intuition refers to insight arising from a combination of literature precedent and hands-on experience). Twelve ligands were selected, including trialkyl, triaryl, ferrocenyl and dialkylbiaryl phosphines. These ligands were evaluated in combination with four continuous parameters, including phosphine ligand, phosphine to palladium ratio, palladium loading, arylboronic acid equivalents, and reaction temperature. The optimization campaign was carried out in fifteen loops of eight reactions, totaling 120 autonomous iterations carried out over 60 hours. This first campaign resulted in conditions to access product **2-E** in 65% yield and 2:1 *E/Z* selectivity upon 118 iterations, under Pd-**L7** catalysis (Figure 3). It was also noted that the optimizer dedicated a higher number of iterations to **L7** (DPPF), the top-performing ligand. Although this optimization should be viewed as

successful, improvements to product **2-E** yields from our initial conditions were nominal (compared to 60% yield and 2:1 *E/Z* selectivity as shown in Figure 1). This nominal improvement was attributed to ligand bias resulting from chemical intuition-based selection, thus, we embarked upon the exploration of a broader set of ligands to access **2-E** in higher yield and selectivity.
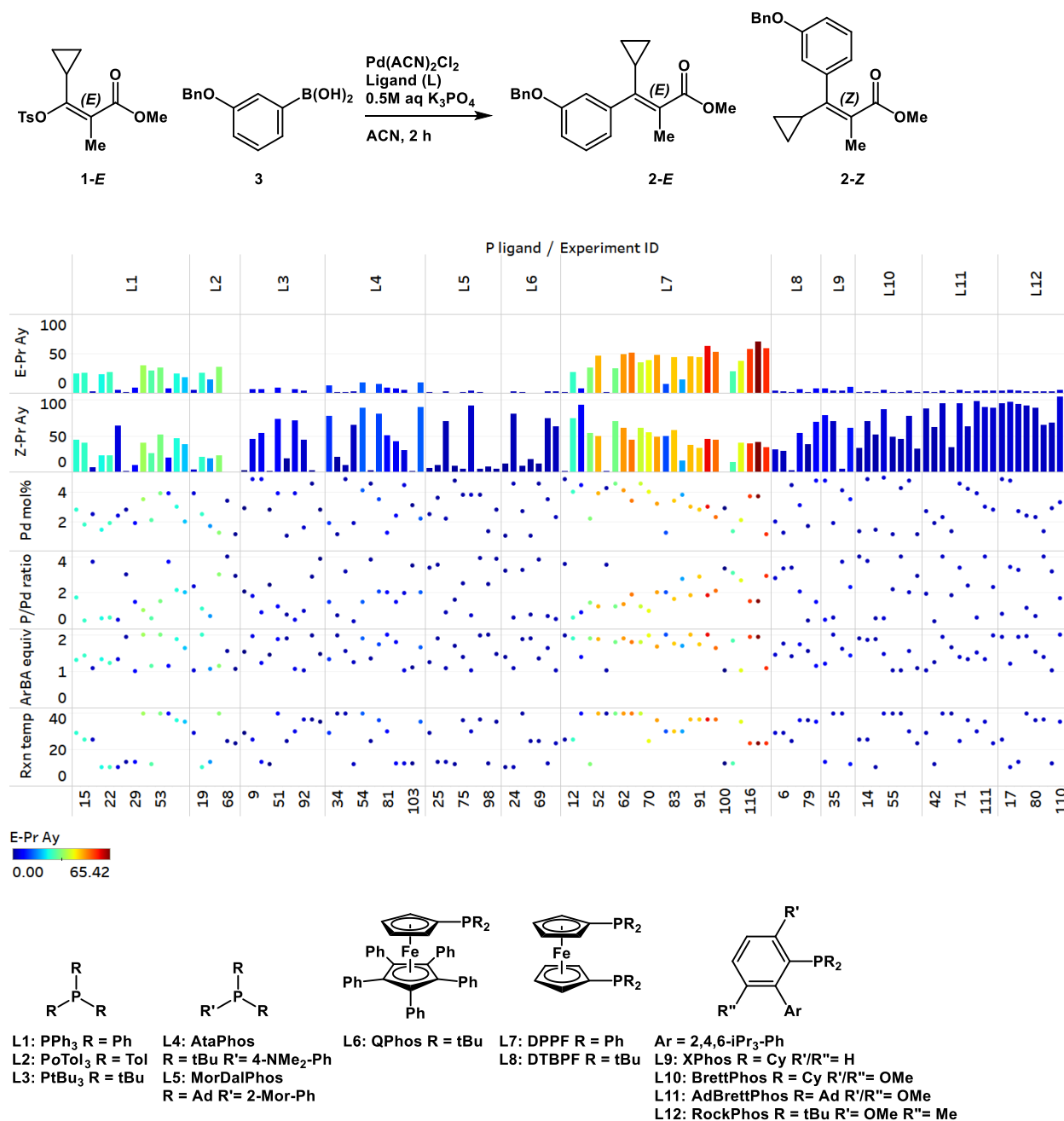

Figure 3. Parameters and results of optimization with ligands selected through chemical intuition (campaign 1).
Conditions: 10 μmol **1-E**, 1 μmol 1,3,5-trimethoxybenzene, 10 - 20 μmol **3**, 0.1 - 0.5 μmol Pd(ACN)$_2$Cl$_2$, 0.05 - 2 μmol **L**, 30 μmol K$_3$PO$_4$ (0.5M aq) in ACN (0.05M), 2 h at 10 - 40 °C.

Autonomous optimization with ligands selected through computed molecular features

To this end, we sought a systematic method for the selection of a diverse set of phosphines for autonomous evaluation. A particularly attractive approach would leverage computed molecular features of phosphines as these have been applied to reaction optimization through predictive modeling.[20] In our current study, 365 commercially available monodentate phosphines were used to define the chemical space (the focus was limited to monodentate phosphines in order to more effectivity control the ligation state of palladium). For each phosphine, features were obtained by computing molecular properties for a representative set of conformers using DFT. Then, k-means clustering was carried out on the first four principal components of this descriptor set to divide the chemical space into 24 regions. A single compound was selected from each cluster for experimental evaluation based on additional considerations such as availability in the Merck phosphine library, price, and anticipated stability (Figure 4).[21] One cluster contained ligands that were deemed too challenging to implement; therefore, candidates from this cluster were not included in the experimental design.[22]
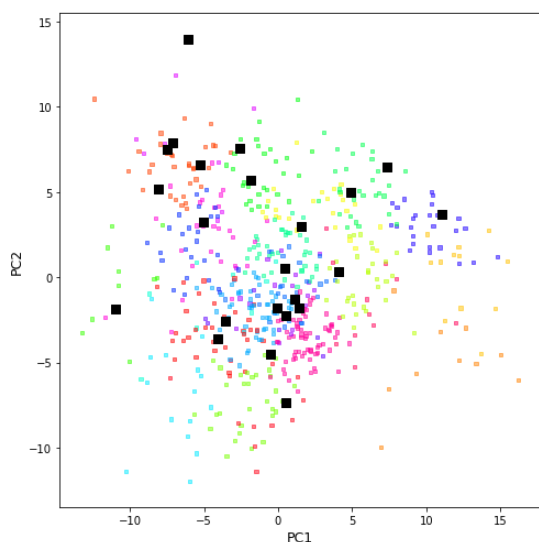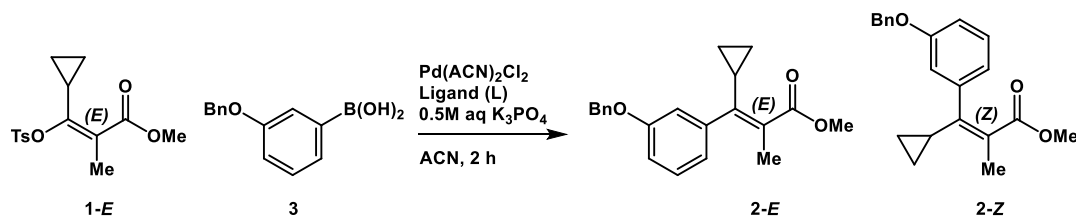

Figure 4. K-means clustering on the first four principal components of the molecular descriptor set for 365 commercial monodentate phosphines. The chemical space is represented by a two-dimensional plot of the first two principal components.
Note: Each cluster is represented by color and highlighted boxes indicate selected candidates.
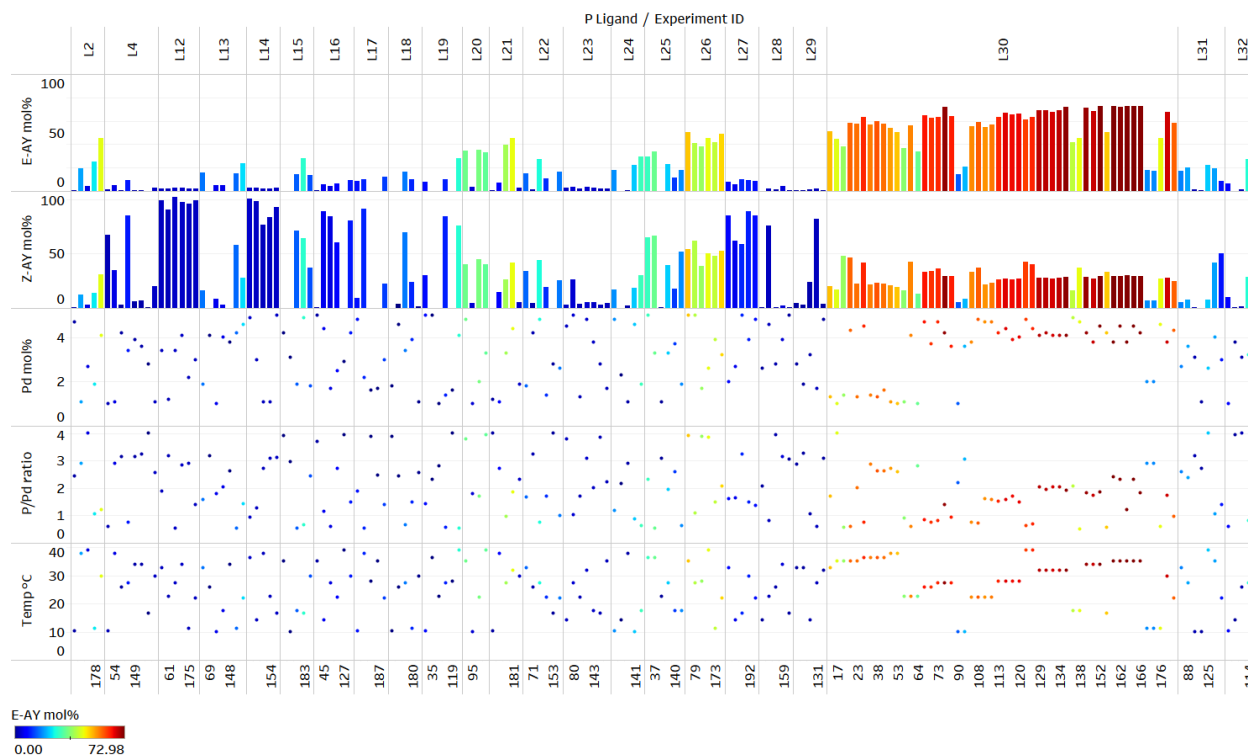
Investigation of a larger ligand set necessitated the expansion of the optimization campaign to 24 loops of eight reactions, totaling 192 autonomous iterations executed over 96 hours (16 of these reactions were designated as test reactions to test for reproducibility, accuracy and precision). In addition, the larger ligand set also drove the removal of the arylboronic acid equivalents parameter and response from the experimental design. This second campaign resulted in the identification of optimal conditions to access product **2-*E*** in 73% yield and 3:1 *E/Z* selectivity upon 161 iterations, using **L30** as the ligand (Figure 5a). Given that the ML algorithm had no information to bias the search, and that ligand selection was unbiased, the improved stereoselectivity of 3:1 highlights the potential of our novel optimization technology. As in the first campaign, the algorithm dedicated a significant portion of iterations to the top-performing ligand; previously **L7** (DPPF), now **L30** (PhSPhos). Additional high-performing ligands also fell under the triaryl phosphine category, with both electron-rich (**L20**, **L21**) and electron-poor (**L26**) triaryl phosphines proving effective. Surprisingly, ligands with significant structural similarity to **L30** (PhSPhos), including **L28** and **L31**
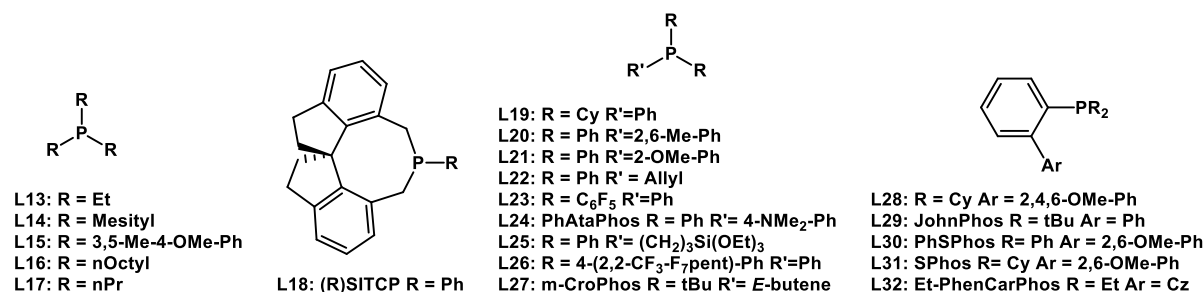
(SPhos), did not selectively yield **2-E**, presumably due to the presence of two electron-rich cyclohexyl substituents.

A deeper look into the influence of the continuous parameters on the yield of product **2-E** (Figure 5b) revealed that phosphine to palladium ratios within the center of the studied range provided optimal outcomes. In fact, phosphine to palladium ratios near the upper and lower bound values resulted in significant declines in yields. Conversely, performing the reaction at temperatures near the upper bound of the evaluated range proved most effective in driving product formation. Finally, and unsurprisingly, higher palladium loadings resulted in improved product **2-E** yields; however, the algorithm appeared to explore a bimodal distribution of palladium loading, focusing on regions close to 1.5 mol% and 3.5 mol%, landing at an optimum of 3.7 mol%. Although we are currently unable to interpret this bimodal exploration behavior, we are pleased to observe that the algorithm does not default to maximizing to upper bound values when presented with potentially linear trends.
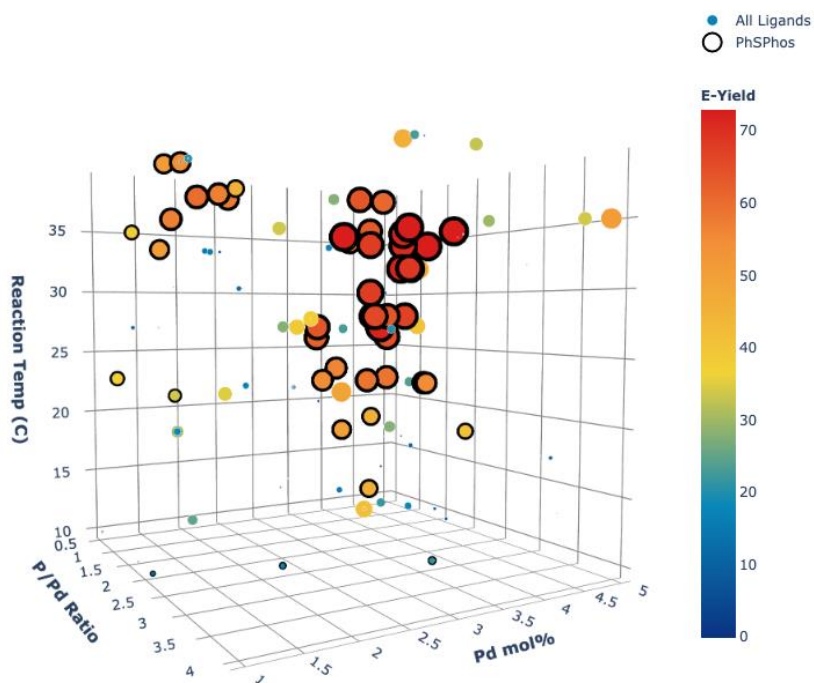

a.

8

L13: R = Et
L14: R = Mesityl
L15: R = 3,5-Me-4-OMe-Ph
L16: R = nOctyl
L17: R = nPr

L18: (R)SITCP R = Ph

L19: R = Cy R'=Ph
L20: R = Ph R'=2,6-Me-Ph
L21: R = Ph R'=2-OMe-Ph
L22: R = Ph R' = Allyl
L23: R = C_6F_5 R'=Ph
L24: PhAtaPhos R = Ph R'= 4-NMe_2-Ph
L25: R = Ph R'= $(CH_2)_3Si(OEt)_3$
L26: R = 4-(2,2-CF_3-F_7pent)-Ph R'=Ph
L27: m-CroPhos R = tBu R'= E-butene

L28: R = Cy Ar = 2,4,6-OMe-Ph
L29: JohnPhos R = tBu Ar = Ph
L30: PhSPhos R= Ph Ar = 2,6-OMe-Ph
L31: SPhos R= Cy Ar = 2,6-OMe-Ph
L32: Et-PhenCarPhos R = Et Ar = Cz

b.

Figure 5a. Parameters and results of optimization with ligands selected through descriptor clustering (campaign 2). 5b. 3D visualization of continuous parameters vs. *E*-product yield (campaign 2). PhSPhos results are outlined; all other ligand results are not outlined.
Conditions: 10 μmol **1-*E***, 1 μmol 1,3,5-trimethoxybenzene, 15 μmol **3**, 0.1 - 0.5 μmol Pd(ACN)$_2$Cl$_2$, 0.05 - 2 μmol **L**, 30 μmol K$_3$PO$_4$ (0.5M aq) in ACN (0.05M), 2 h at 10 - 40 °C.

Comparison of the optimization results from the two campaigns superimposed on the monodentate phosphine ligand space clearly demonstrates the advantage of systematic ligand selection over chemical intuition-based ligand selection (Figure 6). A wider range of product **2-*E*** yields were observed through the systematic exploration of a diverse set of ligands, ultimately leading to the discovery of **L30** as a superior ligand. This is likely due to the challenging nature of the stereoselective Suzuki-Miyaura coupling under evaluation, where selectivity appears to be driven by ligand selection through a poorly understood mechanism, driving the need for empirical ligand exploration.
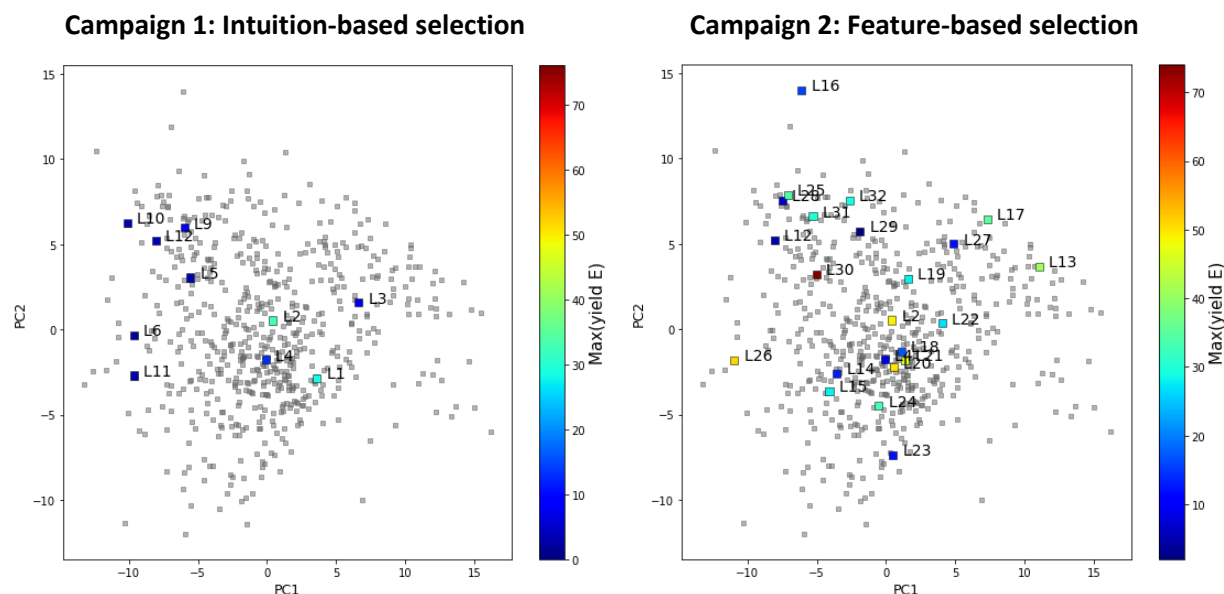
9

Figure 6. The maximum yield of 2-*E* obtained for each monodentate ligand explored in campaigns 1 and 2 as a function of the monodentate phosphine ligand space.

At this stage, the only question that remained unanswered was whether the algorithm's optimization performance could be improved if ligands were parametrized within the algorithm instead of being treated as black-box parameters. We hypothesized that the utilization of computed molecular descriptors as a means for the algorithm to relate among ligands could accelerate convergence. We therefore manually selected 15 from of the hundreds of descriptors utilized in ligand selection and carried out a third 192-iteration optimization campaign.[23] We found that a similar optimum was reached, accessing product **2-*E*** in 74% yield and 3:1 *E/Z* selectivity upon iteration 159, again under Pd-**L30** catalysis. Surprisingly, the algorithm did not appear to converge as clearly as in the previous campaign because fewer iterations focused on the top ligand (**L30**) and, in fact, a high number of iterations focused on unproductive ligands.[24] This observation may be attributed to some form of unproductive bias introduced by the selected set of descriptors, solidifying that an unbiased ligand section was critical to the success of this optimization.

Follow-up experiments with *E*-selective ligands identified through predictive modeling

With two 192-iteration data sets in hand, various modeling strategies were employed to predict additional phosphines that could also promote *E*-selectivity.[25] We found that a reasonable multivariate linear regression model (MLR) was generated, and multiple ligands with the potential for *E*-selectivity were proposed.[26] Manual experiments employing six of the predicted ligands were carried out utilizing the optimal conditions identified from autonomous optimization to measure the selectivity outcome. Additionally, **L30** was employed in a control experiment. Upon experimental execution, we were pleased to find that experimental results agreed with the predictions (Figure 7). Although Pd-**L30** still offered the highest olefin **2-*E*** yield and E/Z ratio, two ligands (**L34**, **L36**) out of six were identified that surpassed the 50% threshold for product **2-*E*** yield, meaning that one-third of the ligands proposed offered *E*-selectivity. This result is a significant improvement over the ligand set identified in the initial clustering analysis, where employment of only one-twelfth of the ligands surpassed the 50% yield threshold for generating

product **2-E**. It is important to note that the newly identified *E*-selective ligands are structurally distinct from **L30** and, thus, it would have been unlikely to arrive at these selections through chemical intuition.
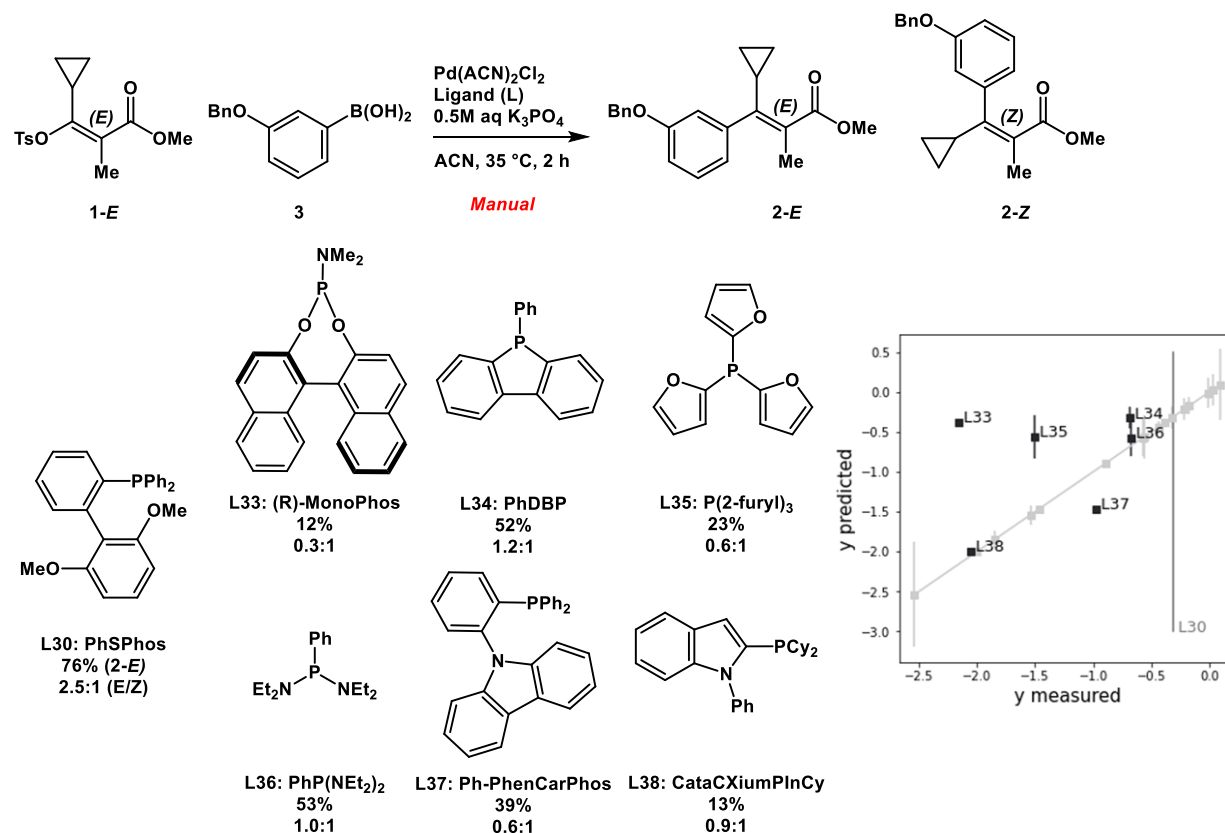

Figure 7. Manual experiments under optimized conditions with predicted *E*-selective ligands.
Note: y is the ln of the highest yield of **2-E** per ligand in any experiment.
Conditions: 10 µmol **1-E**, 1 µmol 1,3,5-trimethoxybenzene, 15 µmol **3**, 0.4 µmol Pd(ACN)$_2$Cl$_2$, 0.9 µmol **L**, 30 µmol K$_3$PO$_4$ (0.5M aq) in ACN (0.05M), 2 h at 35 °C.

Discussion

We have demonstrated the human-intervention free multivariate optimization of a stereoselective Suzuki-Miyaura coupling in batch through the autonomous evaluation of a large phosphine ligand set and continuous parameters in tandem. This success was accomplished through a series of technological advances. First, the implementation of a robotic system capable of carrying out parallel experiments with online analytics significantly reduced cycle times. Second, the development of seamless communication between the ML algorithm and robotic hardware established a closed-loop system. Third, the incorporation of Bayesian ML algorithms, here Phoenics and Gryffin, facilitated tandem categorical and continuous parameter optimization. Finally, the employment of computed molecular features enabled the systematic and unbiased definition of the phosphine ligand search space. The application of these advances resulted in the rapid identification of optimal conditions and ligand clusters to maximize the yield of product **2-E**.

It is relevant to ask what advantage autonomous optimization offers over more well-established technologies such as HTE (high-throughput experimentation) or DOE (design of experiments). Although we view algorithmic optimization as a complementary technology, we also believe its advantage lies in the multivariate optimization of categorical and continuous parameters in tandem within the fewest

11

number of iterations. To elaborate on this point, a theoretical comparison among the technologies in the context of this Suzuki-Miyaura optimization is warranted. The goal of this optimization was to evaluate 23 categorical and three continuous parameters in tandem to maximize the yield of product **2-E**. In our studies, we observed that the yield of product **2-E** followed a bell-shaped distribution with respect to one of the continuous parameters, phosphine to palladium ratio. In an HTE or DOE paradigm, studying the lower, center and upper bound values of each continuous parameter with each ligand would have resulted in 621 experiments, much higher than the number of experiments reported in our successful autonomous optimization (192 experiments). DOE factorial designs with fewer experiments could have been implemented, but this approach would have sacrificed center points, in this case leading to the potential to miss the global optimum of phosphine to palladium ratio. Thus, at least in the context of this optimization, the autonomous approach theoretically proved most efficient in the identification of the global optimum within the fewest number of iterations.

The autonomous optimization approach can easily be replicated to solve a multitude of multivariate process optimization problems. Once widely adopted, the technology has the potential to empower modern-day chemists to shift their focus away from routine experimental execution and towards higher-complexity problem-solving. Areas for future improvement include further extensions of algorithmic schemes to facilitate process-constrained batch optimization as well as enhancements to multivariate data analysis to drive a better understanding of reaction trends, in essence, to learn what the algorithm learned.

Methods

Instrumentation

Autonomous optimization experiments were executed using a Chemspeed SWING robotic system equipped with a four-needle dispense head and four 1 mL syringe pumps to enable accurate dispenses at low volumes. Slurry dispensing was enabled through 0.8 mm needle inner diameters. Agitation was carried out through an integrated custom V&P scientific 2-position tumble stirring module and temperature control was achieved through an integrated Huber Unistat chiller with temperature feedback control. Online HPLC analysis was carried out through an integrated Agilent 1100 HPLC equipped with a photodiode array detector and a custom sampling valve installed on the robot deck.

Automation of data flow

ChemOS was utilized as the scheduler that packaged proposed experiments from the machine learning algorithm into batches of eight for execution by the Chemspeed SWING robotic module. Communication of ChemOS with the Chemspeed SWING robot and Agilent HPLC was established through a lightweight Python tool. This script parsed experimental parameters proposed by ChemOS, calculated dispense volumes based on stock solution concentrations, and wrote those volumes to a CSV file actively being monitored by the Chemspeed Autosuite software. Upon automated experimental execution and subsequent HPLC analysis, this script also parsed Agilent HPLC report files for peak area counts to and calculated product yields based on predetermined response factors, which were then reported back to ChemOS for interpretation and proposal of new experimental parameters.

Experimental planning algorithms

The ML algorithms used for experiment planning in this study, Phoenics and Gryffin, leverage fundamental concepts from Bayesian optimization in combination with kernel density estimation. Bayesian

optimization is an approach to global optimization for applications where the evaluation of a single parameter point is highly time or resource demanding. While several formulations exist, Bayesian optimization follows a two-step strategy to suggest parameter points for future evaluation: (1) constructing a statistical approximation to the considered experiment based on collected measurements, and (2) locating parameter points for which the approximation predicts promising performance. Phoenics and Gryffin construct the statistical approximation based on kernel density estimates of evaluated parameters and suggest promising parameter points with an explicit balance of exploitative and explorative sampling behavior with native support for batch optimization.

Automated experimental procedure

General: Stock solutions or slurries were prepared manually in anhydrous ACN under N_2 atmosphere and placed on the robot deck for autonomous execution. Two fluoropolymer and PFA mat-sealed 96-well metal blocks with 1 mL glass vial inserts were equilibrated at the designated reaction temperature under 20 psig of N_2 with 500 rpm agitation.

Representative procedure for test reactions: In campaigns involving 192 iterations, eight wells from each 96-well reaction block were dedicated to standard reactions to test for reproducibility. To each well was dispensed Pd(ACN)$_2$Cl$_2$ (0.25 µmol, 25 µl of 0.01 M stock solution), and **L2** (PoTol$_3$) (0.38 µmol, 19 µl of 0.02 M stock mixture), followed by 7 min of age time. Then, *E*-tosylate **1-*E*** (10 µmol) with 1,3,5-trimethoxybenzene (1 µmol) was dispensed (20 µl of 0.5 M/ 0.05 M stock solution), followed by (3-(benzyloxy)phenyl)boronic acid **3** (15 µmol, 30 µl of 0.5 M stock solution in degassed ACN 5% H_2O), followed by anhydrous ACN (106 µl) to ensure a total organic solvent volume of 200 µl. Then, a dispense of degassed aqueous K$_3$PO$_4$ (30 µmol, 60 µl of 0.5 M stock solution) was carried out to initiate the reaction. This procedure was executed sequentially for each well within a loop of eight replicates in 15-minute intervals. Each replicate was time stamped individually, aged for 120 min, and sampled for online analysis.

Representative procedure for optimization reactions: In campaigns involving 192 iterations, 88 wells from each 96-well reaction block were dedicated to the optimization reactions. To each well was dispensed Pd(ACN)$_2$Cl$_2$ (0.1 - 0.5 µmol, 10 - 50 µl of 0.01 M stock solution), and phosphine ligand (0.05 - 2.0 µmol, 3 - 100 µl of 0.02 M stock mixture), followed by 7 min of age time. Then, E-tosylate **1-*E*** (10 µmol) with 1,3,5-trimethoxybenzene (1.0 µmol) was dispensed (20 µl of 0.5 M/ 0.05 M stock solution), followed by (3-(benzyloxy)phenyl)boronic acid **3** (15 µmol, 30 µl of 0.5 M stock solution in degassed ACN 5% H_2O), followed by anhydrous ACN (0 - 138 µl) to ensure a total organic solvent volume of 200 µl. Then, a dispense of degassed aqueous K$_3$PO$_4$ (30 µmol, 60 µl of 0.5 M stock solution) was carried out to initiate the reaction. This procedure was executed sequentially for each well within a loop of eight experiments in 15-minute intervals. Each reaction was time stamped individually, aged for 120 min, and sampled for online analysis.

Sampling and analysis: Two polypropylene 96-well collection blocks sealed with a silicone mats were manually prefilled with 800 µl of acetonitrile 10% aqueous pH 3.5 ammonium formate buffer and placed on the robot deck. Upon reaching the reaction end point at 120 min, 10 ul of reaction mixture was aliquoted and dispensed into the 800 ul quench solution in the collection block. Upon needle-mixing, 40 ul of quenched sample from the collection block was aliquoted and injected to the on-deck sampling valve outfitted with a 5 µl loop. The valve was automatically switched to transfer the sample to the Agilent 1100 HPLC for analysis.

Phosphine descriptor calculations and predictive modeling

Electronic and steric descriptor calculations for 365 commercially available phosphine ligands have been

described previously.[27] Additionally, employed predictive modeling strategies surrounding parametrized phosphines have also been described previously.[28]

Conflicts of interest
There are no conflicts of interest to declare.

Acknowledgements

[1] Jensen, K. F., Coley, C. W. & Eyke, N. S. Autonomous discovery in the chemical sciences part I: Progress. *Angew. Chem. Int. Ed.* anie.201909987 (2019); Coley, C. W., Eyke, N. S. & Jensen, K. F. Autonomous discovery in the chemical sciences part II: Outlook. *Angew. Chem. Int. Ed.* anie.201909989 (2019); Häse, F., Roch, L. M. & Aspuru-Guzik, A. Next-Generation Experimentation with Self-Driving Laboratories. *Trends Chem.* **1,** 282–291 (2019); Stein, H. S. & Gregoire, J. M. Progress and prospects for accelerating materials science with automated and autonomous workflows. *Chem. Sci.* **10,** 9640–9649 (2019).

[2] Desai, B. *et al.* Rapid discovery of a novel series of Abl kinase inhibitors by application of an integrated microfluidic synthesis and screening platform. *J. Med. Chem.* **56,** 3033–3047 (2013); Weber, L., Wallbaum, S., Broger, C. & Gubernator, K. Optimization of the Biological Activity of Combinatorial Compound Libraries by a Genetic Algorithm. *Angew. Chem. Int. Ed.* **34,** 2280–2282 (1995).

[3] Burger, B. *et al.* A mobile robotic chemist. *Nature* **583,** 237–241 (2020); MacLeod, B. P. *et al.* Self-driving laboratory for accelerated discovery of thin-film materials. *Sci. Adv.* **6,** eaaz8867 (2020); Porwol, L. *et al.* An Autonomous Chemical Robot Discovers the Rules of Inorganic Coordination Chemistry without Prior Knowledge. *Angew. Chem. Int. Ed.* anie.202000329 (2020); Gongora, A. E. *et al.* A Bayesian experimental autonomous researcher for mechanical design. *Sci. Adv.* **6,** eaaz1708 (2020); Langner, S. *et al.* Beyond Ternary OPV: High-Throughput Experimentation and Self-Driving Laboratories Optimize Multicomponent Systems. *Adv. Mater.* **32,** 1907801 (2020); Tabor, D. P. *et al.* Accelerating the discovery of materials for clean energy in the era of smart automation. *Nat. Rev. Mater.* **3,** 5–20 (2018); Nikolaev, P. *et al.* Autonomy in materials research: A case study in carbon nanotube growth. *npj Comput. Mater.* **2,** 1–6 (2016).

[4] Granda, J. M., Donina, L., Dragone, V., Long, D. L. & Cronin, L. Controlling an organic synthesis robot with machine learning to search for new reactivity. *Nature* **559,** 377–381 (2018); Sans, V. & Cronin, L. Towards dial-a-molecule by integrating continuous flow, analytics and self-optimisation. *Chem. Soc. Rev.* **45,** 2032–2043 (2016); Sans, V., Porwol, L., Dragone, V. & Cronin, L. A self optimizing synthetic organic reactor system using real-time in-line NMR spectroscopy. *Chem. Sci.* **6**, 1258–1264 (2015).

[5] Mateos, C., Nieves-Remacha, M. J. & Rincón, J. A. Automated platforms for reaction self-optimization in flow. *React. Chem. Eng.* **4,** 1536–1544 (2019); Clayton, A. D. et al. Algorithms for the self-optimisation of chemical reactions. *React. Chem. Eng.* **4**, 1545–1554 (2019); Bédard, A.-C., Adamo, A., Aorh, 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* **361**, 1220–1225 (2018); Baumgartner, L. M., Coley, C. W., Reizman, B. J., Gao, K. W. & Jensen, K. F. Optimum catalyst selection over continuous and categorical process variables with a single droplet microfluidic reaction platform. *React. Chem. Eng.* **3,** 301–311 (2018); Cortés-Borda, D. *et al.* An Autonomous Self-Optimizing Flow Reactor for the Synthesis of Natural Product Carpanone. *J. Org. Chem.* **83,** 14286–14289 (2018); 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.* **22,** 542–550 (2018); Zhou, Z., Li, X. & Zare, R. N. Optimizing Chemical Reactions with Deep Reinforcement Learning. *ACS Cent. Sci.* **3,** 1337–1344 (2017); Reizman, B. J. & Jensen, K. F. Feedback in Flow for Accelerated Reaction Development. *Acc. Chem. Res.* **49,** 1786–1796 (2016); Fitzpatrick, D. E., Battilocchio, C. & Ley, S. V. A Novel Internet-Based Reaction Monitoring, Control and Autonomous Self-Optimization Platform for Chemical Synthesis. *Org. Process Res. Dev.* **20,** 386–394 (2016); Reizman, B. J., Wang, Y.-M., Buchwald, S. L. & Jensen, K. F. Suzuki–Miyaura cross-coupling optimization enabled by automated feedback. *React. Chem. Eng.* **1,** 658–666 (2016); Cortés-Borda, D. *et al.* Optimizing the Heck–Matsuda Reaction in Flow with a Constraint-Adapted Direct Search Algorithm. *Org. Process Res. Dev.* **20,** 1979-1987 (2016); Reizman, B. J. & Jensen, K. F. Simultaneous solvent screening and reaction optimization in microliter slugs. *Chem. Commun.* **51,** 13290–13293 (2015); McMullen, J. P. & Jensen, K. F. An Automated Microfluidic System for Online Optimization in Chemical Synthesis. *Org. Process Res. Dev.* **14,** 1169–1176 (2010).

[6] Baumgartner, L. M., Coley, C. W., Reizman, B. J., Gao, K. W. & Jensen, K. F. Optimum catalyst selection over continuous and categorical process variables with a single droplet microfluidic reaction platform. *React. Chem. Eng.* **3,** 301–311 (2018); Reizman, B. J., Wang, Y.-M., Buchwald, S. L. & Jensen, K. F. Suzuki-Miyaura–Miyaura cross-coupling optimization enabled by automated feedback. *React. Chem. Eng.* **1,** 658–666 (2016); Reizman, B. J. & Jensen, K. F. Simultaneous solvent screening and reaction optimization in microliter slugs. *Chem. Commun.* **51,** 13290–13293 (2015).

[7] See Reference 5.

[8] Burger, B. *et al.* A mobile robotic chemist. *Nature* **583,** 237–241 (2020); MacLeod, B. P. *et al.* Self-driving laboratory for accelerated discovery of thin-film materials. *Sci. Adv.* **6,** eaaz8867 (2020); Gongora, A. E. *et al.* A Bayesian experimental autonomous researcher for mechanical design. *Sci. Adv.* **6,** eaaz1708 (2020); Langner, S. *et al.* Beyond Ternary OPV: High-Throughput Experimentation and Self-Driving Laboratories Optimize Multicomponent Systems. *Adv. Mater.* **32,** 1907801 (2020).

[9] Häse, F., Roch, L. M., Kreisbeck, C. & Aspuru-Guzik, A. Phoenics: A Bayesian Optimizer for Chemistry. *ACS Cent. Sci.* **4,** 1134–1145 (2018).

[10] Häse, F., Roch, L. M. & Aspuru-Guzik, A. Gryffin: An algorithm for Bayesian optimization for categorical variables informed by physical intuition with applications to chemistry. arXiv:2003.12127 (2020).

[11] Plummer, C. W. *et al.* Design and Synthesis of Novel, Selective GPR40 AgoPAMs. *ACS Med. Chem. Lett.* **8,** 221–226 (2017); Christensen, M. *et al.* Enantioselective synthesis of α-methyl-β-cyclopropyldihydrocinnamates. *J. Org. Chem.* **81,** 824-830 (2016).

[12] For additional palladium-catalyzed stereoselective cross-couplings influenced by phosphine ligands, see: Chehal, N. K., Budzelaar, P. H. M. & Hultin, P. G E - Z isomerization in Suzuki cross-couplings of haloenones: Ligand effects and evidence for a separate catalytic cycle. *Org. Biomol. Chem.* **16,** 1134–1143 (2018); Lu, G. P., Voigtritter, K. R., Cai, C. & Lipshutz, B. H. Ligand effects on the stereochemistry of Stille couplings, as manifested in reactions of Z-alkenyl halides. *Chem. Commun.* **48,** 8661–8663 (2012); Lu, G. P., Voigtritter, K. R., Cai, C. & Lipshutz, B. H. Ligand effects on the stereochemical outcome of suzuki-miyaura couplings. *J. Org. Chem.* **77,** 3700–3703 (2012); Krasovskiy, A. & Lipshutz, B. H. Ligand effects on Negishi couplings of alkenyl halides. *Org. Lett.* **13,** 3818–3821 (2011).

[13] Johansson Seechurn, C. C. C., Kitching, M. O., Colacot, T. J. & Snieckus, V. Palladium-catalyzed cross-coupling: A historical contextual perspective to the 2010 nobel prize. *Angew. Chem. Int. Ed.* **51,** 5062–5085 (2012). Martin, R. & Buchwald, S. L. Palladium-catalyzed suzuki-miyaura cross-coupling reactions employing dialkylbiaryl phosphine ligands. *Acc. Chem. Res.* **41,** 1461–1473 (2008).

[14] Gomes, *et. al.* Mapping the Property Space of Monodentate Organophosphorus Ligands for Catalysis. DOI: 10.26434/chemrxiv.12996665 (2020).

[15] Roch, L. M. *et al.* ChemOS: An orchestration software to democratize autonomous discovery. *PLoS One* **15,** e0229862 (2020).

[16] Christensen, M. *et al.* Development of an automated kinetic profiling system with online HPLC for reaction optimization. *React. Chem. Eng.* **4,** 1555–1558 (2019).

[17] Häse, F., Roch, L. M. & Aspuru-Guzik, A. Chimera: Enabling hierarchy based multi-objective optimization for self-driving laboratories. *Chem. Sci.* **9,** 7642–7655 (2018).

[18] Vellanki, P. *et al.* Process-constrained batch Bayesian Optimisation. (2017).

[19] See reference 13.

[20] Zhao, S. *et al.* Enantiodivergent Pd-catalyzed C–C bond formation enabled through ligand parameterization. *Science* **362,** 670–674 (2018); Jover, J. *et al.* Expansion of the Ligand Knowledge Base for Monodentate P-Donor Ligands (LKB-P). *Organometallics* **29,** 6245–6258 (2010); Wu, K. & Doyle, A. G. Parameterization of phosphine ligands demonstrates enhancement of nickel catalysis via remote steric effects. *Nat. Chem.* **9,** 779–784 (2017); Niemeyer, Z. L., Milo, A., Hickey, D. P. & Sigman, M. S. Parameterization of phosphine ligands reveals mechanistic pathways and predicts reaction outcomes. *Nat. Chem.* **8,** 610–617 (2016); Santiago, C. B., Guo, J.-Y. & Sigman, M. S. Predictive and mechanistic multivariate linear regression models for reaction development. *Chem. Sci.* **9,** 2398–2412 (2018); Durand, D. J. & Fey, N. Computational Ligand Descriptors for Catalyst Design. *Chem. Rev.* acs.chemrev.8b00588 (2019).

[21] See Figure 5 for ligand structures.

[22] This cluster contained three low-boiling phosphines, including $P(CF_3)_3$, $PMe(CF_3)_2$ and PMe_2CF_3. The same cluster also contained PMe_2Ac, a phosphine that was difficult to access commercially or synthetically.

[23] See supplementary information section 6.6 for the selected descriptors and computed properties.

[24] See supplementary information figure SI-9 for the results of campaign 3.

[25] Gensch, *et. al.* Design and Application of a Training Set for Monophosphine Ligands in Metal Catalysis. DOI: 10.26434/chemrxiv.13160939 (2020).

[26] See supplementary information figure SI-11 for the multivariate linear regression model.

[27] See Reference 14.

[28] See Reference 25.