ORGANIC CHEMISTRY

Closed-loop optimization of general reaction conditions for heteroaryl Suzuki-Miyaura coupling

Nicholas H. Angello[1,2]†, Vandana Rathore[1,2]†, Wiktor Beker[3], Agnieszka Wołos[3,4], Edward R. Jira[2,5], Rafał Roszak[3,4], Tony C. Wu[6,7], Charles M. Schroeder[1,2,5,8], Alán Aspuru-Guzik[6,7,9,10,11,12], Bartosz A. Grzybowski[3,4,13,14]*, Martin D. Burke[1,2,15,16,17]*

General conditions for organic reactions are important but rare, and efforts to identify them usually consider only narrow regions of chemical space. Discovering more general reaction conditions requires considering vast regions of chemical space derived from a large matrix of substrates crossed with a high-dimensional matrix of reaction conditions, rendering exhaustive experimentation impractical. Here, we report a simple closed-loop workflow that leverages data-guided matrix down-selection, uncertainty-minimizing machine learning, and robotic experimentation to discover general reaction conditions. Application to the challenging and consequential problem of heteroaryl Suzuki-Miyaura cross-coupling identified conditions that double the average yield relative to a widely used benchmark that was previously developed using traditional approaches. This study provides a practical road map for solving multidimensional chemical optimization problems with large search spaces.

The development of automated synthesis methods for peptides (*1*), nucleic acids (*2*), and polysaccharides (*3*) required discovery of highly general reaction conditions applicable to a wide range of building block combinations. In contrast, in the synthesis of small organic molecules, bespoke reaction conditions are usually developed to maximize the yield of each target molecule, minimize side products, and/or minimize the cost of the corresponding process. Reaction optimization per target is often necessary because synthetic methods are typically optimized on only one or a few pairs of substrates and then applied to a wider range of substrate combinations with the rarely fulfilled hope that the same conditions will generally lead to high yields (*4*). Even the application of machine learning to optimization protocols (*5*–*9*) does not ensure generality, which is critical for automating, accelerating, and ultimately democratizing the small molecule–making process. Identification of such general conditions is difficult because the search space—spanning all possible combinations of substrates multiplied by all possible combinations of reaction conditions—is enormous and thus impractical to navigate using standard approaches.

Heteroaryl molecular fragments are ubiquitous in many industrially relevant functional molecules, including pharmaceuticals, materials, catalysts, dyes, and natural products. In all of these spaces, synthesis remains a key bottleneck. Finding general conditions for (hetero)aryl Suzuki-Miyaura cross-coupling (SMC) is therefore an important problem. It is also a challenging and largely unsolved problem, primarily owing to variable degrees of both desired and undesired reactivities across the very large and diverse range of potential heteroaryl and aryl substrates (*10*–*12*). We recently attempted, but failed, to use machine learning (ML) to discover general reaction conditions by mining the extensive chemical literature on (hetero)aryl SMC (*13*). This is mainly because the choices of conditions reported in the literature lacked causal links to the substrates' structures, and because of a lack of published (or otherwise accessibly archived) negative results.

Here, we report a simple closed-loop workflow that can efficiently navigate vast substrate-condition space to discover general reaction conditions. The approach leverages: (i) data-guided matrix down-selection to render the vast search space tractable while retaining validity to the whole; (ii) uncertainty-minimizing ML to efficiently drive prediction optimization; and (iii) robotic experimentation to increase throughput, precision, and reproducibility of datasets recursively generated on demand (Fig. 1). We demonstrate that this workflow succeeds in identifying general reaction conditions for the (hetero)aryl SMC reaction. The optimized solution doubled the average yield compared with benchmark general conditions that had previously been developed through traditional human-guided experimentation (hereafter referred to as *JACS* 2009) (*14*) and that have since been used extensively in academic and industrial laboratories worldwide (cited in >590 papers and patent applications). This approach can thus find powerful

[1]Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [2]Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [3]Allchemy, Inc., Highland, IN, USA. [4]Institute of Organic Chemistry, Polish Academy of Sciences, Warsaw, Poland. [5]Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [6]Department of Chemistry, University of Toronto, Toronto, ON, Canada. [7]Department of Computer Science, University of Toronto, Toronto, ON, Canada. [8]Department of Materials Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [9]Vector Institute for Artificial Intelligence, Toronto, ON, Canada. [10]Canadian Institute for Advanced Research, Toronto, ON, Canada. [11]Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON, Canada. [12]Department of Materials Science and Engineering, University of Toronto, Toronto, ON, Canada. [13]Center for Soft and Living Matter, Institute for Basic Science, Ulsan, Republic of Korea. [14]Department of Chemistry, Ulsan Institute of Science and Technology, Ulsan, Republic of Korea. [15]Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [16]Cancer Center at Illinois, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [17]Carle Illinois College of Medicine, University of Illinois at Urbana-Champaign, Urbana, IL, USA.
***Corresponding author. Email: mdburke@illinois.edu (M.D.B.); nanogrzybowski@gmail.com (B.A.G.)**
†These authors contributed equally to this work.

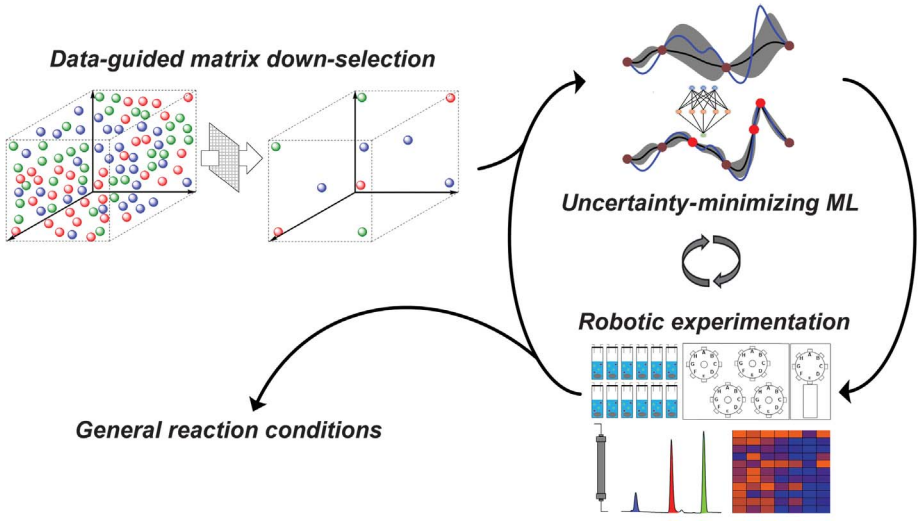

Fig. 1. Problem definition and substrate scope for generalized heterocyclic cross-coupling. Workflow developed in this work for the discovery of general reaction conditions.

solutions that lie in vast multidimensional search spaces and stands to accelerate the field of organic chemistry's march toward automated and democratized small molecule synthesis (*15*–*28*), which critically requires more general reaction conditions.

Data-guided down-selection of substrates

To enable practical pursuit of general hetero(aryl) SMC reaction conditions, we first strategically down-selected both the matrix of possible building block combinations and the matrix of possible reaction conditions in a way that preserved relevance of the subsets to their wholes (Fig. 1). Specifically, we first data mined the inventories of common fine chemical suppliers and assembled a list of ~5400 (hetero)aryl halide building blocks that were practically purchasable and therefore accessible for study [supplementary materials (SM) section 4]. To define a representative subset of this chemical space, we applied a stratified clusterization strategy (fig. S21) to algorithmically cluster the building blocks by their common (hetero)aromatic ring substructures and pendant functionalities, down-selecting 54 "centroid" molecules most representative of each section of the available chemical space. Combining these molecules with a selection of 54 commercially available (hetero)aryl *N*-methyliminodiacetic acid (MIDA) boronates defined a down-selected substrate scope composed of 2688 representative cross-coupling products (figs. S22 and S23). Mapping this potential product space and comparing it to all previously reported heteroaryl products in the literature revealed substantial overlap between both sets, suggesting that it is representative of heteroaryl chemical space as a whole (Fig. 2A). However, testing even this initially down-selected collection of cross-coupling products against many possible reaction conditions is technically unfeasible. Accordingly, we pursued a second layer of down-selection. Specifically, we used a greedy algorithm based on the Tanimoto similarity (*29*) to identify from this larger collection a set of 11 representative substrate pairs that maximize mutual dissimilarity of the resulting products (Fig. 2B). For all of these products, we determined liquid chromatography–mass

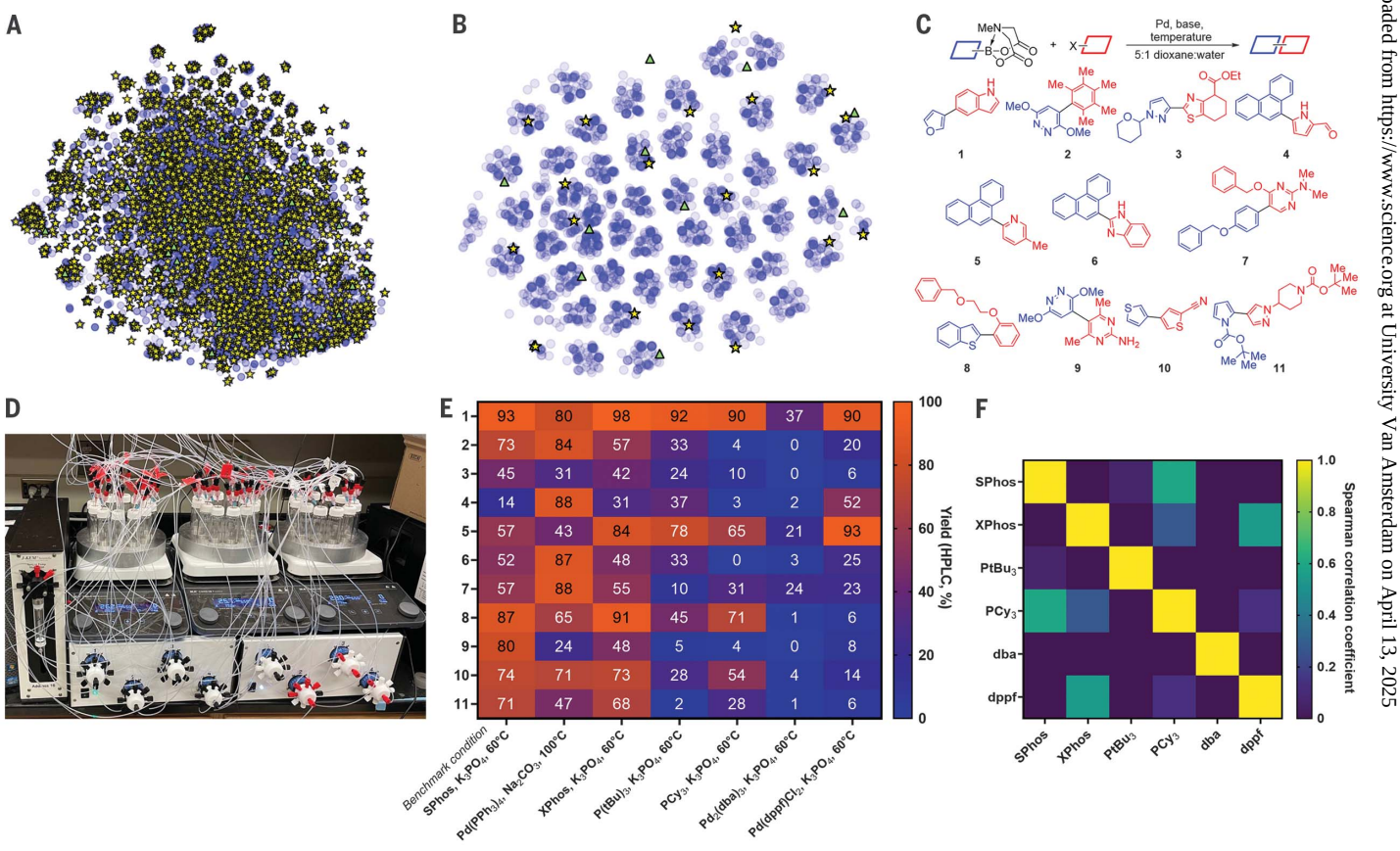

Fig. 2. Automated synthesis of the initial training set. (**A**) T-distributed stochastic neighbor embedding (t-SNE) mapping of the substrate combinations (2688 heteroaryl products) examined in this work versus all (hetero)aryl products previously reported in the literature. Blue circles represent literature-reported products, yellow stars represent products exclusively belonging in this reported search space, and green triangles represent products present in both sets. (**B**) t-SNE mapping of the product space synthesized during the training and test sets versus the overall reaction space. Blue circles represent products belonging to the reported search space, green triangles represent products belonging to the training set, and yellow stars represent products belonging to the test set. (**C**) Reaction scheme and chemical structures of the initial training set. Me, methyl; Et, ethyl. (**D**) Photo of the automated synthesis instrument used in this work. (**E**) Initial training set with the benchmark condition; all other common palladium catalysts reported in the literature; and a condition with the most common catalyst [$Pd(PPh_3)_4$], base (Na_2CO_3), temperature (100°C), and solvent (dioxane:water) used in the literature. Yields are the average of two automated repetitions (±2% deviation), measured by LCMS-UV/Vis with an authentic product standard (response factor to phenanthrene; SM section 9). HPLC, high-performance liquid chromatography. (**F**) Yield-based similarity between Pd-based catalysts differing only in organic ligands. Each square quantifies the Spearman rank correlation coefficient between yields obtained for each of the 11 substrate pairs. Two pairs of ligands (XPhos and dppf, and SPhos and PCy_3) were highly correlated and redundant. $PtBu_3$, tri-tert-butylphosphine; PCy_3, tricyclohexylphosphine; dba, dibenzylideneacetone; dppf, 1,1′-bis(diphenylphosphino)ferrocene.

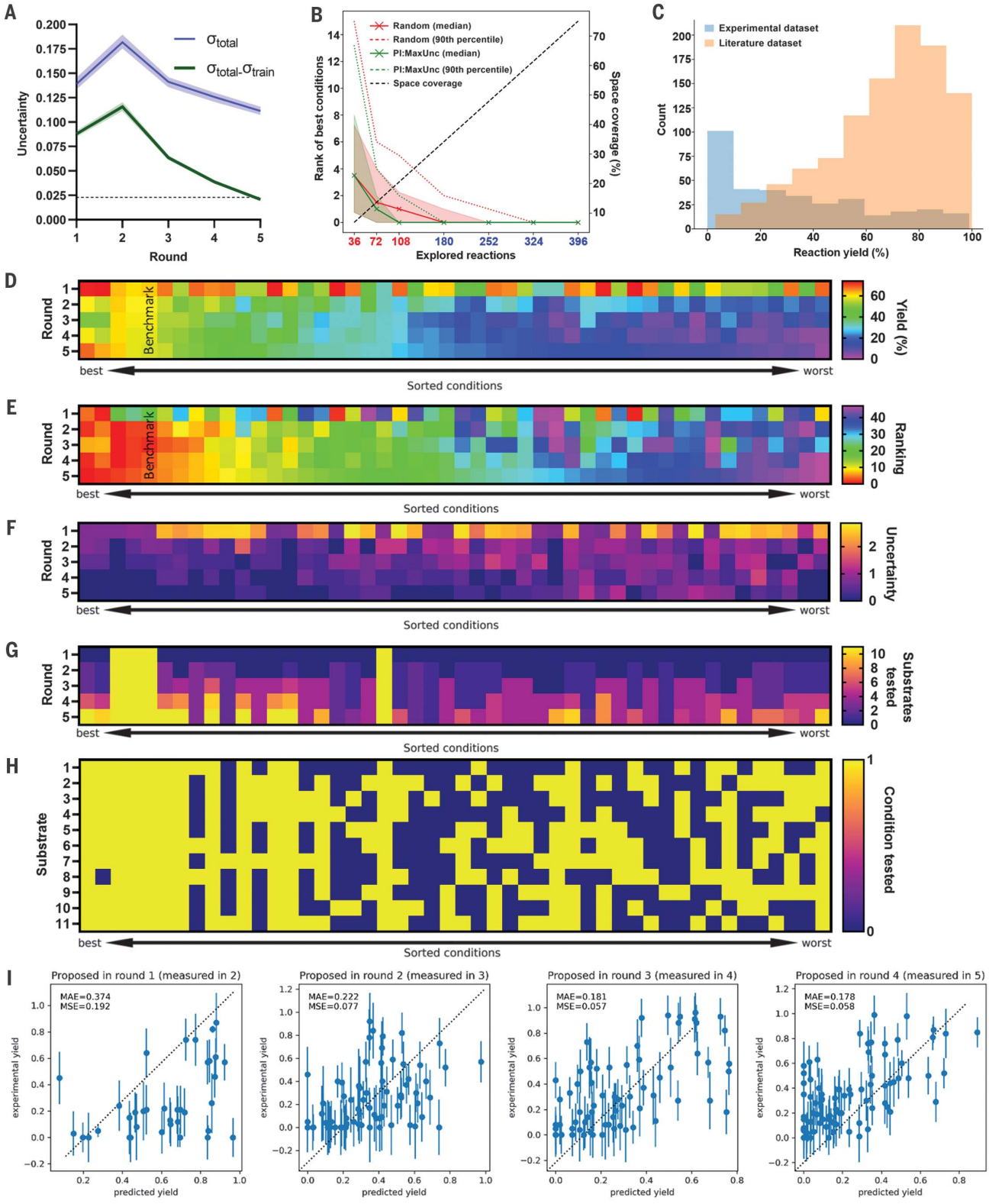

Fig. 3. Closed-loop experimentation and analysis. (**A**) Convergence of the model's uncertainty. Dashed horizontal line depicts threshold for $\bar{\sigma}_{total} - \bar{\sigma}_{train}$ obtained in calibration simulations. The shaded areas correspond to 95% confidence interval computed by repeatedly training the model 10 times. (**B**) Comparison of ML-guided searches versus random searches for general conditions. Simulations with both model selection policy [probability of improvement (PI) in conditions space and maximum uncertainty (MaxUnc) in substrate space for given conditions; abbreviated as PI:MaxUnc and corresponding to green lines] and random selection of the next reactions (red lines) were repeated 100 times to evaluate the random factor in the algorithm (random initialization of neural network weights as well as selection of the next step in the random baseline). The shaded areas mark the

interquartile range. (**C**) Comparison of yield distribution between literature-reported reactions and those explored in this work. (**D**) Color scale indicates the percent yield per general condition as perceived by the ML model at the conclusion of a given optimization round, 1 to 5. Along the horizontal axis, the conditions are ranked according to the ML prediction after round 5. (**E**) Ranking per general condition per round as perceived by the ML model. (**F**) Uncertainty per general condition per round as perceived by the ML model (computed from 10 repetitions). (**G**) Number of substrates tested per general condition per round of closed-loop optimization. (**H**) Coverage of reaction space tested by round 5 of closed-loop experimentation. A value of 1 indicates that the condition was tested, and a value of 0 indicates that the condition was not tested. (**I**) Distributions of yields measured and predicted prior to the measurement for each round of closed-loop optimization. MAE, mean absolute error; MSE, mean squared error.

spectrometry ultraviolet/visible spectroscopy (LCMS-UV/Vis) response factor curves, which enabled us to automatically measure the yields of automated reactions (SM sections 8 and 9).

Data-guided down-selection of conditions

We considered four condition variables—solvent, base, catalyst and ligand, and temperature. As our aim was to test a broad range of conditions, we initially down-selected representatives of condition classes on the basis of not only their extent of prior use from our earlier comprehensive literature analysis (*13*) but also structural and functional diversity. For instance, whereas the two most commonly used solvents in the literature are dioxane and dimethoxyethane, they both belong to the same solvent class of ethers, and so we selected only one of them (dioxane). Similar reasoning led us to keep only one carbonate base [in (*13*), we showed that the nature of the cation did not alter the yields]. We selected 100°C as the most frequently used temperature in the literature, as well as 60°C, which was used in the previously developed benchmark protocol (*14*). In the end, we selected three solvents (dioxane, toluene, and dimethylformamide, all used in 5:1 mixture with water), two bases (sodium carbonate and potassium phosphate), two temperatures (60° and 100°C), and seven catalysts [Pd SPhos G4, $Pd(PPh_3)_4$, Pd XPhos G4, Pd P(tBu)$_3$ G4, Pd PCy_3 G4, $Pd_2(dba)_3$, and $Pd(dppf)Cl_2$; G4 refers to the fourth-generation Buchwald precatalyst] to evaluate. The down-selected 11 building block combinations described above were tested under an initial set of conditions to "seed" the ML optimization (Fig. 2C) and then tested iteratively under a much broader set of conditions during the ML-guided optimization phase.

Seeding experiments, reaction standardization, and conditions space

All reactions were performed automatically on the robotic system shown in Fig. 2D. Before solvent addition, heating, and stirring, reaction mixtures were purged with 10 automated vacuum and argon cycles, which led to highly reproducible reaction yields (fig. S11). This automated Schlenk process was necessary—even when using air-stable precatalysts and building blocks—for reproducibility. To "seed" the optimization procedure, we performed all couplings between the aforementioned 11 pairs of substrates, each under seven different conditions: those corresponding to the *JACS* 2009 benchmark (5:1 dioxane:water, 60°C, K_3PO_4, Pd SPhos G4); same base and solvents but with the other selected palladium catalysts [Pd XPhos G4, Pd P(tBu)$_3$ G4, Pd PCy_3 G4, $Pd_2(dba)_3$, and $Pd(dppf)Cl_2$]; and a condition with the most common catalyst [$Pd(PPh_3)_4$], base (Na_2CO_3), temperature (100°C), and solvent (dioxane:water) used in the literature (Fig. 2E). When each reaction was repeated twice, the yields exhibited only ±2% deviation, underscoring one of the key advantages of automated experimentation [indeed, it has previously been reported (*30*) that repetition of the same reaction even by the same human experts entails variability of ~10 to 15%].

This initial round of experiments also allowed us to identify catalysts that, for different substrate pairs, systematically gave similar yields and could thus be redundant. Such functional rather than structural similarity is quantified by the Spearman rank matrix shown in Fig. 2F and correlating yields obtained for all 11 substrate pairs using two different catalyst ligands—in this representation, redundant catalysts correspond to high-correlation, off-diagonal elements (e.g., XPhos and dppf, or PCy_3 and SPhos). On the basis of this analysis, we eliminated PCy_3 and dppf from our pool of ligands to decrease redundancy, and we eliminated $Pd_2(dba)_3$ because of poor performance (<5% yield for 8 of 11 substrates), yielding a full space to be explored of 528 reactions (11 substrates × 2 temperatures × 2 bases × 3 solvents × 4 catalysts).

Uncertainty-minimizing ML for generality

A reaction condition can be considered maximally general when it provides the highest average yield across the widest range of chemical space. Optimization for generality is an unsolved and underexplored challenge in the evolving field of ML. We thus considered an alternative de novo approach, where small sets of highly reproducible data are generated on demand during ML-guided closed-loop optimization, including negative data vastly underrepresented in existing datasets. We also decided to strategically focus the ML algorithm on decreasing model uncertainty and thereby maximize the efficiency of the learning process.

Denoting the set of possible reaction conditions as $C = \{c\}$, a set of substrate pairs as $S = \{s\}$, and reaction yield as $y(s,c)$, our aim is to maximize the objective function given by

$$f(c) = \frac{1}{|S|}\sum\nolimits_{s\in S} y(s,c) \quad (1)$$

Then, the general conditions c_{general} are given as

$$c_{\text{general}} = \arg\max_{c\in C} f(c) \quad (2)$$

At first glance, the problem of identifying c_{general} in the least number of experiments resembles standard Bayesian optimization (BO). However, there is a substantial difference: In all BO algorithms, each experiment or measurement performed immediately provides information about the objective function desired for optimization. In contrast, experimental evaluation of $f(c)$ in our problem requires multiple experiments (because summation in Eq. 1 runs over the entire set S)—that is, determination of $f(c)$ for given conditions requires experiments with every pair of substrates in the S set. To address this problem, we modified the standard BO approach by constructing a surrogate model for predicting reaction yields, $\hat{y}(s,c)$. We then used its predictions to estimate $\hat{f}(c)$ according to Eq. 1 and using the model's prediction for the yet-unperformed reactions. Note that in standard BO, we would have observed $f(c)$ for the "seen" conditions and estimated for the "unseen" ones; in our case, $f(c)$ was estimated even for the already-tested conditions, unless the entire substrate space S had already been tested. On the basis of these considerations, the optimization over C (selection of the next conditions to examine) is performed with standard BO techniques, whereas sampling of S is achieved using an active learning approach—each of these techniques is known on its own (*8*, *31*), but the combination of the two (where the observation of the BO objective can be incomplete) seems to be unknown. In particular, we decided to choose substrate pairs on the basis of the model's prediction of uncertainty for given substrates under given reaction conditions: the highly uncertain (low-confidence) predictions indicate missing information, and providing the model with the corresponding experimental data should decrease its uncertainty the most.

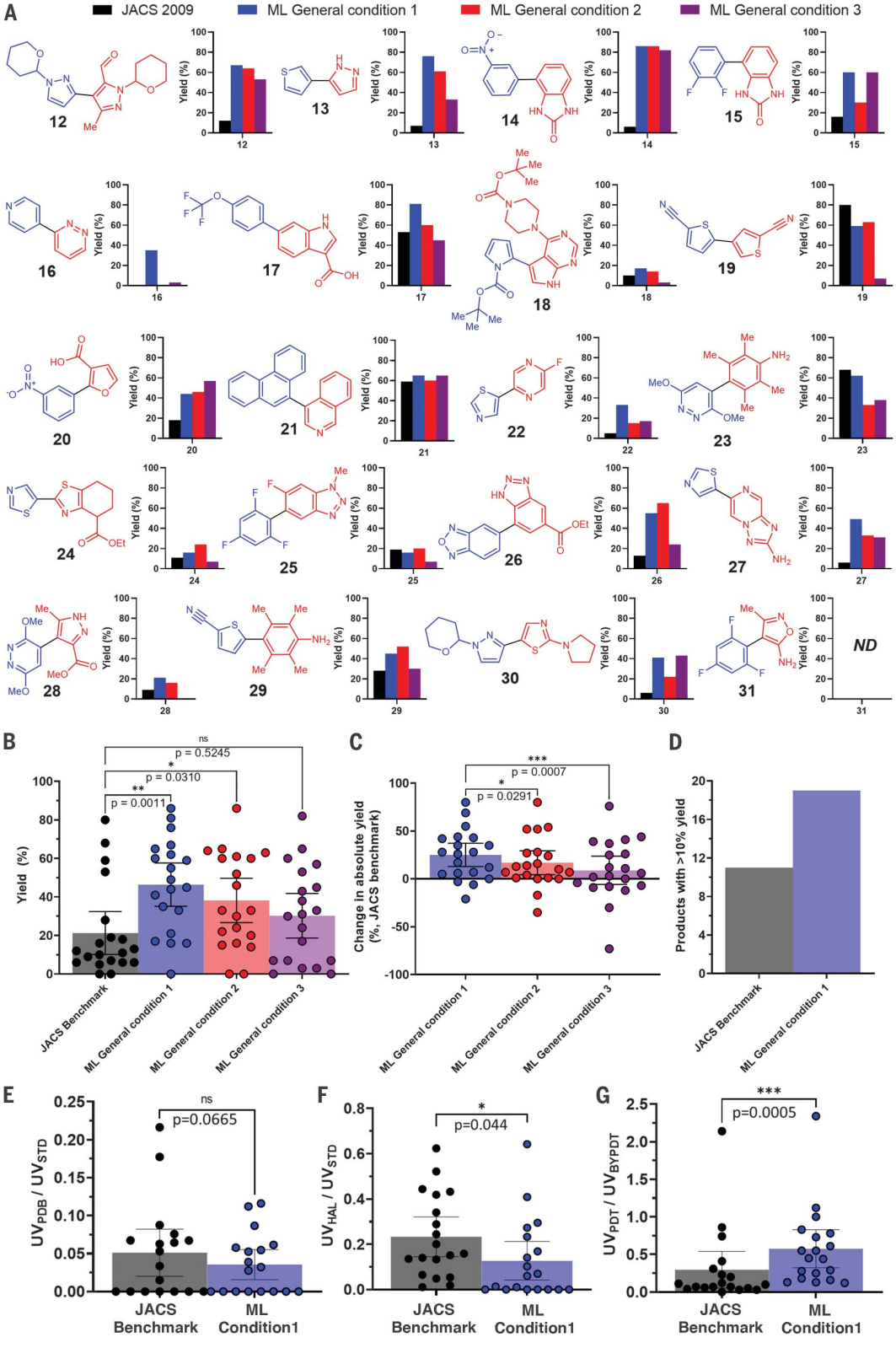

Fig. 4. Test set for ML-discovered reaction conditions. (**A**) Set of 20 diverse compounds from outside the training set selected to test whether the discovered general reaction conditions translate to other diverse heteroaryl product classes. *JACS* 2009 condition: 5:1 dioxane:water, 60°C, K_3PO_4, Pd SPhos G4. ML General condition 1: 5:1 dioxane:water, 100°C, Na_2CO_3, Pd XPhos G4. ML General condition 2: 5:1 dioxane:water, 100°C, Na_2CO_3, Pd SPhos G4. ML general condition 3: 5:1 dioxane:water, 100°C, Na_2CO_3, $Pd(PPh_3)_4$. ND, not detected. (**B**) Jitter plot showing the performance of the top ML conditions versus the benchmark. Brackets indicate 95% confidence interval. (**C**) Jitter plot showing the relative performance in change of yield of the top ML conditions versus the benchmark. Brackets indicate 95% confidence interval. (**D**) The number of products per general condition with >10% yield measured. (**E**) Relative protodeboronation per condition, as measured by integrated UV peak area (UV_{PDB}) standardized to the internal standard (UV_{STD}). (**F**) Relative remaining halide per condition, as measured by integrated UV peak area (UV_{HAL}) standardized to the internal standard (UV_{STD}). (**G**) Relative product formation per condition (UV_{PDT}) relative to by-product formation (UV_{BYPDT}).

For uncertainty estimation, we sought a model offering prediction uncertainty commensurate with prediction error—for instance, highly confident predictions with high error are undesirable. Per the analyses of numerous neural-network (NN) and Gaussian process (GP) models (SM section 3), we ultimately selected an ensemble of GP supplemented with a NN kernel component [GPE(NN)]; similar approaches were recently used in BO (*32*) and interactive learning (*33*). Such a model is particularly appealing because of its flexibility (the similarity metric between different conditions will be learned from the data) and the reliability of the prediction uncertainties

(e.g., it is guaranteed that the prediction of a test sample will not be more confident than a training sample). For the selection of conditions to be tested, we chose as an acquisition function the probability of improvement (PI).

Closed-loop, ML-driven optimization with robotic experimentation

The GPE(NN)/PI model guided the automated experiments over the down-selected search space. We worked in experimental batches, meaning that multiple experiments were performed before the theoretical model was updated. Within each batch, the algorithm formed a "priority queue" of unexplored reactions by sorting and selecting conditions according to the computed PI and substrate pairs according to the prediction uncertainty. The batch size for rounds 1 and 2 of optimization was 36 duplicated reactions, followed by 72 and 84 unduplicated reactions for rounds 3 and 4 and round 5, respectively.

Over the closed-loop rounds, the model's uncertainty decreased and converged at the fifth round to the threshold obtained during calibration simulations (Fig. 3A and fig. S20), suggesting that the model gained sufficient knowledge about the whole space, at which point the optimization was terminated. This strategy converges to this optimum in about half as many reactions as does random sampling (Fig. 3B) and with a higher likelihood of success compared with typical BO strategies (fig. S30). As the algorithm explored the reaction-condition space, reaction yields for our dataset were distributed more or less uniformly over the range of possible values (Fig. 3C). In other words, our protocol learned by probing both low- and high-yielding conditions. This contrasts with the distribution of yields in published reaction sets—such yields are heavily skewed toward positive outcomes, which, as we discussed in many of our previous works on computerized synthesis (*34*–*36*), limits the usefulness of approaches aiming to learn from published datasets.

In this dataset, the discovered top-1 condition conferred 72% average yield across all 11 substrates, whereas the benchmark condition [found to also be the top-5 (i.e., fifth best) condition] conferred 64% average yield. To understand how the model arrived at this optimum, we examined the model's perception of the average yield and ranking of each general condition per round (Fig. 3, D and E). Within the first two rounds, the model gains the ability to accurately categorize these conditions into high, medium, and low overall average yield and, in the subsequent rounds, establishes the correct ranking within these categories. The increasing accuracy of the model over the course of the experiment is recapitulated in Fig. 3F, which shows the model's ranking uncertainty decreasing between rounds and is especially apparent for the top conditions. The model chose to test a few substrates per round across many conditions for the first three rounds, followed by primarily filling in the top conditions in the latter rounds (Fig. 3G). By the fifth round, the model explored nearly all of the top-7 conditions, which corresponds to every condition with >50% overall average yield, as estimated by the model (Fig. 3H).

Finally, we analyzed the yields of reactions the model requested in order to gain more information about the reaction-condition space (Fig. 3I). These values are not expected to increase as the optimization progresses, because the yield of a single experiment is not our objective. Given the uncertainty-guided selection of substrates, one could even expect the opposite: Once a set of suitable conditions is identified, further exploitation should involve lower-yielding reactions to verify that the found candidate conditions are indeed better, as well as to increase confidence of the estimate of $f(c)$. The results shown in Fig. 3I indicate that after exploring good reactions in the second iteration, the model gradually shifted its attention toward the parts of the reaction-condition space that can be considered as "negative examples" (and, in doing so, improved its prediction accuracy). From these results, it appears that (i) relatively good candidate solutions were identified early, (ii) the model initially tried to look for better-yielding reactions (to find better candidates), and (iii) more and more attention was dedicated to decreasing the uncertainty of its estimates as the "loop" progressed.

Quantifying generality

After the discovery of higher-yielding general conditions in the training set, we next sought to determine whether the learning would transfer to substrates outside of the optimization—specifically, over 20 substrate pairs chosen [by the Butina algorithm (*37*)] to maximize dissimilarity to the training set while ensuring coverage of the heterocyclic substructure and functional group space (Fig. 2B). We then set out to synthesize and purify all of the computer's suggestions and test them against the benchmark condition and the top-3 highest yielding general reaction conditions discovered during the closed-loop optimization (Fig. 4A), as ranked by the model after the completion of round 5. Despite including some very challenging building block combinations, this process was 95% successful, with only one product having no measurable yield under all four conditions.

The ML-discovered general reaction conditions performed substantially better than the previously reported and widely used benchmark condition (*14*). The top-2 conditions provided statistically significant increases in average yield compared with the benchmark, with the top condition doubling the overall average yield from 21% to 46% (Fig. 4B). Comparing the relative increase in yield reveals statistically significant differences between the top-1 and both the top-2 and top-3 conditions (Fig. 4C). Notably, the experimental yields correlate with the predicted ranking of the conditions such that the yield for the top-1 is higher than that for the top-2, which, in turn, is higher than that for the top-3. In the context of functional discovery efforts, the binary capacity to isolate or not isolate testable quantities of purified targeted compounds is arguably even more important than the specific percent yield. We estimate that the practical limit for isolating purified products is 10% yield. For the benchmark condition, only 11/20 targeted products cleared this bar, whereas this fraction rose to 19/20 for the top-1 condition (Fig. 4D).

Extending the reaction times for couplings that were low yielding under the benchmark conditions did not increase yields (fig. S33). Comprehensive analysis of by-products and product formation for all 20 reactions (SM section 10) demonstrated that a favorable shift from the former to the latter accompanies the shift from the benchmark to the ML-discovered reaction conditions. Specifically, the ML-discovered conditions were associated with a trend toward decreased protodeboronation (Fig. 4E), increased halide conversion (Fig. 4F), and an overall statistically significant increase in the ratio of product to total by-products formation (Fig. 4G) (0.30 ± 0.12 for *JACS* 2009 versus 0.58 ± 0.12 for ML conditions; $P = 0.0005$).

Outlook

The straightforward workflow developed here has enabled the accelerated discovery of improved general reaction conditions for difficult C–C bond forming reactions, representing a key step toward increasing the efficiency, generality, and accessibility of small molecule synthesis. This result also highlights the power of down-selection as an entry point into large multidimensional search spaces, the distinct advantages of a de novo ML approach for navigating such spaces by generating datasets that evenly reflect the reality of positive and negative data during optimization, and the particular suitability of robotized chemistry for generating high-quality, reproducible data. Future studies will incorporate next-generation ligands and reagents to yield further improved general reaction conditions, creating an actionable path for automated small molecule synthesis to achieve reaction efficiencies approaching that of automated peptide synthesis. This general workflow should be applicable to

the optimization of reactions beyond SMC, as it did not require extensive prior literature data to be successful.

REFERENCES AND NOTES

1. R. B. Merrifield, *Science* **150**, 178–185 (1965).
2. M. H. Caruthers, *Science* **230**, 281–285 (1985).
3. O. J. Plante, E. R. Palmacci, P. H. Seeberger, *Science* **291**, 1523–1527 (2001).
4. I. W. Davies, *Nature* **570**, 175–181 (2019).
5. D. T. Ahneman, J. G. Estrada, S. Lin, S. D. Dreher, A. G. Doyle, *Science* **360**, 186–190 (2018).
6. C. W. Coley *et al*., *Science* **365**, eaax1566 (2019).
7. M. Christensen *et al*., *Commun. Chem.* **4**, 112 (2021).
8. B. J. Shields *et al*., *Nature* **590**, 89–96 (2021).
9. A. B. Clayton *et al*., *React. Chem. Eng.* **4**, 1545–1554 (2019).
10. P. A. Cox, A. G. Leach, A. D. Campbell, G. C. Lloyd-Jones, *J. Am. Chem. Soc.* **138**, 9145–9157 (2016).
11. P. A. Cox *et al*., *J. Am. Chem. Soc.* **139**, 13156–13165 (2017).
12. H. L. D. Hayes *et al*., *J. Am. Chem. Soc.* **143**, 14814–14826 (2021).
13. W. Beker *et al*., *J. Am. Chem. Soc.* **144**, 4819–4827 (2022).
14. D. M. Knapp, E. P. Gillis, M. D. Burke, *J. Am. Chem. Soc.* **131**, 6961–6963 (2009).
15. J. Li *et al*., *Science* **347**, 1221–1226 (2015).
16. A. Buitrago Santanilla *et al*., *Science* **347**, 49–53 (2015).
17. A.-C. Bédard *et al*., *Science* **361**, 1220–1225 (2018).
18. J. M. Granda, L. Donina, V. Dragone, D.-L. Long, L. Cronin, *Nature* **559**, 377–381 (2018).
19. S. Lin *et al*., *Science* **361**, eaar6236 (2018).
20. D. Perera *et al*., *Science* **359**, 429–434 (2018).
21. M. Trobe, M. D. Burke, *Angew. Chem. Int. Ed.* **57**, 4192–4214 (2018).
22. S. Steiner *et al*., *Science* **363**, eaav2211 (2019).
23. S. Chatterjee, M. Guidi, P. H. Seeberger, K. Gilmore, *Nature* **579**, 379–384 (2020).
24. S. H. M. Mehr, M. Craven, A. I. Leonov, G. Keenan, L. Cronin, *Science* **370**, 101–108 (2020).
25. C. Liu *et al*., *Nat. Chem.* **13**, 451–457 (2021).
26. D. J. Blair *et al*., *Nature* **604**, 92–97 (2022).
27. V. Fasano *et al*., ChemRxiv [Preprint] (2020). https://doi.org/10.26434/chemrxiv-2022-zwctv-v2.
28. A. Bender *et al*., *Nat. Rev. Chem.* **6**, 428–442 (2022).
29. D. Bajusz, A. Rácz, K. Héberger, *J. Cheminform.* **7**, 20 (2015).
30. F. S. Emami *et al*., *Angew. Chem. Int. Ed.* **54**, 10797–10801 (2015).
31. Y. Gong, D. Xue, G. Chuai, J. Yu, Q. Liu, *Chem. Sci.* **12**, 14459–14472 (2021).
32. M. Wistuba, J. Grabocka, arXiv:2101.07667 [cs.LG] (2021).
33. J. Bowden, J. Song, Y. Chen, Y. Yue, T. A. Desautels, "Deep kernel Bayesian optimization," 37th Conference on Uncertainty in Artificial Intelligence (UAI 2021), 27 to 30 July 2021.
34. T. Klucznik *et al*., *Chem* **4**, 522–532 (2018).
35. B. Mikulak-Klucznik *et al*., *Nature* **588**, 83–88 (2020).
36. F. Strieth-Kalthoff *et al*., *Angew. Chem. Int. Ed.* **61**, e202204647 (2022).
37. D. Butina, *J. Chem. Inf. Comput. Sci.* **39**, 747–750 (1999).
38. N. H. Angello *et al*., Closed-loop optimization of general conditions for heteroaryl Suzuki coupling, version 0.1.2, Zenodo (2022); https://doi.org/10.5281/zenodo.7099435.
39. N. H. Angello *et al*., Closed-loop optimization of general conditions for heteroaryl Suzuki coupling, version 2, Zenodo (2022); https://doi.org/10.5281/zenodo.7106075.
40. verysure, P. Schleich, aspuru-guzik-group/b3: Release for Science Publication, version 1.0, Zenodo (2022); https://doi.org/10.5281/zenodo.7094835.
41. N. H. Angello *et al*., Closed-loop optimization of general conditions for heteroaryl Suzuki coupling, version 1.0, Zenodo (2022); https://doi.org/10.5281/zenodo.6517012.

ACKNOWLEDGMENTS

Funding: This work was supported by the Defense Advanced Research Projects Agency under the Accelerated Molecular Discovery Program (cooperative agreement no. HR00111920027 dated 1 August 2019) to M.D.B., B.A.G., and A.A.-G. The content of the information presented in this work does not necessarily reflect the position or the policy of the US government. This work was also supported by funding from the Molecule Maker Lab Institute: An AI Research Institutes program supported by NSF under award no. 2019897 to C.M.S. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect those of the National Science Foundation. N.H.A. was supported by the Department of Defense (DoD) through the National Defense Science and Engineering Graduate (NDSEG) Fellowship Program. B.A.G. was supported by the Institute for Basic Science, Korea (project code IBS-R020-D1). **Author contributions:** N.H.A., V.R., W.B., A.W., R.R., B.A.G., and M.D.B. conceived of the study. N.H.A. and E.R.J. developed and validated the automated synthesizer. V.R. and N.H.A. performed the automated synthesis experiments. V.R. and N.H.A. purified, characterized, and generated response factor curves for new compounds. W.B., A.W., R.R., and B.A.G. developed the closed-loop machine-learning strategy, building block clustering, and chemical space quantifications. N.H.A. and T.C.W. data mined the substrate scope building block set. M.D.B., B.A.G., A.A.-G., and C.M.S. supervised the research. The manuscript was written by N.H.A., V.R., W.B., A.W., R.R., B.A.G., and M.D.B., with contributions from all other authors. **Competing interests:** The University of Illinois has filed patent applications related to MIDA boronates (US20170002026A1, US9845317B2, WO2011103435A3, and US20160207943A1 with M.D.B. as inventor). A.A.-G. is chief visionary officer and board member of Kebotix, Inc., a company that carries out closed-loop molecular materials discovery. **Data and materials availability:** All data and code generated as part of this study are freely accessible either in the supplementary materials or in open repositories. Code for the simulation example, model selection, and substrates clusterization as well machine- and human-readable versions of the reaction data are freely accessible at Zenodo (*38*). The automated synthesis code, machine parts list and build guide, datamined commercially available building block set, checklist for reporting and evaluating machine learning models, and tabulated numerical data underlying Figs. 2 to 4 are deposited at Zenodo (*39*). The code used in the building block selection process is available at Zenodo (*40*). All code related to the closed-loop optimization is freely available at Zenodo (*41*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adc8743
Materials and Methods
Supplementary Text
Figs. S1 to S56
Tables S1 to S3
References (*42*–*50*)

Submitted 5 May 2022; resubmitted 23 August 2022
Accepted 29 September 2022
10.1126/science.adc8743