Delocalized, Asynchronous, Closed-Loop Discovery of Organic Laser Emitters

Felix Strieth-Kalthoff,[1,2,†] Han Hao,[1,2,†] Vandana Rathore,[3,4] Joshua Derasp,[5] Théophile Gaudin,[2] Nicholas H. Angello,[3,4] Martin Seifrid,[1,2] Ekaterina Trushina,[6] Mason Guy,[5] Junliang Liu,[5] Xun Tang,[7] Masashi Mamada,[7] Wesley Wang,[3,4] Tuul Tsagaantsooj,[7] Cyrille Lavigne,[1,2] Robert Pollice,[1,2] Tony C. Wu,[1,2] Kazuhiro Hotta,[1,2] Leticia Bodo,[1] Shangyu Li,[1] Mohammad Haddadnia,[1,8] Agnieszka Wołos,[9] Rafał Roszak,[9] Cher-Tian Ser,[1,2] Carlota Bozal-Ginesta,[1,2,10] Riley J. Hickman,[1,2] Jenya Vestfrid,[1,2] Andrés Aguilar-Granda,[1,2] Elena L. Klimareva,[6] Ralph C. Sigerson,[6] Wenduan Hou,[6] Daniel Gahler,[6] Slawomir Lach,[6] Adrian Warzybok,[6,11] Oleg Borodin,[6] Simon Rohrbach,[6] Benjamin Sanchez-Lengeling,[12] Chihaya Adachi,[7]* Bartosz A. Grzybowski,[9,13,14,15,17]* Leroy Cronin,[6,17]* Jason E. Hein,[5,16,17]* Martin D. Burke,[3,4,17,18,19,20]* Alán Aspuru-Guzik[1,2,8,17,21,22,23]*

[1] Department of Chemistry, University of Toronto, Toronto, ON, Canada.

[2] Department of Computer Science, University of Toronto, Toronto, ON, Canada.

[3] Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL, USA.

[4] Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL, USA.

[5] Department of Chemistry, University of British Columbia, Vancouver, BC, Canada.

[6] School of Chemistry, University of Glasgow, Glasgow, UK.

[7] Center for Organic Photonics and Electronics Research (OPERA), Kyushu University, Fukuoka, Japan.

[8] Vector Institute for Artificial Intelligence, Toronto, ON, Canada.

[9] Allchemy Inc., Highland, IN, USA.

[10] Catalonia Institute for Energy Research; Barcelona, Spain.

[11] Department of Chemical Physics, Jagiellonian University, Krakow, Poland.

[12] Google Research, Brain Team, USA.

[13] Institute of Organic Chemistry, Polish Academy of Sciences, Warsaw, Poland.

[14] Center for Soft and Living Matter, Institute for Basic Science, Ulsan, Republic of Korea.

[15] Department of Chemistry, Ulsan Institute of Science and Technology, Ulsan, Republic of Korea.

[16] Department of Chemistry, University of Bergen, Norway.

[17] Acceleration Consortium, University of Toronto, Toronto, ON, Canada.

[18] Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA.

[19] Cancer Center at Illinois, University of Illinois at Urbana-Champaign, Urbana, IL, USA.

[20] Carle Illinois College of Medicine, University of Illinois at Urbana-Champaign, Urbana, IL, USA.

[21] Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON, Canada.

[22] Department of Materials Science and Engineering, University of Toronto, Toronto, ON, Canada.

[23] Canadian Institute for Advanced Research (CIFAR), Toronto, ON, Canada.

† These authors contributed equally to this work. The order of authorship was determined through a coin flip.

* Corresponding authors: Emails: adachi@cstf.kyushu-u.ac.jp (C.A.), nanogrzybowski@gmail.com (B.A.G.), Lee.Cronin@glasgow.ac.uk (L.C.), jhein@chem.ubc.ca (J.E.H.), mdburke@illinois.edu (M.D.B), aspuru@utoronto.ca (A.A.-G.).

Contemporary materials discovery requires intricate sequences of synthesis, formulation and characterization that often span multiple locations with specialized expertise or instrumentation. To accelerate these workflows, we present a cloud-based strategy that enables delocalized and asynchronous design–make–test–analyze cycles. We showcase this approach through the exploration of molecular gain materials for organic solid-state lasers as a frontier application in molecular optoelectronics. Distributed robotic synthesis and in-line property characterization, orchestrated by a cloud-based AI experiment planner, resulted in the discovery of 21 new state-of-the-art materials. Automated gram-scale synthesis ultimately allowed for the verification of best-in-class stimulated emission in a thin-film device. Demonstrating the asynchronous integration of five laboratories across the globe, this workflow provides a blueprint for delocalizing – and democratizing – scientific discovery.

Introduction and Background

Efficient molecular discovery for diverse applications in medicine,[1] optoelectronics,[2] or energy storage,[3] requires intertwined loops of molecular synthesis, property characterization, formulation and system-level testing. There is an undebated necessity for accelerating these human-centric and often laborious workflows, in order to meet the societal demands for enhanced materials.[4] In response to these demands, there has been a surge in the development of computational [5] and artificial intelligence (AI) tools for materials science,[6,7] along with major advances in automation and high-throughput experimentation (HTE),[8,9] – and eventually, the integration of both automated experimentation and automated decision making into "self-driving laboratories" (SDLs).[10–13] Such efforts have shown the potential to significantly accelerate local units within these high-level workflows – examples include the optimization of reaction conditions, or the identification of ideal formulation parameters.[14–17]

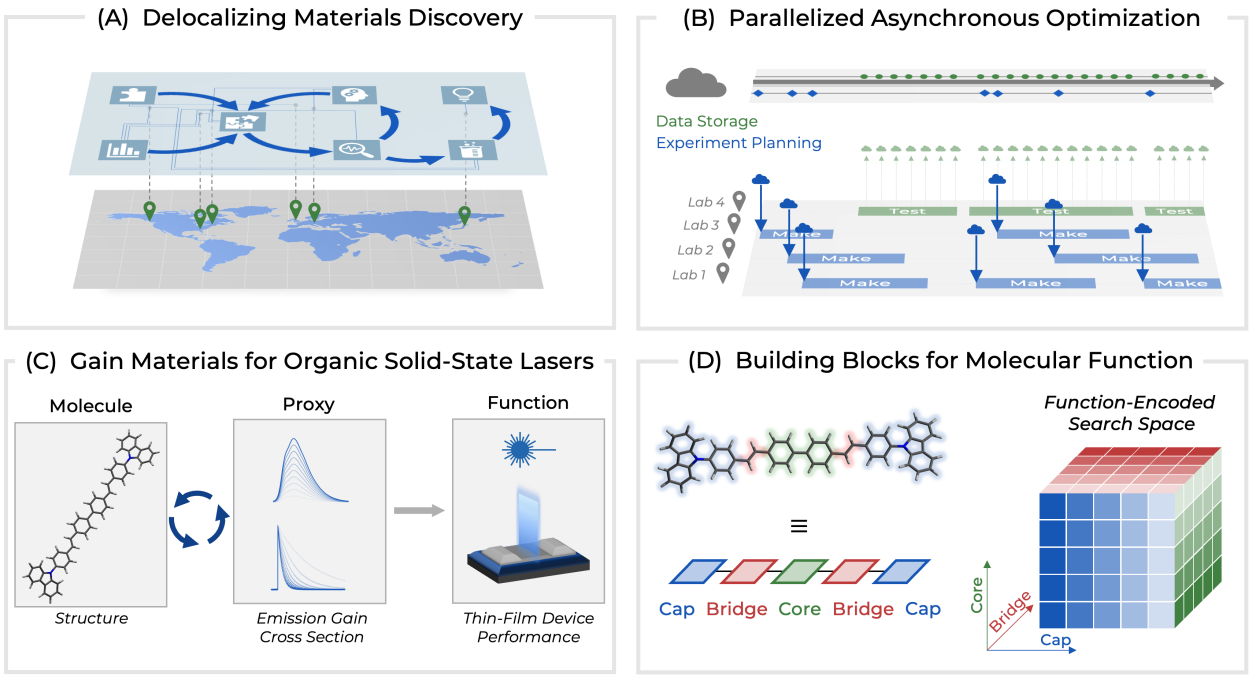

Figure 1: Delocalizing molecular materials discovery targeting OSL emitters. (**A**) Delocalized complex materials discovery workflows (traditional "design–make–test-analyze" cycles) over multiple sites, orchestrated by a single cloud-based application. Markers on the map correspond to the geographical locations of the laboratories participating in the current study. (**B**) Experimental bottlenecks can be bypassed by distributing experiments over multiple asynchronous worker "threads" running in different laboratories. In this work, we distribute molecular synthesis over multiple sites, orchestrated by a central experiment planning algorithm. (**C**) Schematic depiction of discovering gain materials for organic solid-state lasers (OSLs) by optimizing a proxy objective (the emission gain cross section) over multiple cycles before evaluating top candidates in thin-film devices. (**D**) Formal disconnection of 4,4'-Bis[*(E)*-4-(*N*-carbazoyl)styryl]biphenyl (BSBCz) into symmetric *cap*, *bridge* and *core* building blocks, upon which a vast search space of functional BSBCz-like molecules can be enumerated from sets of building blocks.

The evolution of the second generation of SDLs is driven by distribution and delocalization as major paradigms. As the complexity of discovery workflows continues to grow, the integration of advanced experimental and computational modules becomes essential. These units often rely on domain expertise and specific instrumentation, resulting in their dispersion over multiple geographical locations and time zones (Fig. 1a). In addition, the capacity to parallelize experiments over multiple modules offers solutions to enhance throughput and rapidly replicate experimental findings, ultimately facilitating democratized discovery. In both scenarios, *distributed experimentation* is the key factor for accelerating more complex discoveries. However, the implementation of distributed experimentation necessitates a central platform with clearly defined standards for communication, data transfer, and experiment planning,[18–21] while flexibly accounting for the inevitable delays and asynchrony between different sites (Fig. 1b). The discussed aspects hold particular significance in expediting molecular materials discovery, where synthesis remains the main bottleneck.[12] In fact, generalizable and readily automatable synthetic protocols have remained elusive for all but the most prominent classes of biomolecules, i.e. peptides,[22] oligosaccharides,[23] or oligonucleotides[24] – impeding broad molecular materials discovery endeavors.

Against this background, we herein demonstrate a decentralized discovery workflow, showcasing the automated design, synthesis, and testing of gain materials for organic solid-state lasers (OSL), which are characterized by best-in-class emission gain cross section in solution and amplified spontaneous emission in thin-film. The workflow relies on a closed-loop protocol encompassing synthesis planning, automated synthesis, proxy characterization, and molecular function optimization through machine learning (ML). Importantly, the discussed synthesis bottleneck is bypassed by segmenting the OSL candidate space into a building block framework,[25,26] which enables rapid, parallelizable assembly of OSL gain candidates, following a "synthesis–to–function" paradigm. Although all tasks are delocalized across five physical laboratories on three continents, they are orchestrated by a cloud server to ensure continuous learning from the incoming data, and continuous prioritization of informative experiments. This approach heralds future research campaigns in which the expertise and experimental capabilities of different SDL sites will work synergistically to expedite the discovery of functional materials.

The Experimental Engine for OSL Candidate Discovery

OSLs represent an emerging technology to provide flexible, readily processable and color-tunable lasing devices with potential applications in displays, medical devices, spectroscopy, or LiFi telecommunication (Fig. 1c).[27–29] Crucial to the development of OSLs is the emissive gain material – typically a large, conjugated molecule such as 4,4'-Bis[*(E)*-4-(*N*-carbazoyl)styryl]biphenyl (commonly referred to as BSBCz, Fig. 1d).[30,31] These linear, symmetric molecular structures are inherently amenable to modularization into LEGO-like building blocks that can be subjected to automated syntheses based on iterative Suzuki–Miyaura couplings (SMC), which have been developed in previous works.[25,26] With this general scheme, and by analogy to state-of-the-art emitters,[32,33] we conceived a palindromic Cap–Bridge–Core–Bridge–Cap architecture as a generalizable and synthesizable blueprint for powerful OSL gain materials (Fig. 1d). SMC assembly of this framework requires cap building blocks carrying a boronic acid ester, bridge precursors featuring both a halide and a protected boronic acid, and dibromo core building blocks (see Fig. 2a and Supplementary Information).

We began by surveying the catalogs of specialty chemical suppliers and defined a fragment library comprising 32 cap, 30 bridge and 161 core building blocks, spanning a hypothetical candidate space of > 150,000 putative gain materials. Building on recent advances in iterative SMC,[16,34] we conceived a generalizable two-step, five-component one-pot synthesis protocol optimized for parallel high-throughput screening, avoiding the necessity for intermediate purification, and enabling facile adaptation on different automated experimental platforms. This two-step protocol consists of an initial SMC between a cap building block and a bifunctional bridge unit, followed by an *in-situ* deprotection and double coupling with the core building block (Fig. 2a).

First-generation conditions for the two-step one-pot coupling were derived from literature reports on the well-established iterative SMC of *N*-methyliminodiacetic acid (MIDA) protected boronic acids.[25,35] This protocol showed

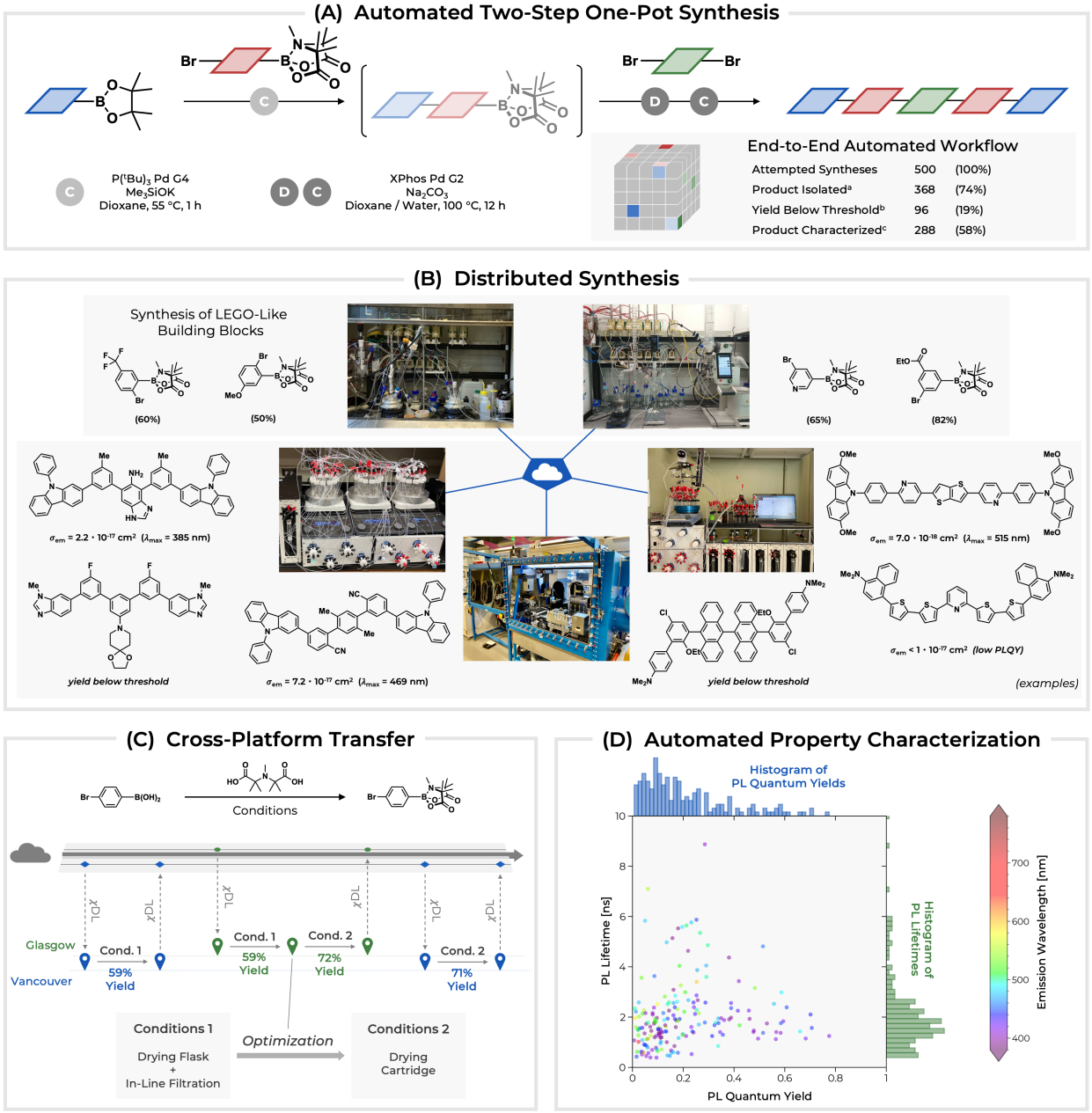

Figure 2: Overview of the modules for automated synthesis and characterization of OSL candidate molecules. (**A**) Conditions of the iterative two-step one-pot SMC coupling for synthesizing pentameric structures, and evaluation of the conditions on a representative subset of 500 target compounds, as obtained from Latin Hypercube Sampling (see Supplementary Section 3.5 for details). C: Coupling. D: Deprotection. (**B**) Selected examples of automated gram-scale synthesis of LEGO-like building blocks (top), and parallelized small-scale synthesis of OSL target molecules (bottom). (**C**) Cross-platform optimization of reproducible automated building block syntheses, enabled by the execution of standardized χDL protocols. (**D**) Scatter plot and histograms of measured photoluminescence quantum yields (horizontal axis) and photoluminescence lifetime (vertical axis) for the seed dataset of 500 attempted OSL candidate compounds (cf. a). Data points are colored according to their emission wavelength.

decent applicability for a set of target molecules with high similarity to the parent BSBCz scaffold, with product formation observed for 43 of 81 target compounds (see Supplementary Tab. S6 for further details). Across the overall candidate space, however, these conditions proved to be less effective (successful target compound detection in 32% of all cases, see Supplementary Tab. S6 for further details). Circumventing this limitation, the recently developed

2,2,2',2'-tetramethyl-N-methyliminodiacetic acid (TIDA) protecting group for boronic acids[36] allowed the use of potassium trimethylsilanoxide (TMSOK) in the first coupling step,[37,38] significantly reducing reaction time from 12 hours to 1 hour, thus minimizing side reactions (Fig. 2a). Notably, the challenging second *in-situ* coupling step was enabled by the general slow-release coupling conditions for SMCs with protected boronates, which we recently developed through AI-guided optimization.[16] In an exploratory seed campaign across 500 representative candidate pentamers selected through Latin Hypercube Sampling (see Supplementary Section 3.5 for further details), this second-generation protocol led to a substantial increase in the global synthetic hit rate (75%, Fig. 2a and Fig 2b).

The transition to bifunctional TIDA boronates as the bridge building blocks led to a shift of the initial building block space, as it required access to a library of bifunctional TIDA-protected haloboronic acids. Unlike their MIDA analogs, these are not commercially available. Exploiting our capabilities to perform automated gram-scale synthesis and purification,[39] the respective derivatization of commercially available reagents was rapidly performed across multiple laboratories, following general procedures encoded in the χDL language.[40] Notably, information transfer in the form of χDL emphasizes the importance of standardization for reproducible synthesis: executing the identical χDL protocol across different laboratories resulted in identical reaction yields, and rapidly verifiable and transferable "global" reaction optimization (Fig. 2c).

Large-scale building block preparation and parallelized small-scale target syntheses were performed on five different automated platforms across four laboratories (Fig. 2b), orchestrated from a single, readily accessible central database[41] for storing reagent availability, experiment progress, as well as all experimental and computational results (for details on the database integration, see Supplementary Section 4.1). Functional characterization of the obtained OSL target compounds was then performed from the crude reaction mixture through the automated analysis, purification and characterization workflow, as reported previously by our groups.[34] Following separation and peak identification by HPLC-MS, the collected product fraction was subjected to down-stream spectroscopic characterization in solution. From steady-state absorption and emission spectroscopy, relative quantum yield measurements and transient emission spectroscopy, lasing performance can be approximated through the emission gain cross section σ_{em} (see Supplementary Section 3.3 for further details).[42,43] This proxy objective is maximized by those molecules that simultaneously exhibit a narrow emission spectrum, a high photoluminescence quantum yield ϕ and a short emission lifetime τ. Arguably, this non-adaptive workflow of automated synthesis, purification and spectroscopic characterization represents a tradeoff between throughput and accuracy – as full spectroscopic data could only be obtained for 48% of all characterized compounds, whereas the remaining 52% of molecules were either too weak emitters (enabling only partial characterization), or the collected product fraction was too low in concentration (Fig. 2a, see Supplementary Tab. S7 for further details). An overview of the obtained data is given in Fig. 2d and Supplementary Fig. S9. While a series of target molecules with short emission lifetimes (< 1.5 ns) were found within the exploratory seed dataset, no state-of-the-art emitters were discovered in that data, as the obtained quantum yields were predominantly very low.

Asynchronous, Data-Driven Experiment Planning for OSL Candidate Discovery

Having established the experimental engine to synthesize and characterize organic laser pentamers, we sought to develop a robust computational workflow for planning and orchestrating the synthesis of novel, improved OSL gain materials, and navigate the search within the space of >150,000 potential target compounds. In the context of experiment planning, Bayesian Optimization (BO) is regarded as the gold standard for the sample-efficient optimization of unknown spaces.[44,45] Whereas BO research has largely focused on optimization over continuous parameter domains, the space of molecules is inherently discrete[46] – and therefore requires a vectorized encoding of the molecular structure for training the surrogate model, as well as a scheme to optimize recommended candidates across a discrete space.

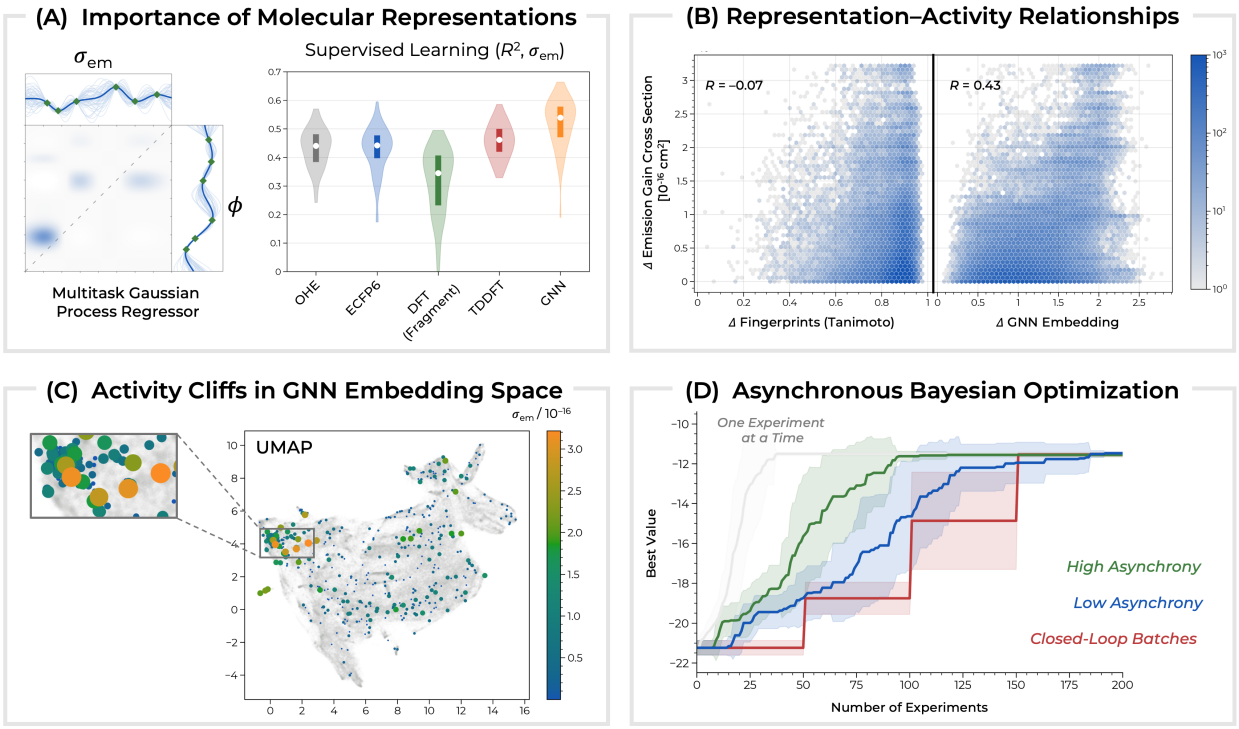

Figure 3: Bayesian optimization for OSL gain material discovery. (**A**) Supervised learning performance of a multi-task Gaussian process regressor for predicting emission gain cross section across the experimental seed dataset of 287 data points, comparing different molecular representations (OHE: one-hot encoding, ECFP6: extended-connectivity fingerprint with diameter 6, DFT (Fragment): per-fragment ground-state DFT descriptors, TDDFT: calculated excited-state descriptors, GNN: graph neural network embeddings). Results are given as R^2, averaged over 20 x 3-fold cross validation runs. (**B**) Scatter plot of pairwise molecular distances (in fingerprint or GNN embedding space), and pairwise functional distances (as difference in emission gain cross section). (**C**) Uniform manifold approximation and projection (UMAP) of the GNN embedding space (gray), and depiction of all experimentally observed data points (colored). (**D**) Benchmark of asynchronous and batch-wise Bayesian optimization strategies on the synthetic Ackley surface (6 dimensions, discretized). See Supplementary Section 4.4 for further details on the simulation of asynchronous optimization.

We evaluated a series of established structural representations (molecular fingerprints, graph-level descriptors, computed building-block descriptors) for the supervised learning of emission gain cross section. However, the regression performance of models built on these representations failed to surpass a simple one-hot encoding of building block identity (Fig. 3a, see Supplementary Section 4.3 for a full evaluation of different surrogate model types). This indicates the absence of unambiguous structure–activity relationships within the experimental data, which is well reflected in the lack of clear canonical design principles for OSL gain materials in the literature.[33] In fact, we did not observe any correlation between the structural similarity for all pairs of experimentally characterized target molecules (measured as Tanimoto similarity on fingerprints), and their respective functional similarity (Fig. 3b).

In order to obtain better predictivity, we envisioned that additional physically meaningful information obtained from quantum chemistry simulations could improve the performance of our models.[47] In a high-throughput computational campaign, excited-state properties of a large catalog of possible target molecules were approximated using time-dependent density functional theory (TD-DFT) with a vertical-gradient (VG) approximation for vibrational coupling (for details on the workflow, see Supplementary Section 4.2).[48] Including these simulated properties as molecular descriptors indeed resulted in a slightly increased predictive performance. The most significant boost in surrogate model predictivity, however, was obtained by learning a new molecular embedding from the computed data. In this manner, a graph neural network (GNN), trained on a randomized subset of the entire candidate space (training set of 92,880 molecules) to predict a set of TD-DFT properties, achieved high accuracy (R^2 = 0.86, averaged over seven computed properties, for network architecture and prediction performance, see Supplementary Section 4.2). Extracting

the embedding vector from this GNN in a transfer learning approach led to increased performance in predicting emission gain cross sections using a Gaussian process (GP) regressor (Fig. 3a). Importantly, a correlation between pairwise embedding distances and experimentally observed molecular function differences was observed (Fig. 3b), emphasizing the physical relevance of GNN embeddings as a functional molecular representation. Whilst analyzing this learned representation across the entire candidate space indicates localized domains of high lasing performance (visualized through uniform manifold approximation and projection, UMAP, see Fig. 3c), it also features a range of activity cliffs, i.e. pairs of molecules that are close in embedding space, but dissimilar in lasing performance (see Fig. 3b and Fig. 3c). The existence of such activity cliffs in embedding space leaves room for learning more globally informed representations, and emphasizes the importance of an explorative search strategy for navigating the space of OSL candidates.

Orchestrating the required experiments across experimental platforms was realized through an asynchronous Bayesian optimization workflow, using the described GNN–GP as the surrogate model (for details, see Supplementary Section 4.4). Parallelized optimization using multiple threads of experimentation is supported by maintaining a ranked catalog of recommended, currently synthesizable target compounds at all times, that can be allocated to experimental resources as available. Allocated, but incomplete experiments are handled by conditioning the model posterior on predicted "fantasy" values[45] for these data points, enabling a more global exploration of the candidate space. In benchmark experiments on synthetic surfaces, this strategy was demonstrated to be superior to conventional, batch-wise closed-loop experimentation (Fig. 3d), while maximizing the use of available experimental capacities. Notably, this strategy allows for full flexibility with respect to the number of parallel threads, as well as throughputs and instrument down-times, thereby paving the path towards global democratized experimentation.

Asynchronous Closed-Loop Optimization of OSL Gain Materials

With the overall synthesis, characterization and experiment planning engine in hand, a two-month optimization campaign for OSL gain materials was carried out, starting from the available seed data. Synthesis of the cap–bridge–core–bridge–cap pentamers was performed in the described multi-threaded, asynchronous fashion across three different robotic platforms at two different sites. Batches of target syntheses from the pool of recommended candidates were manually allocated to the robotic platforms as available, executed in an automated fashion, and subjected to the end-to-end characterization workflow.

Already from the first pool of recommendations, the Bayesian optimizer identified sets of compounds with state-of-the-art lasing performance (Fig. 4b). In fact, from the start of the optimization campaign, a total of 12 new compounds with higher solution gain cross section than the parent BSBCz were discovered, indicating accelerated discovery of high-gain materials through the inclusion of BO and automated experimentation. Structures and solution-state optical properties of selected molecules discovered in the course of this study are shown in Fig. 4c. These candidates represent the small-molecule emitters with the highest emission gain cross sections in solution known to date. Since the optimization campaign did not include any wavelength constraints, most OSL candidates were identified in the violet–blue region of the emission spectrum. Notably, in this wavelength range, our optimization campaign approaches the upper limit of the proxy's linear range,[49] i.e. the solution-state emission gain cross section at room temperature cannot be optimized much further due to physical constraints to emission lifetime and spectral width (see Supplementary Section 3.4 for a detailed discussion).

A visual analysis of the full dataset of potential OSL emitters (Fig. 4a) suggests the existence of a Pareto front between the emission color and the gain cross section within our current search space. This leaves room for further optimization campaigns with tailored building blocks, targeting the development of yellow OSL emitters, which have remained largely elusive, even for inorganic solid-state lasers.[50]

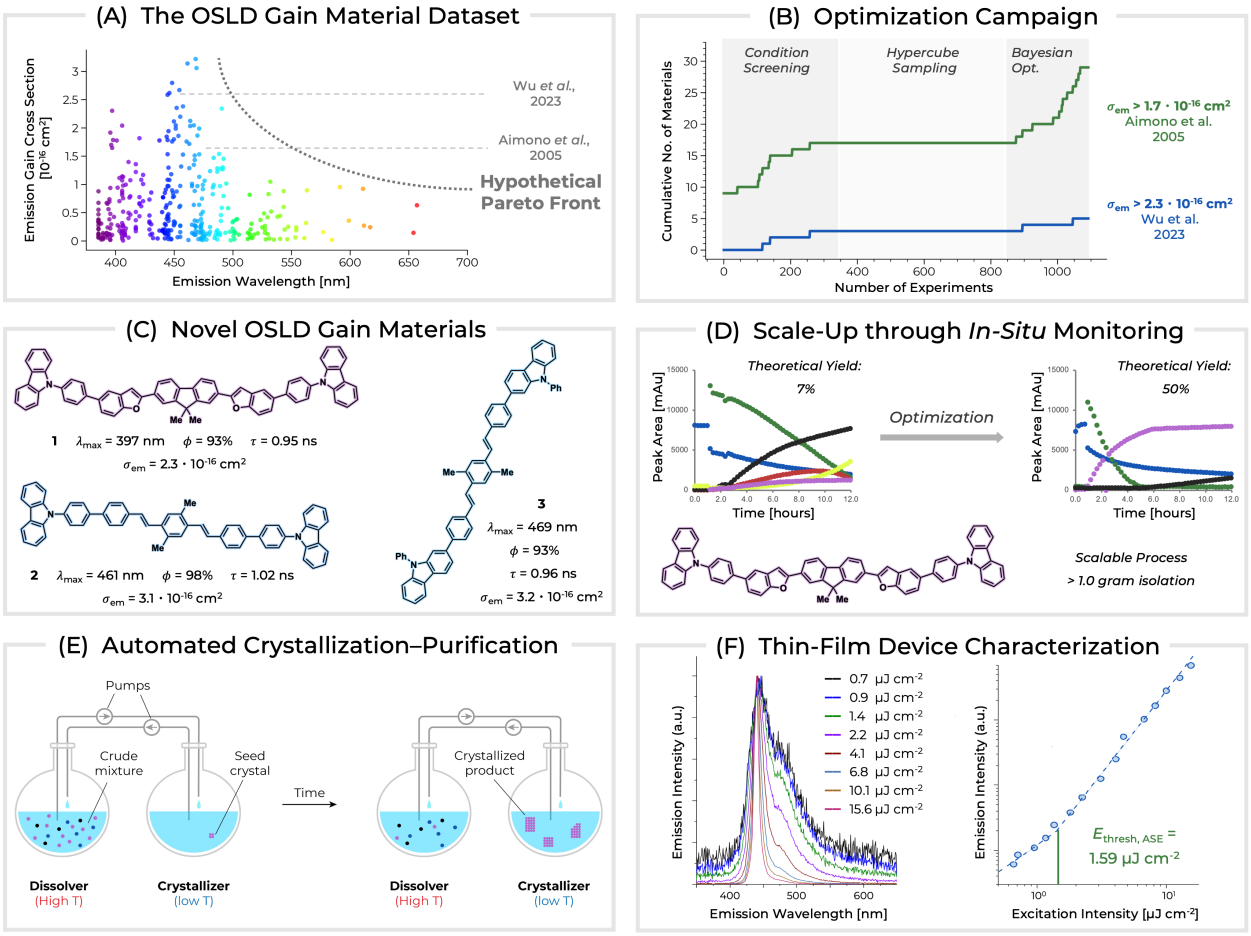

Figure 4: Discovered OSL gain materials. (**A**) Scatter plot of emission wavelength vs. emission gain cross section for all candidates the OSL gain materials dataset obtained throughout the optimization campaign. (**B**) Cumulative number of discovered materials with emission gain cross sections greater than BSBCz (green) and the top candidate from Wu *et al.* (blue). (**C**) Molecular structure and optical properties of selected OSL gain materials discovered in the course of the optimization campaign. (**D**) *In-situ* monitoring of reactant, intermediate and product concentrations for the synthesis of **1** by HPLC-MS as a function of time. Blue: Bridge building block, green: Core building block, red: mono-dehalogenation side product Br–[Core]–H, yellow: bis-dehalogenation side product H–[Core]–H, black: proto-deboronation product Cap–Bridge–H homocoupling product, violet: product **1**. (**E**) Schematic depiction of the automated crystallization module for scalable purification of OSL emitter materials. (**F**) Thin-film spectral data of **3** (3 wt% in matrix of 4,4'-Bis(*N*-carbazolyl)-1,1'-biphenyl). Emission spectra (left) and emission intensity (right) as a function of excitation intensity, demonstrating amplified spontaneous emission.

Evaluating Candidates in Thin-Film Devices

In order to evaluate the identified lasing candidates in an actual OSL device, synthesis and isolation of larger materials quantities at high purity are required. In the static two-step one-pot synthesis protocol used throughout the optimization campaign, reaction yields for different target molecules can vary substantially. Real-time tracking of starting material and reaction intermediate quantities allows for more precise control of the reaction outcome, particularly for a two-step protocol, in which competing reactions and intermediate decomposition can be minimized by adjusting the timing at which reactants and reagents are added, or reactions are terminated. Against this background, candidate compounds **1–3** were prepared on scale.[51] At the example of compound **1**, tracking intermediate concentrations by means of on-line HPLC-MS (Fig. 4d)[52,53] revealed reductive dehalogenation and protodeboronation as major side reactions, which could be suppressed by adjusting the reaction duration in real-time, and adding further XPhos ligand in the second coupling step. Using these specifically adapted conditions allowed for the isolation of **1** on a larger scale (460 mg, 63% yield).

At the same time, purification of the obtained materials evolved as a second major challenge, particularly when high purity (>99.5%) is required for device fabrication. In the case of our OSL material candidates, well established chromatography or sublimation protocols are hampered by poor solubility and decomposition upon sublimation, respectively. For this purpose, we developed a module for continuous preferential crystallization purification (CPC, Fig. 4e, see Supplementary Section 3.8 for further details)[54,55]. Using this automated workflow, gram-scale quantities of **1, 2** and **3** could be obtained, which were used for the preparation of thin-film devices.

Eventually, device-level properties were determined by spin-coating thin films of **1–3** (3 wt%) in a matrix of 4,4'-Bis(*N*-carbazolyl)-1,1'-biphenyl (CBP), and subsequent spectroscopic characterization. For all materials, amplified spontaneous emission (ASE) could be observed, and low ASE thresholds of 1.5–1.7 μJ cm^{-2} were determined (Fig. 4f and Supplementary Fig. S56–S63). Notably, material **3**, initially identified as the highest-gain emitter in solution, exhibited the lowest ASE threshold in thin film, and significantly outperformed a solution-processable BSBCz derivative used as a reference material (E_{th} = 1.71 μJ cm^{-2}).[32] This underscores its best-in-class lasing performance, and highlights the effectiveness of our proxy-based materials discovery workflow. In addition, it is worth noting that the ability to prepare thin-film devices through spin-coating represents a further advancement over the parent BSBCz, where thin films need to be prepared through vapor deposition.

Outlook

In summary, we demonstrated an asynchronous, delocalized discovery campaign for gain materials for OSL devices, integrating multiple automated synthesis and characterization modules across different laboratories and time zones with a central, cloud-based AI optimizer. Key to the success of the discovery campaign was the identification of a robust two-step, five-component, one-pot synthesis protocol for assembling functional targets from pre-fabricated building blocks. By optimizing a solution-state proxy for lasing performance, BO enabled the efficient navigation of a large virtual space of synthesizable OSL candidate compounds. Overall, this study resulted in the discovery of a library of 21 novel gain materials with state-of-the-art lasing performance. A set of optimized candidate molecules was prepared and purified on gram scale, and device-level performance confirms the identification of best-in-class gain materials (in terms of stimulated emission threshold).

Drawing from these findings, we envisage three major directions for next-generation workflows towards improved OSL devices: a) While our current synthesis module operates on a static set of conditions for building block assembly, adaptive treatment of synthesizability and substrate-dependent condition selection can lead to a significant improvement of reaction yields and robust quantification of molecular properties. b) Advanced proxy measurements, e.g. assessing optical properties in thin films rather than in solution, are required to provide a more realistic estimation of the lasing performance, taking into account important parameters such as solid packing or matrix effects. c) The systematic identification of sets of function-infused building blocks would enable encoding an optimized candidate space, enabling interpretability and hybrid human–AI molecular design.

Most importantly, our work demonstrates a blueprint for delocalized discovery campaigns. The integration of various synthesis and characterization modules delocalized over multiple sites across the globe enabled a complex discovery workflow, synergistically merging the capacities of automated and human-centric experimentation. At the heart of such a campaign must be an accessible, cloud-based platform for centralized data storage and AI experiment planning, to harness the full potential of automated experimentation and data-driven experiment design – while more rigorous and ideally automated policies for ensuring robustness and reproducibility should be enacted. Eventually, the abstraction of integrating multiple reproducible experimental modules to *fog* units, and their assembly into high-level *cloud* computing workflows, can provide a scalable framework for the democratization of (materials) discovery.[56]

References and Notes

(1) Zhavoronkov, A.; Ivanenkov, Y. A.; Aliper, A.; Veselov, M. S.; Aladinskiy, V. A.; Aladinskaya, A. V.; Terentiev, V. A.; Polykovskiy, D. A.; Kuznetsov, M. D.; Asadulaev, A.; Volkov, Y.; Zholus, A.; Shayakhmetov, R. R.; Zhebrak, A.; Minaeva, L. I.; Zagribelnyy, B. A.; Lee, L. H.; Soll, R.; Madge, D.; Xing, L.; Guo, T.; Aspuru-Guzik, A. Deep Learning Enables Rapid Identification of Potent DDR1 Kinase Inhibitors. *Nat. Biotechnol.* 2019, *37* (9), 1038–1040. https://doi.org/10.1038/s41587-019-0224-x.

(2) Uoyama, H.; Goushi, K.; Shizu, K.; Nomura, H.; Adachi, C. Highly Efficient Organic Light-Emitting Diodes from Delayed Fluorescence. *Nature* 2012, *492* (7428), 234–238. https://doi.org/10.1038/nature11687.

(3) Larcher, D.; Tarascon, J.-M. Towards Greener and More Sustainable Batteries for Electrical Energy Storage. *Nat. Chem.* 2015, *7* (1), 19–29. https://doi.org/10.1038/nchem.2085.

(4) *The United Nations Sustainable Development Goals*. https://www.un.org/sustainabledevelopment/ (accessed 2023-07-04).

(5) Aspuru-Guzik, A.; Lindh, R.; Reiher, M. The Matter Simulation (R)Evolution. *ACS Cent. Sci.* 2018, *4* (2), 144–152. https://doi.org/10.1021/acscentsci.7b00550.

(6) Wang, H.; Fu, T.; Du, Y.; Gao, W.; Huang, K.; Liu, Z.; Chandak, P.; Liu, S.; Van Katwyk, P.; Deac, A.; Anandkumar, A.; Bergen, K.; Gomes, C. P.; Ho, S.; Kohli, P.; Lasenby, J.; Leskovec, J.; Liu, T.-Y.; Manrai, A.; Marks, D.; Ramsundar, B.; Song, L.; Sun, J.; Tang, J.; Veličković, P.; Welling, M.; Zhang, L.; Coley, C. W.; Bengio, Y.; Zitnik, M. Scientific Discovery in the Age of Artificial Intelligence. *Nature* 2023, *620* (7972), 47–60. https://doi.org/10.1038/s41586-023-06221-2.

(7) Zhang, X.; Wang, L.; Helwig, J.; Luo, Y.; Fu, C.; Xie, Y.; Liu, M.; Lin, Y.; Xu, Z.; Yan, K.; Adams, K.; Weiler, M.; Li, X.; Fu, T.; Wang, Y.; Yu, H.; Xie, Y.; Fu, X.; Strasser, A.; Xu, S.; Liu, Y.; Du, Y.; Saxton, A.; Ling, H.; Lawrence, H.; Stärk, H.; Gui, S.; Edwards, C.; Gao, N.; Ladera, A.; Wu, T.; Hofgard, E. F.; Tehrani, A. M.; Wang, R.; Daigavane, A.; Bohde, M.; Kurtin, J.; Huang, Q.; Phung, T.; Xu, M.; Joshi, C. K.; Mathis, S. V.; Azizzadenesheli, K.; Fang, A.; Aspuru-Guzik, A.; Bekkers, E.; Bronstein, M.; Zitnik, M.; Anandkumar, A.; Ermon, S.; Liò, P.; Yu, R.; Günnemann, S.; Leskovec, J.; Ji, H.; Sun, J.; Barzilay, R.; Jaakkola, T.; Coley, C. W.; Qian, X.; Qian, X.; Smidt, T.; Ji, S. Artificial Intelligence for Science in Quantum, Atomistic, and Continuum Systems. *arXiv* 2023. https://doi.org/10.48550/arXiv.2307.08423.

(8) Macarron, R.; Banks, M. N.; Bojanic, D.; Burns, D. J.; Cirovic, D. A.; Garyantes, T.; Green, D. V. S.; Hertzberg, R. P.; Janzen, W. P.; Paslay, J. W.; Schopfer, U.; Sittampalam, G. S. Impact of High-Throughput Screening in Biomedical Research. *Nat. Rev. Drug Discov.* 2011, *10* (3), 188–195. https://doi.org/10.1038/nrd3368.

(9) Buitrago Santanilla, A.; Regalado, E. L.; Pereira, T.; Shevlin, M.; Bateman, K.; Campeau, L.-C.; Schneeweis, J.; Berritt, S.; Shi, Z.-C.; Nantermet, P.; Liu, Y.; Helmy, R.; Welch, C. J.; Vachal, P.; Davies, I. W.; Cernak, T.; Dreher, S. D. Nanomole-Scale High-Throughput Chemistry for the Synthesis of Complex Molecules. *Science* 2015, *347* (6217), 49–53. https://doi.org/10.1126/science.1259203.

(10) Häse, F.; Roch, L. M.; Aspuru-Guzik, A. Next-Generation Experimentation with Self-Driving Laboratories. *Trends Chem.* 2019, *1* (3), 282–291. https://doi.org/10.1016/j.trechm.2019.02.007.

(11) Stach, E.; DeCost, B.; Kusne, A. G.; Hattrick-Simpers, J.; Brown, K. A.; Reyes, K. G.; Schrier, J.; Billinge, S.; Buonassisi, T.; Foster, I.; Gomes, C. P.; Gregoire, J. M.; Mehta, A.; Montoya, J.; Olivetti, E.; Park, C.; Rotenberg, E.; Saikin, S. K.; Smullin, S.; Stanev, V.; Maruyama, B. Autonomous Experimentation Systems for Materials Development: A Community Perspective. *Matter* 2021, *4* (9), 2702–2726. https://doi.org/10.1016/j.matt.2021.06.036.

(12) Seifrid, M.; Pollice, R.; Aguilar-Granda, A.; Morgan Chan, Z.; Hotta, K.; Ser, C. T.; Vestfrid, J.; Wu, T. C.; Aspuru-Guzik, A. Autonomous Chemical Experiments: Challenges and Perspectives on Establishing a Self-Driving Lab. *Acc. Chem. Res.* 2022, *55* (17), 2454–2466. https://doi.org/10.1021/acs.accounts.2c00220.

(13) Abolhasani, M.; Kumacheva, E. The Rise of Self-Driving Labs in Chemical and Materials Sciences. *Nat. Synth.* 2023, *2* (6), 483–492. https://doi.org/10.1038/s44160-022-00231-0.

(14) Bédard, A.-C.; Adamo, A.; Aroh, K. C.; Russell, M. G.; Bedermann, A. A.; Torosian, J.; Yue, B.; Jensen, K. F.; Jamison, T. F. Reconfigurable System for Automated Optimization of Diverse Chemical Reactions. *Science* 2018, *361* (6408), 1220–1225. https://doi.org/10.1126/science.aat0650.

(15) Christensen, M.; Yunker, L. P. E.; Adedeji, F.; Häse, F.; Roch, L. M.; Gensch, T.; dos Passos Gomes, G.; Zepel, T.; Sigman, M. S.; Aspuru-Guzik, A.; Hein, J. E. Data-Science Driven Autonomous Process Optimization. *Commun. Chem.* 2021, *4* (1), 1–12. https://doi.org/10.1038/s42004-021-00550-x.

(16) Angello, N. H.; Rathore, V.; Beker, W.; Wołos, A.; Jira, E. R.; Roszak, R.; Wu, T. C.; Schroeder, C. M.; Aspuru-Guzik, A.; Grzybowski, B. A.; Burke, M. D. Closed-Loop Optimization of General Reaction Conditions for Heteroaryl Suzuki-Miyaura Coupling. *Science* 2022, *378* (6618), 399–405. https://doi.org/10.1126/science.adc8743.

(17) MacLeod, B. P.; Parlane, F. G. L.; Morrissey, T. D.; Häse, F.; Roch, L. M.; Dettelbach, K. E.; Moreira, R.; Yunker, L. P. E.; Rooney, M. B.; Deeth, J. R.; Lai, V.; Ng, G. J.; Situ, H.; Zhang, R. H.; Elliott, M. S.; Haley, T. H.; Dvorak, D. J.; Aspuru-Guzik, A.; Hein, J. E.; Berlinguette, C. P. Self-Driving Laboratory for Accelerated Discovery of Thin-Film Materials. *Sci. Adv.* 2020, *6* (20), eaaz8867. https://doi.org/10.1126/sciadv.aaz8867.

(18) Caramelli, D.; Salley, D.; Henson, A.; Camarasa, G. A.; Sharabi, S.; Keenan, G.; Cronin, L. Networking Chemical Robots for Reaction Multitasking. *Nat. Commun.* 2018, *9* (1), 3406. https://doi.org/10.1038/s41467-018-05828-8.

(19) Roch, L. M.; Häse, F.; Kreisbeck, C.; Tamayo-Mendoza, T.; Yunker, L. P. E.; Hein, J. E.; Aspuru-Guzik, A. ChemOS: An Orchestration Software to Democratize Autonomous Discovery. *PLOS ONE* 2020, *15* (4), e0229862. https://doi.org/10.1371/journal.pone.0229862.

(20) Bai, J.; Cao, L.; Mosbach, S.; Akroyd, J.; Lapkin, A. A.; Kraft, M. From Platform to Knowledge Graph: Evolution of Laboratory Automation. *JACS Au* 2022, *2* (2), 292–309. https://doi.org/10.1021/jacsau.1c00438.

(21) Vogler, M.; Busk, J.; Hajiyani, H.; Jørgensen, P. B.; Safaei, N.; Castelli, I. E.; Ramirez, F. F.; Carlsson, J.; Pizzi, G.; Clark, S.; Hanke, F.; Bhowmik, A.; Stein, H. S. Brokering between Tenants for an International Materials Acceleration Platform. *Matter* 2023. https://doi.org/10.1016/j.matt.2023.07.016.

(22) Merrifield, R. B. Automated Synthesis of Peptides. *Science* 1965, *150* (3693), 178–185. https://doi.org/10.1126/science.150.3693.178.

(23) Plante, O. J.; Palmacci, E. R.; Seeberger, P. H. Automated Solid-Phase Synthesis of Oligosaccharides. *Science* 2001, *291* (5508), 1523–1527. https://doi.org/10.1126/science.1057324.

(24) Caruthers, M. H. Gene Synthesis Machines: DNA Chemistry and Its Uses. *Science* 1985, *230* (4723), 281–285. https://doi.org/10.1126/science.3863253.

(25) Gillis, E. P.; Burke, M. D. A Simple and Modular Strategy for Small Molecule Synthesis: Iterative Suzuki−Miyaura Coupling of B-Protected Haloboronic Acid Building Blocks. *J. Am. Chem. Soc.* 2007, *129* (21), 6716–6717. https://doi.org/10.1021/ja0716204.

(26) Li, J.; Ballmer, S. G.; Gillis, E. P.; Fujii, S.; Schmidt, M. J.; Palazzolo, A. M. E.; Lehmann, J. W.; Morehouse, G. F.; Burke, M. D. Synthesis of Many Different Types of Organic Small Molecules Using One Automated Process. *Science* 2015, *347* (6227), 1221–1226. https://doi.org/10.1126/science.aaa5414.

(27) Clark, J.; Lanzani, G. Organic Photonics for Communications. *Nat. Photonics* 2010, *4* (7), 438–446. https://doi.org/10.1038/nphoton.2010.160.

(28) Kuehne, A. J. C.; Gather, M. C. Organic Lasers: Recent Developments on Materials, Device Geometries, and Fabrication Techniques. *Chem. Rev.* 2016, *116* (21), 12823–12864. https://doi.org/10.1021/acs.chemrev.6b00172.

(29) Zhang, Q.; Tao, W.; Huang, J.; Xia, R.; Cabanillas-Gonzalez, J. Toward Electrically Pumped Organic Lasers: A Review and Outlook on Material Developments and Resonator Architectures. *Adv. Photonics Res.* 2021, *2* (5), 2000155. https://doi.org/10.1002/adpr.202000155.

(30) Aimono, T.; Kawamura, Y.; Goushi, K.; Yamamoto, H.; Sasabe, H.; Adachi, C. 100% Fluorescence Efficiency of 4,4′-Bis[(N-Carbazole)Styryl]Biphenyl in a Solid Film and the Very Low Amplified Spontaneous Emission Threshold. *Appl. Phys. Lett.* 2005, *86* (7), 071110. https://doi.org/10.1063/1.1867555.

(31) Sandanayaka, A. S. D.; Matsushima, T.; Bencheikh, F.; Terakawa, S.; Potscavage, W. J.; Qin, C.; Fujihara, T.; Goushi, K.; Ribierre, J.-C.; Adachi, C. Indication of Current-Injection Lasing from an Organic Semiconductor. *Appl. Phys. Express* 2019, *12* (6), 061010. https://doi.org/10.7567/1882-0786/ab1b90.

(32) Mamada, M.; Fukunaga, T.; Bencheikh, F.; Sandanayaka, A. S. D.; Adachi, C. Low Amplified Spontaneous Emission Threshold from Organic Dyes Based on Bis-Stilbene. *Adv. Funct. Mater.* 2018, *28* (32), 1802130. https://doi.org/10.1002/adfm.201802130.

(33) Oyama, Y.; Mamada, M.; Shukla, A.; Moore, E. G.; Lo, S.-C.; Namdas, E. B.; Adachi, C. Design Strategy for Robust Organic Semiconductor Laser Dyes. *ACS Mater. Lett.* 2020, *2* (2), 161–167. https://doi.org/10.1021/acsmaterialslett.9b00536.

(34) Wu, T. C.; Aguilar-Granda, A.; Hotta, K.; Yazdani, S. A.; Pollice, R.; Vestfrid, J.; Hao, H.; Lavigne, C.; Seifrid, M.; Angello, N.; Bencheikh, F.; Hein, J. E.; Burke, M.; Adachi, C.; Aspuru-Guzik, A. A Materials Acceleration Platform for Organic Laser Discovery. *Adv. Mater.* 2023, *35* (6), 2370042. https://doi.org/10.1002/adma.202370042.

(35) Ballmer, S. G.; Gillis, E. P.; Burke, M. D. B-Protected Haloboronic Acids for Iterative Cross-Coupling. In *Organic Syntheses*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2009; pp 344–359. https://doi.org/10.1002/0471264229.os086.33.

(36) Blair, D. J.; Chitti, S.; Trobe, M.; Kostyra, D. M.; Haley, H. M. S.; Hansen, R. L.; Ballmer, S. G.; Woods, T. J.; Wang, W.; Mubayi, V.; Schmidt, M. J.; Pipal, R. W.; Morehouse, G. F.; Palazzolo Ray, A. M. E.; Gray, D. L.; Gill, A. L.; Burke, M. D. Automated Iterative Csp3–C Bond Formation. *Nature* 2022, *604* (7904), 92–97. https://doi.org/10.1038/s41586-022-04491-w.

(37) Delaney, C. P.; Kassel, V. M.; Denmark, S. E. Potassium Trimethylsilanolate Enables Rapid, Homogeneous Suzuki−Miyaura Cross-Coupling of Boronic Esters. *ACS Catal.* 2020, *10* (1), 73–80. https://doi.org/10.1021/acscatal.9b04353.

(38) Wang, W.; Angello, N.; Blair, D.; Medine, K.; Tyrikos-Ergas, T.; Laporte, A.; Burke, M. Rapid Automated Iterative Small Molecule Synthesis. *ChemRxiv* 2023. https://doi.org/10.26434/chemrxiv-2023-qpf2x.

(39) Steiner, S.; Wolf, J.; Glatzel, S.; Andreou, A.; Granda, J. M.; Keenan, G.; Hinkley, T.; Aragon-Camarasa, G.; Kitson, P. J.; Angelone, D.; Cronin, L. Organic Synthesis in a Modular Robotic System Driven by a Chemical Programming Language. *Science* 2019, *363* (6423), eaav2211. https://doi.org/10.1126/science.aav2211.

(40) Mehr, S. H. M.; Craven, M.; Leonov, A. I.; Keenan, G.; Cronin, L. A Universal System for Digitization and Automatic Execution of the Chemical Synthesis Literature. *Science* 2020, *370* (6512), 101–108. https://doi.org/10.1126/science.abc2986.

(41) *Molar – A Python Database Management System for PostgreSQL.* https://molar.readthedocs.io/en/latest/.

(42) Deshpande, A. V.; Beidoun, A.; Penzkofer, A.; Wagenblast, G. Absorption and Emission Spectroscopic Investigation of Cyanovinyldiethylaniline Dye Vapors. *Chem. Phys.* 1990, *142* (1), 123–131. https://doi.org/10.1016/0301-0104(90)89075-2.

(43) Nakanotani, H.; Adachi, C.; Watanabe, S.; Katoh, R. Spectrally Narrow Emission from Organic Films under Continuous-Wave Excitation. *Appl. Phys. Lett.* 2007, *90* (23), 231109. https://doi.org/10.1063/1.2746958.

(44) Kushner, H. J. A New Method of Locating the Maximum Point of an Arbitrary Multipeak Curve in the Presence of Noise. *J. Basic Eng.* 1964, *86* (1), 97–106. https://doi.org/10.1115/1.3653121.

(45) Snoek, J.; Larochelle, H.; Adams, R. P. Practical Bayesian Optimization of Machine Learning Algorithms. *arXiv* 2012. https://doi.org/10.48550/arXiv.1206.2944.

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

(47) Huang, B.; von Rudorff, G. F.; von Lilienfeld, O. A. The Central Role of Density Functional Theory in the AI Age. *Science* 2023, *381* (6654), 170–175. https://doi.org/10.1126/science.abn3445.

(48) Baiardi, A.; Bloino, J.; Barone, V. General Time Dependent Approach to Vibronic Spectroscopy Including Franck–Condon, Herzberg–Teller, and Duschinsky Effects. *J. Chem. Theory Comput.* 2013, *9* (9), 4097–4115. https://doi.org/10.1021/ct400450k.

(49) Lyu, J.; Wang, S.; Balius, T. E.; Singh, I.; Levit, A.; Moroz, Y. S.; O'Meara, M. J.; Che, T.; Algaa, E.; Tolmachova, K.; Tolmachev, A. A.; Shoichet, B. K.; Roth, B. L.; Irwin, J. J. Ultra-Large Library Docking for Discovering New Chemotypes. *Nature* 2019, *566* (7743), 224–229. https://doi.org/10.1038/s41586-019-0917-9.

(50) Castellano-Hernández, E.; Metz, P. W.; Demesh, M.; Kränkel, C. Efficient Directly Emitting High-Power Tb^{3+}:$LiLuF_4$ Laser Operating at 587.5 Nm in the Yellow Range. *Opt. Lett.* 2018, *43* (19), 4791–4794. https://doi.org/10.1364/OL.43.004791.

(51) Materials 2 and 3 Were Synthesized Manually (See Supplementary Section 3.6 for Further Details).

(52) Malig, T. C.; Yunker, L. P. E.; Steiner, S.; Hein, J. E. Online High-Performance Liquid Chromatography Analysis of Buchwald–Hartwig Aminations from within an Inert Environment. *ACS Catal.* 2020, *10* (22), 13236–13244. https://doi.org/10.1021/acscatal.0c03530.

(53) Sato, Y.; Liu, J.; Kukor, A. J.; Culhane, J. C.; Tucker, J. L.; Kucera, D. J.; Cochran, B. M.; Hein, J. E. Real-Time Monitoring of Solid–Liquid Slurries: Optimized Synthesis of Tetrabenazine. *J. Org. Chem.* 2021, *86* (20), 14069–14078. https://doi.org/10.1021/acs.joc.1c01098.

(54) Hein, J. E.; Cao, B. H.; Van Der Meijden, M. W.; Leeman, M.; Kellogg, R. M. Resolution of Omeprazole Using Coupled Preferential Crystallization: Efficient Separation of a Nonracemizable Conglomerate Salt under Near-Equilibrium Conditions. *Org. Process Res. Dev.* 2013, *17* (6), 946–950. https://doi.org/10.1021/op400081c.

(55) Rougeot, C.; Hein, J. E. Application of Continuous Preferential Crystallization to Efficiently Access Enantiopure Chemicals. *Org. Process Res. Dev.* 2015, *19* (12), 1809–1819. https://doi.org/10.1021/acs.oprd.5b00141.

(56) Sim, M.; Ghazi Vakili, M.; Strieth-Kalthoff, F.; Hao, H.; Hickman, R.; Miret, S.; Pablo-García, S.; Aspuru-Guzik, A. ChemOS 2.0: An Orchestration Architecture for Chemical Self-Driving Laboratories. *ChemRxiv* 2023. https://doi.org/10.26434/chemrxiv-2023-v2khf.

Data and Materials Availability

All data and code generated as part of this study are openly accessible either in the supplementary materials or in open repositories. Raw characterization data (NMR spectra of all building blocks and scaled-up materials, raw HPLC-MS data in open-source format) are available at Zenodo (https://doi.org/10.5281/zenodo.8357283). Synthesis and optical spectroscopy results for all target compounds have been deposited at Zenodo. Results of the high-throughput computational analysis are available at Zenodo. All training data for machine learning models are deposited at Zenodo. All software utilized in this work is freely available on *GitHub* (https://github.com/aspuru-guzik-group/acdc_laser.git), and a snapshot of the code has been deposited at Zenodo (https://doi.org/10.5281/zenodo.8357375).

Acknowledgments and Funding

The authors thank Dr. Anne Fischer for conceiving the DARPA Accelerated Molecular Discovery Program and for numerous fruitful discussions. Dr. Rachel Keunen (University of Toronto) is acknowledged for experimental and administrative assistance. Computations were performed on the *Niagara* supercomputer (SciNet HPC consortium; Canada Foundation for Innovation; Government of Ontario; Ontario Research Fund – Research Excellence; University of Toronto), the *Cedar* supercomputer (WestGrid consortium; Digital Research Alliance of Canada), and the *Beluga* and *Narval* supercomputers (École de technologie supérieure, Calcul Québec; Digital Research Alliance of Canada; Canada Foundation for Innovation; Ministère de l'Économie, des Sciences et de l'innovation du Québec; Fonds de recherche du Québec – Nature et technologies).

The authors acknowledge the Defense Advanced Research Projects Agency (DARPA) under the Accelerated Molecular Discovery Program under Cooperative Agreement No. HR00111920027 dated August 1, 2019. The content of the information presented in this work does not necessarily reflect the position or the policy of the Government. This research was undertaken thanks in part to funding provided to the University of Toronto's Acceleration Consortium from the Canada First Research Excellence Fund CFREF-2022-00042. FSK is a postdoctoral fellow in the Eric and Wendy Schmidt AI in Science Postdoctoral Fellowship Program, a program by Schmidt Futures. CA thanks the Japan Science and Technology Agency (JST) CREST (grant no. JPMJCR22B3) Specially Promoted Research (grant no. 23H05406). LC and the Glasgow team thank EPSRC (grant nos. EP/L023652/1, EP/R01308X/1, EP/S019472/1, and EP/P00153X/1), and the ERC (project 670467 SMART-POM). LC thanks Schmidt Futures for an Innovation Fellowship. AAG. thanks Anders G. Frøseth for his generous support. AAG also acknowledges the generous support of Natural Resources Canada and the Canada 150 Research Chairs program.

Competing Interests

L.C. is a founder of Chemify Ltd. J.E.H is the founder and CEO of Telescope Innovations, an enabling technologies startup located in Vancouver, BC, Canada. A.A.-G. is chief visionary officer and board member of Kebotix Inc., a company that carries out closed-loop molecular materials discovery.

Author Contributions (CRediT)

Conceptualization:	F.S.-K., H.H., V.R., J.D., T.G., N.H.A., M.S., R.P., T.C.W., C.A., B.A.G., L.C., J.E.H., M.D.B., A.A.-G.
Data curation:	F.S.-K., H.H., V.R., J.D., N.H.A., M.S., E.T., M.G., J.L., X.T., M.M., W.W., T.T., L.B., A.W., R.R., C.-T.S., L.C., J.E.H.
Formal analysis:	F.S.-K., H.H., V.R., J.D., N.H.A., M.S., E.T., M.G., J.L., X.T., M.M., T.T., A.W., R.R., C.-T.S.
Funding acquisition:	C.A., B.A.G., L.C., J.E.H., M.D.B., A.A.-G.
Investigation:	F.S.-K., H.H., V.R., J.D., T.G., N.H.A., M.S., E.T., M.G., J.L., X.T., M.M., W.W., T.T., L.B., A.W., R.R., C.-T.S., L.C., J.E.H.
Methodology:	F.S.-K., H.H., V.R., J.D., T.G., N.H.A., M.S., E.T., M.G., J.L., X.T., M.M., W.W., C.L., R.P., T.W., K.H., M.H., A.W., R.R., C.-T.S., C.B.-G, R.J.H., J.V., A.A., E.L.K., R.C.S., W.H., D.G., S.L., A.W., O.B., S.R., B.S.-L., C.A., B.A.G., L.C., J.E.H., M.D.B., A.A.-G.
Project administration:	F.S.-K., H.H., V.R., J.D., N.H.A., M.S., R.P., T.C.W., C.A., B.A.G., L.C., J.E.H., M.D.B., A.A.-G.
Resources:	C.A., B.A.G., L.C., J.E.H., M.D.B., A.A.-G.
Software:	F.S.-K., H.H., T.G., M.S., C.L., R.P., T.C.W., K.H., M.H., A.W., R.R., C.-T.S., D.G.
Supervision:	C.A., B.A.G., L.C., J.E.H., M.D.B., A.A.-G.
Validation:	F.S.-K., H.H., V.R., J.D., N.H.A., M.S., E.T., M.G., J.L., W.W., L.B., S.L., C.-T.S., C.B.-G.
Visualization:	F.S.-K., T.G.

Writing (Original Draft): F.S.-K., H.H., C.A., B.A.G., L.C., J.E.H., M.D.B, A.A.-G.

Writing (Review and Editing): *all authors*

A.A. refers to A. Aguilar-Gránda, whereas A.A.-G. refers to A. Aspuru-Guzik.