RESEARCH ARTICLE SUMMARY

ORGANIC CHEMISTRY

Delocalized, asynchronous, closed-loop discovery of organic laser emitters

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

INTRODUCTION: Modern materials discovery necessitates intricate workflows that integrate synthesis, property characterization, formulation, and system-level testing. Often, the required expertise and research infrastructure are dispersed across multiple locations and time zones, which hinders their integration into advanced discovery pipelines. This challenge is particularly pronounced in the context of automated experimentation and data-driven decision-making with artificial intelligence (AI), which usually rely on streamlined flows of data and materials at a single site.

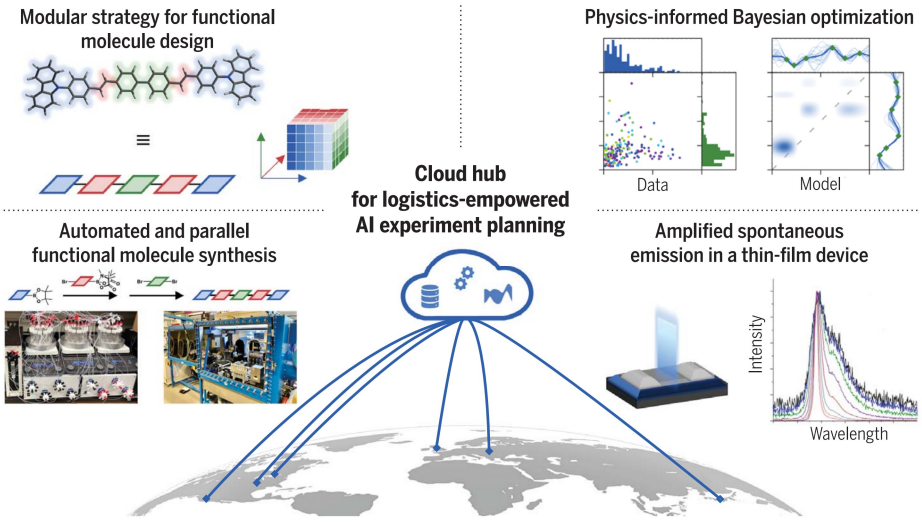
RATIONALE: Synergistically integrating experimental research infrastructure at geographically distributed sites requires a central, globally accessible cloud hub for data transfer; AI-informed experiment design; and logistics management. Such a decentralized engine can effectively orchestrate global design-make-test-analyze loops for materials discovery that would be impractical to establish in a single laboratory.

RESULTS: We demonstrate the distributed, AI-guided discovery of best-in-class small-molecule gain materials for organic solid-state lasers (OSLs). To overcome the prevalent synthesis bottleneck in molecular discovery, we devised a building block–based strategy to rapidly construct molecular function. Leveraging iterative Suzuki–Miyaura couplings, we developed a generalizable two-step one-pot protocol for assembling pentameric OSL gain materials from modular precursors, spanning a candidate space of >150,000 target materials. The preparation of building blocks was distributed over the available experimental resources at four geographic locations. Similarly, the assembly of building blocks to form the gain material candidates was parallelized and automated on different robotic synthesis platforms. This multisite synthesis engine was complemented by a resource-efficient, end-to-end automated testing workflow, encompassing (i) purification to a degree that allows for reliable spectroscopic measurements and (ii) solution-phase characterization through steady-state and time-resolved spectroscopy.

The resulting laser-like property data, asynchronously obtained from the different "workers," fed into the central, machine learning–based experiment planning module hosted on the cloud hub. This module was supplemented with physical knowledge from quantum-chemical simulations and logistic constraints from multisite experimentation. Taking into account what could be done, what was being done, and what had been done at any point in time, the module maintained a ranked catalog of most-informative next experiments—from which the following sets of experiments were allocated to the available robotic platforms at the different sites.

Throughout the development and operation of this multisite discovery engine, we found a total of 21 small-molecule emitters with improved emission gain cross sections compared with state-of-the-art OSL gain materials. For their property evaluation on a device level, on-scale synthesis and purification introduced further challenges, necessitating the development of automated workflows for synthesis with on-line decision-making and purification through continuous preferential crystallization. Ultimately, laser-like properties of three identified materials were measured in thin films, confirming the discovery of gain materials with best-in-class amplified spontaneous emission thresholds.

CONCLUSION: Addressing a frontier challenge in molecular optoelectronics, this work demonstrates a blueprint for decentralized (molecular) discovery in the age of AI. A central, readily accessible hub that manages data transfer, the logistics of materials transport and availability, and the experimental design process has proven to be of central importance. Eventually, we are convinced that scaling such a framework to flexibly include distributed human and robotic research resources can pave the way for democratizing (materials) discovery. ■



Overview of the described decentralized workflow for the accelerated discovery of small-molecule gain materials for OSLs.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: 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.)
†These authors contributed equally to this work.
Cite this article as F. Strieth-Kalthoff *et al.*, *Science* **384**, eadk9227 (2024). DOI: 10.1126/science.adk9227

RESEARCH ARTICLE

ORGANIC CHEMISTRY

Delocalized, asynchronous, closed-loop discovery of organic laser emitters

Felix Strieth-Kalthoff[1,2]†, Han Hao[1,2,3]†, Vandana Rathore[4,5], Joshua Derasp[6], Théophile Gaudin[2], Nicholas H. Angello[4,5,7], Martin Seifrid[1,2,8], Ekaterina Trushina[9], Mason Guy[6], Junliang Liu[6], Xun Tang[10], Masashi Mamada[10], Wesley Wang[4,5,7], Tuul Tsagaantsooj[10], Cyrille Lavigne[1,2], Robert Pollice[1,2], Tony C. Wu[1,2], Kazuhiro Hotta[1,2,11], Leticia Bodo[1], Shangyu Li[1], Mohammad Haddadnia[1,12], Agnieszka Wołos[13,14], Rafał Roszak[13,14], Cher Tian Ser[1,2], Carlota Bozal-Ginesta[1,2,15], Riley J. Hickman[1,2], Jenya Vestfrid[1,2], Andrés Aguilar-Granda[1,2], Elena L. Klimareva[9], Ralph C. Sigerson[9], Wenduan Hou[9], Daniel Gahler[9], Slawomir Lach[9], Adrian Warzybok[9,16], Oleg Borodin[9], Simon Rohrbach[9], Benjamin Sanchez-Lengeling[17], Chihaya Adachi[10]*, Bartosz A. Grzybowski[14,18,19]*, Leroy Cronin[3,9]*, Jason E. Hein[3,6,20]*, Martin D. Burke[3,4,5,7,21,22]*, Alán Aspuru-Guzik[1,2,3,12,23,24,25]*

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 enabled delocalized and asynchronous design-make-test-analyze cycles. We showcased 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 artificial intelligence experiment planner, resulted in the discovery of 21 new state-of-the-art materials. 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.

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 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*). Early examples have indicated the ability to substantially accelerate narrow tasks within these high-level workflows, such as the optimization of reaction conditions or the identification of ideal processing or formulation parameters (*14*–*17*).

The evolution of the second generation of AI-guided experimentation is driven by the principles of distribution and delocalization, acting as major paradigms in two distinct directions. First, with the growing complexity of discovery workflows, the integration of advanced experimental and computational modules becomes essential. These units often rely on domain expertise and specific instrumentation, resulting in their distribution over multiple geographical locations and time zones (Fig. 1A). Second, the capacity to parallelize experiments over multiple modules offers solutions to enhance throughput and rapidly validate experimental findings, as showcased, for example, for chemical reaction optimization with multiple (distributed) workers (*18*–*20*). Both scenarios—the integration of multiple modules into intricate experimental workflows and the parallelization of experimental tasks across multiple sites—rely on distributed experimentation as the key factor for accelerated materials discovery. However, the implementation of distributed experimentation necessitates a central, readily accessible platform with clearly defined standards for communication, data transfer, and experiment planning (*21*–*24*). This platform must also be adaptable to integrate with the specific preexisting infrastructures at each site, to account for the logistics of interdependent experiments, and to accommodate asynchrony between sites (Fig. 1B) to minimize disruptions and enhance overall efficiency. The discussed aspects hold particular significance in expediting molecular materials discovery, where (automated) synthesis remains the primary 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 (*25*), oligosaccharides (*26*), or oligonucleotides (*27*)—impeding broad molecular materials discovery endeavors.

Against this background, we demonstrated a decentralized discovery workflow, showcasing the design, synthesis, and testing of gain materials for organic solid-state lasers (OSLs), which are characterized by best-in-class emission gain cross section in solution and amplified spontaneous emission (ASE) in thin film. The workflow relied on a central closed-loop protocol encompassing synthesis planning, automated synthesis, proxy characterization, and molecular function optimization through machine learning (ML). Notably, the discussed synthesis bottleneck was bypassed by segmenting the OSL candidate space into a building block framework (*28*, *29*), which enabled rapid, parallelizable assembly of OSL gain material candidates following a "synthesis-to-function" paradigm. This closed-loop campaign was embedded in a broader discovery workflow, encompassing upstream and downstream operations for building block supply, on-scale materials fabrication, and device-level characterization. Although all tasks were delocalized across five physical laboratories on three continents, they were orchestrated by a cloud

[1]Department of Chemistry, University of Toronto, Toronto, ON, Canada. [2]Department of Computer Science, University of Toronto, Toronto, ON, Canada. [3]Acceleration Consortium, University of Toronto, Toronto, ON, Canada. [4]Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [5]Molecule Maker Lab, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [6]Department of Chemistry, University of British Columbia, Vancouver, BC, Canada. [7]Molecule Maker Lab Institute, Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [8]Department of Materials Science and Engineering, North Carolina State University, Raleigh, NC, USA. [9]School of Chemistry, University of Glasgow, Glasgow, UK. [10]Center for Organic Photonics and Electronics Research (OPERA), Kyushu University, Fukuoka, Japan. [11]Mitsubishi Chemical Corporation Science & Innovation Center, Kanagawa, Japan. [12]Vector Institute for Artificial Intelligence, Toronto, ON, Canada. [13]Allchemy Inc., Highland, IN, USA. [14]Institute of Organic Chemistry, Polish Academy of Sciences, Warsaw, Poland. [15]Catalonia Institute for Energy Research, Barcelona, Spain. [16]Department of Chemical Physics, Jagiellonian University, Krakow, Poland. [17]Google Research, Brain Team, Cambridge, MA, USA. [18]Center for Algorithmic and Robotized Synthesis, Institute for Basic Science, Ulsan, Republic of Korea. [19]Department of Chemistry, Ulsan Institute of Science and Technology, Ulsan, Republic of Korea. [20]Department of Chemistry, University of Bergen, Bergen, Norway. [21]Cancer Center at Illinois, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [22]Carle Illinois College of Medicine, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [23]Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON, Canada. [24]Department of Materials Science and Engineering, University of Toronto, Toronto, ON, Canada. [25]Canadian Institute for Advanced Research (CIFAR), Toronto, ON, Canada.
*Corresponding author. Email: 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.)
†These authors contributed equally to this work.

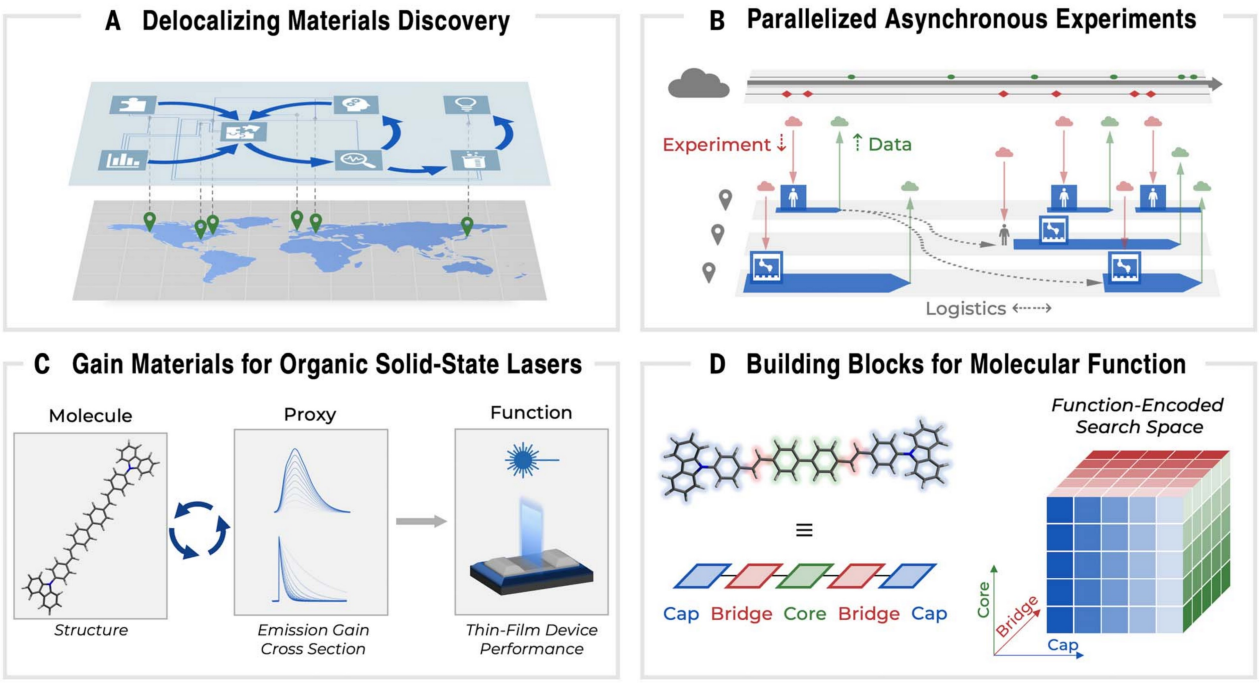


Fig. 1. Delocalizing molecular materials discovery targeting OSL emitters. (**A**) Delocalized complex materials discovery workflows (traditional design-make-test-analyze cycles) over multiple sites, orchestrated by a single cloud-based application. Markers on the map correspond to the geographical locations 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. Prerequisite for this is a cloud platform that can interoperate with existing (automated) research infrastructure in different laboratories while managing the logistics of asynchronous operation. (**C**) Schematic depiction of discovering gain materials for 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.

server to ensure continuous learning from the incoming data and effective 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 light fidelity (LiFi) telecommunication (Fig. 1C) (*30*–*33*). 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) (*34*, *35*). 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 (SMCs), which were developed by Burke and co-workers (*28*, *29*). With this general scheme, and by analogy to state-of-the-art emitters (*36*, *37*), 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 (Fig. 2A and data S1 to S3).

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 SMCs (*16*, *38*), 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 (*28*, *39*). This protocol showed 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. Across the overall candidate space, however, these conditions proved to be less effective (successful target compound detection in 32% of all cases; see table S7 for further details). Circumventing this limitation, the recently developed 2,2,2',2'-tetramethyl-*N*-methyliminodiacetic acid (TIDA) protecting group for boronic acids (*40*) allowed the use of potassium trimethylsilanolate (TMSOK) in the first coupling step (*41*, *42*), substantially 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 were recently developed through AI-guided optimization by Angello *et al.* (*16*). In an exploratory seed campaign across 500 representative candidate pentamers selected through Latin hypercube sampling (see supplementary text, section 3.5, for further details), this second-generation protocol led to a substantial increase in the global synthetic hit rate (74%) (Fig. 2, A and B).

The transition from MIDA to TIDA boronates as the bridge building blocks led to a shift of the initial building block space because

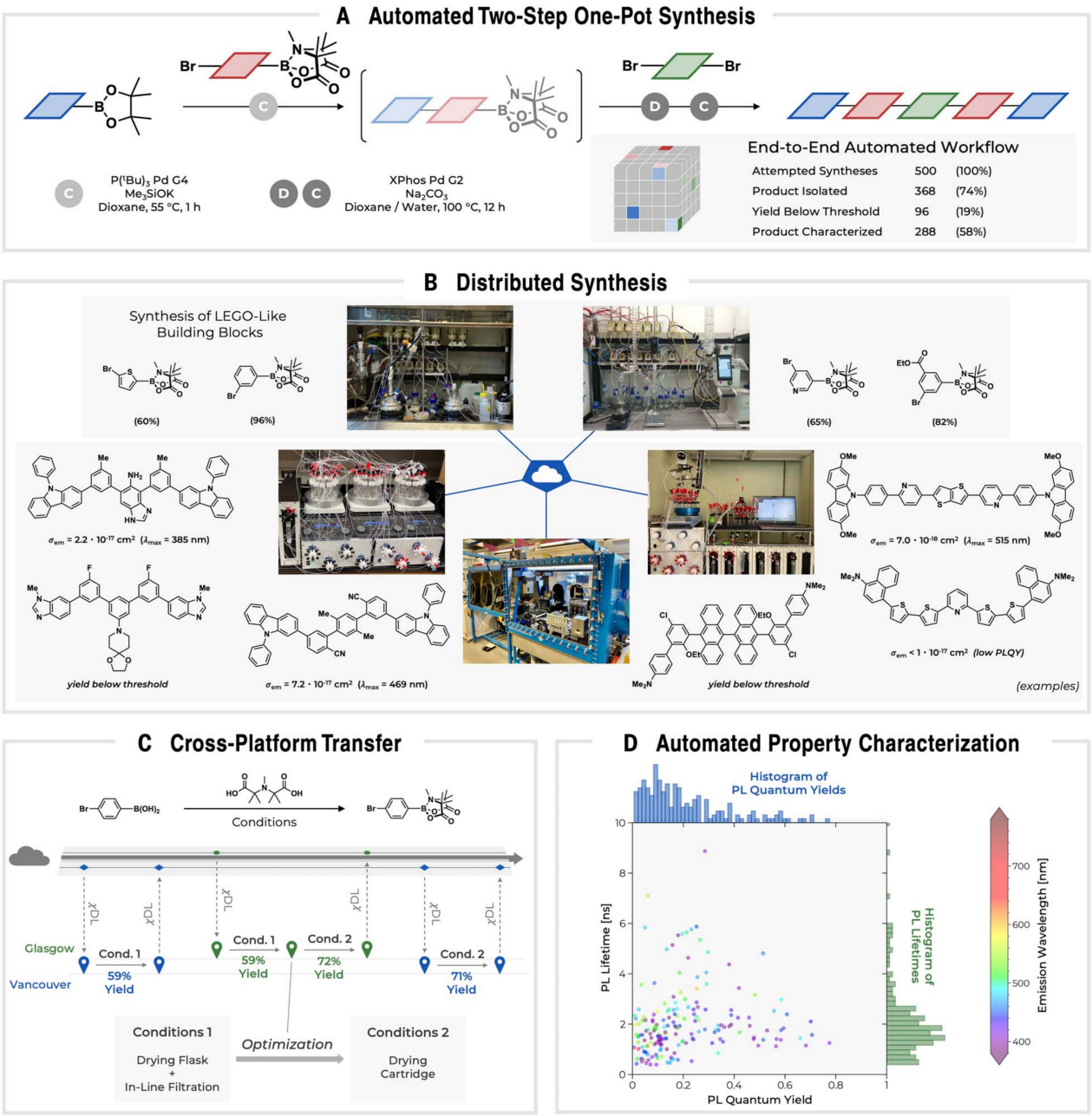


Fig. 2. Overview of the modules for automated synthesis and characterization of OSL candidate molecules. (**A**) Conditions of the iterative two-step one-pot SMC coupling for synthesizing pentameric structures and evaluation of the conditions on a representative subset of 500 target compounds, as obtained from Latin hypercube sampling (see supplementary text, 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). PLQY, photoluminescence quantum yield. (**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 from (A). Data points are colored according to their emission wavelength.

it required access to a library of bifunctional TIDA-protected haloboronic acids. Unlike their MIDA analogs, these compounds were not commercially available. The respective derivatization of commercially available precursors was rapidly performed across multiple laboratories, distributed through our central cloud hub (vide infra) and following general procedures encoded in the χDL language (*43*). In this process, we also explored the execution of these gram-scale synthesis and purification procedures in a fully automated fashion (Fig. 2B) (*44*). Notably, information transfer in the χDL format emphasizes the importance of standardization for reproducible synthesis. As shown in a recent study by Rauschen *et al.* (*45*), executing the identical χDL protocol on synthesis robots in different laboratories resulted in identical reaction yields, through the precise control over synthesis and purification operations. This laid the foundation for a rapidly verifiable and transferable global optimization of synthesis conditions (Fig. 2C) and the automated gram-scale synthesis of 10 building blocks.

From these presynthesized building blocks, small-scale parallel target syntheses were performed on three different automated platforms across two laboratories (Fig. 2B), again orchestrated from a single, readily accessible cloud hub (vide infra). Functional characterization of the obtained OSL target compounds was then performed from the crude reaction mixture through a resource-efficient automated analysis, purification, and characterization workflow, as reported previously by Wu *et al.* (*38*). Rather than subjecting each reaction to a time-consuming sequence of conventional isolation and characterization steps, materials were merely purified to an adequate degree by analytical high-performance liquid chromatography coupled to mass spectrometry (HPLC-MS), and the product fraction was subjected to downstream 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 text, section 3.3, for further details) (*46*, *47*). 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 workflow of automated synthesis, purification, and spectroscopic characterization represents a trade-off between throughput and accuracy, which leaves room for improvements through adaptive decision-making. Full spectroscopic data could only be obtained for 48% of all characterized compounds, whereas the remaining 52% of molecules were either emitters that were too weak (enabling only partial characterization) or had collected product fractions that were too low in concentration (Fig. 2A; see table S8 for further details). An overview of the obtained data across the exploratory seed dataset is given in Fig. 2D and fig. S9. Although a series of target molecules with short emission lifetimes (<1.5 ns) could be found, no state-of-the-art emitters were discovered in those data because the obtained quantum yields were predominantly very low.

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 the synthesis of new, improved OSL gain materials and navigating the search within the space of >150,000 potential target compounds. In the context of data-driven experiment planning, Bayesian optimization (BO) is regarded as the gold standard for the sample-efficient optimization of unknown spaces (*48*, *49*). Its application to efficient molecular discovery within a large virtual library necessitates (i) a vectorized encoding of the molecular structure for training an informed surrogate model (*20*) and (ii) a scheme to optimize recommended candidates across the discrete search space.

First, we evaluated a series of established structural representations (molecular fingerprints, graph-level descriptors, and computed building block descriptors) for the supervised learning of the 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 text, section 4.3, for a full evaluation of different surrogate model types). This observation 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 (*37*). 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).

To obtain a more expressive representation, we envisioned that additional physically meaningful information obtained from quantum chemistry simulations could improve the performance of our models (*50*). 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 text, section 4.2) (*51*). Including these simulated properties as molecular descriptors resulted in a slightly increased predictive performance. The most substantial enhancement for predicting the emission gain cross section, 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 simulated properties, achieved high accuracy [coefficient of determination (R^2) = 0.86 averaged over seven TD-DFT properties; for network architecture and prediction performance, see supplementary text, section 4.2]. Extracting the embedding vector from this GNN in a transfer learning approach led to improved regression and uncertainty calibration for the prediction of experimental emission gain cross sections using a Gaussian process (GP) regressor (Fig. 3A and tables S13 to S15). Notably, a correlation between pairwise embedding distances and experimentally observed molecular function differences was observed (Fig. 3B), which emphasizes the physical relevance of GNN embeddings as a functional molecular representation. This representation space, learned exclusively from simulated data, exhibited localized domains of high experimental lasing performance [visualized through uniform manifold approximation and projection (UMAP); Fig. 3C]. At the same time, it featured a range of activity cliffs—i.e., pairs of molecules that were close in embedding space but dissimilar in lasing performance (Fig. 3, B and C). The existence of such activity cliffs (*52*) leaves room for learning more globally informed representations and emphasizes the importance of an explorative search strategy for navigating the space of OSL candidates.

Asynchronous, closed-loop optimization of OSL gain materials

Eventually, driving the decentralized experimental engine with the developed ML models required a central hub for storing experimental data, overseeing the logistics of interdependent asynchronous experiments, and facilitating informed decision-making for future experiments. Positioned at the core of our workflow (Fig. 3E), this cloud-based hub was readily accessible for all teams through both a web and a software interface. Within the hub, the Bayesian optimizer, using the GNN-GP as the surrogate model, continuously maintained a ranked list of recommended experiments at all times, taking into account logistic constraints originating from building block availability and incomplete experiments.

From this list, top-priority experiments could be flexibly allocated to the available robotic platforms in Toronto and Urbana-Champaign, overseen by human researchers in the respective laboratories. Although in principle it is possible to automate the allocation of experiments entirely, we observed that including a human in the loop led to more-efficient management of experimental resources and dynamic laboratory-specific constraints. After completion of the automated synthesis, the crude samples were transferred (when necessary) and subjected to the end-to-end-automated characterization workflow. Eventually, the obtained characterization data underwent rapid, hybrid human-algorithmic quality control (QC) (see supplementary materials for further details) before being uploaded to the cloud-based hub. A detailed overview of which experiments were performed on which automated platform is provided in data S6.

It is important to note that throughout these processes, the status of an experiment (e.g., available, allocated, on-going, or completed) was continuously synchronized with the cloud hub. Any changes to the experimental database or the building block supply prompted an update of the list of suggested experiments. This dynamic update mechanism ensured continuous

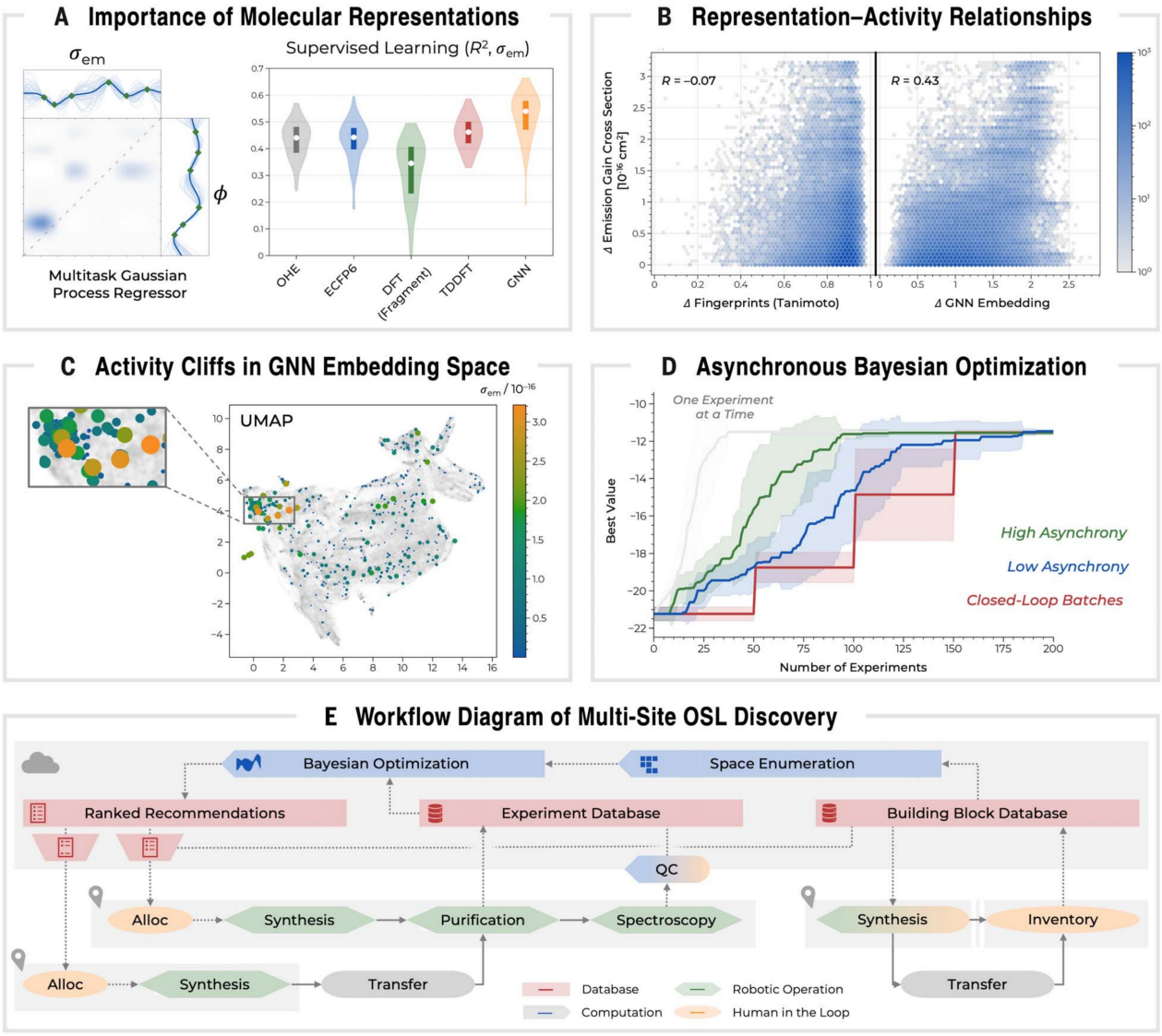


Fig. 3. BO for OSL gain material discovery. (**A**) Supervised learning performance of a multitask GP regressor for predicting emission gain cross section across the experimental seed dataset of 287 data points, comparing different molecular representations. OHE, one-hot encoding; ECFP6, extended-connectivity 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 values, averaged over 20 × 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**) UMAP of the GNN embedding space (gray) and depiction of all experimentally observed data points (colored). (**D**) Benchmark of asynchronous and batch-wise BO strategies on the synthetic Ackley surface (six dimensions, discretized). See supplementary text, section 4.4, for further details on the simulation of asynchronous optimization. (**E**) Schematic representation of the asynchronous, multisite optimization workflow for OSL materials discovery. Solid arrows represent materials transfer, and dashed arrows represent data transfer. Experiment status updates to the database are omitted for clarity. Alloc., manual allocation of experiments.

learning from incoming data, enabling maximally informed decision-making at any point in time. The hypothesis of continuous learning was supported by benchmark experiments on synthetic surfaces, which demonstrate increased sample efficiency if multiple workers with parallel experimentation capacities operate in an asynchronous fashion and constantly learn from incoming data (Fig. 3D). Notably, the described workflow allowed for full flexibility, not only with respect to the number of asynchronous workers but also with respect to the inherent infrastructure diversity across different participating workers. This diversity encompasses, for example, the varying degree of automation and digitization in different laboratories, the involvement of humans in the loop, the throughput and downtime of experimentation capacities, and further laboratory-specific constraints that are often hard to quantify a priori.

With this engine in hand, a 2-month optimization campaign for OSL gain materials was carried out, starting from the available seed data. Already from the first pool of recommendations, the Bayesian optimizer identified sets of compounds with state-of-the-art lasing performance. In fact, from the start of the optimization campaign, a total of 12 new compounds with higher–solution gain cross section compared with that of the parent BSBCz were discovered. Although a set of promising candidates had already resulted from early condition optimization efforts within a narrow space around BSBCz, our results suggest the accelerated discovery of high-gain materials

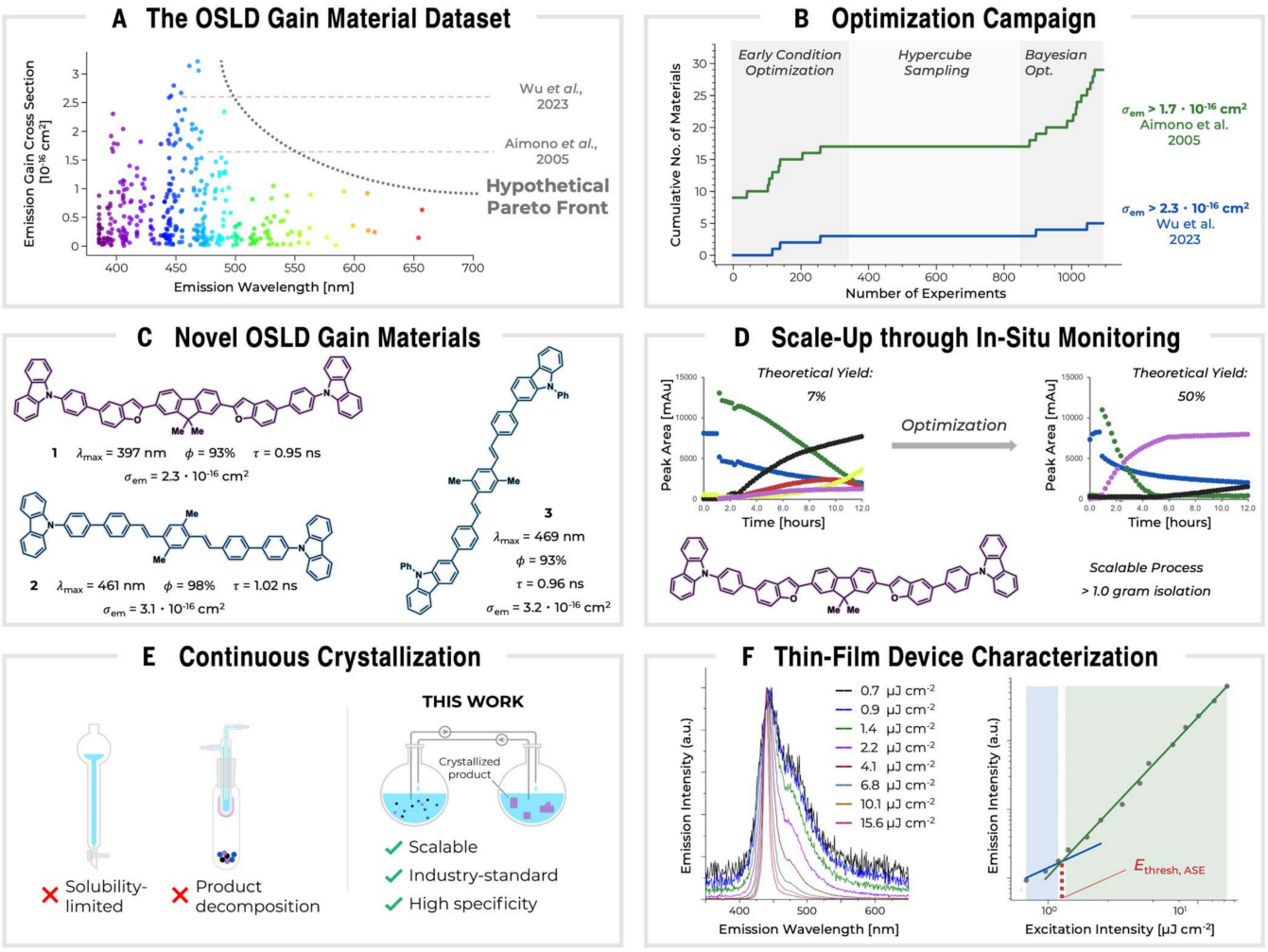


Fig. 4. Discovered OSL gain materials. (**A**) Scatter plot of emission wavelength versus emission gain cross section for all candidates in the OSL gain materials dataset obtained throughout the optimization campaign. (**B**) Cumulative number of discovered materials with emission gain cross sections greater than BSBCz (green) and the top candidate from Wu *et al*. (*38*) (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, reaction intermediate [cap]–[bridge]–B(MIDA); green, Br–[core]–Br 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; violet, product **1**. (**E**) Schematic depiction of the automated crystallization module for scalable purification of OSL emitter materials, owing to the incompatibility of established purification protocols. (**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 ASE.

through the inclusion of BO and automated experimentation (Fig. 4B and supplementary text, section 3.5). Structures and solution-state optical properties of selected molecules discovered in the course of this study are shown in Fig. 4C. Our candidates represent the small-molecule emitters with the highest emission gain cross sections in solution known to date. Because 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 (*53*)—i.e., the solution-state emission gain cross section at room temperature cannot be optimized much further because of physical constraints to emission lifetime and spectral width (see supplementary text, 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 fact 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 (*54*).

Evaluating candidates in thin-film devices

To evaluate the identified lasing candidates in an actual OSL device, synthesis and isolation of larger quantities of material at high purity was required. The two-step one-pot synthesis procedure optimized for generality was ideal to increase throughput of the optimization campaign. However, reaction yields can vary substantially depending on the nature of each particular substrate and the accompanying side product profile, both of which complicate purification. Thus candidates **1** to **3** were prepared on scale using alternative reaction conditions (*55*). Notably, online HPLC was leveraged to enable the large-scale synthesis of candidate **1** using a two-step one-pot procedure (*56*, *57*). This revealed a substantial competitive rate of reductive dehalogenation of the core building block as well as protodeborylation of the intermediate. These side products markedly complicated the purification of the target OSL gain material but could largely be mitigated by avoiding the use of P(*t*-Bu)$_3$ Pd

G4 and the neopentylglycol boronic acid ester cap in favor of Xphos Pd G2 and the pinacol boronic acid ester derivative. Through these modifications, leveraging online HPLC enabled the identification of tailored reaction conditions specifically adapted for the synthesis 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 because high purity (>99%) is required for device fabrication. In the case of our OSL material candidates, well-established chromatography or sublimation protocols were hampered by poor solubility and decomposition upon sublimation. For this purpose, we developed a module for continuous preferential crystallization (CPC) purification (Fig. 4E; see supplementary text, section 3.8, for further details) (*58*, *59*). Although leveraging CPC in this space has largely been overlooked, we believe that this isolation strategy could prove particularly adept at purifying these compound classes. CPCs require only limited solubility of the targeted compound, are often performed at ambient temperature with a high degree of control over supersaturation, and are inherently scalable without increasing the footprint of the isolation setup. Together, the optimized synthetic procedure and continuous preferential purification delivered an exceptionally facile, solution-process route to access gram-scale quantities of **1**, **2**, and **3**, which were used for the preparation of thin-film devices.

Device-level properties were determined by spin-coating thin films of **1** to **3** (3 wt %) in a matrix of 4,4'-Bis(*N*-carbazolyl)-1,1'-biphenyl (CBP) and subsequent spectroscopic characterization. For all materials, ASE could be observed, and low ASE thresholds of 1.5 to 1.9 μJ cm^{-2} were determined (Fig. 4F and figs. S56 to S67). Notably, materials **2** and **3**, initially identified as the highest-gain emitters in solution, exhibited the lowest ASE thresholds in thin film and outperformed a solution-processable BSBCz derivative used as a reference material (E_{th} = 1.49 and 1.50 μJ cm^{-2} versus 1.91 μJ cm^{-2}) (*36*). This observation underscores their 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

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. From an experimental perspective, 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 prefabricated 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 new gain materials with state-of-the-art lasing performance. A set of optimized candidate molecules was prepared and purified on a gram scale, and device-level performance confirmed the identification of best-in-class gain materials (in terms of amplified spontaneous emission threshold).

Drawing from these findings, we envisage three major directions for next-generation workflows toward improved OSL devices: (i) Whereas 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 substantial improvement of reaction yields and robust quantification of molecular properties. (ii) 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. (iii) The systematic identification of sets of function-infused building blocks would enable encoding an optimized candidate space, facilitating interpretability and hybrid human-AI molecular design.

Our work demonstrates a blueprint for delocalized AI-guided discovery campaigns. The multifaceted nature of contemporary materials discovery, characterized by intricate and diverse subtasks, makes it increasingly impractical to centralize all necessary resources in one location. Instead, the integration of existing synthesis, formulation, or characterization modules distributed over multiple sites across the globe can enable increasingly complex discovery workflows, synergistically merging the capacities of automated and human-centric experimentation. At the heart of such a campaign must be a cloud-based hub for data storage, logistics handling, and experiment planning, to enable loss-free data transfer and continuous learning from incoming data and to harness the full potential of AI experiment design. Key requirements for such workflows are the facile accessibility of the central hub [through graphical user interfaces (GUIs) and application programming interfaces (APIs)] and the ability to account for the inherent diversity in digitization and experimentation capacities of different participating laboratories. At the same time, rigorous and, ideally, automated policies for ensuring robustness and reproducibility—e.g., regularly scheduled control experiments with known outcomes or intersite experiment reproduction—should be enacted to ensure high data quality. Eventually, an integrated platform that is open to anybody on the globe and that assembles multiple participating experimental modules to local (fog) units and high-level cloud experimentation workflows can provide a scalable framework for democratizing scientific discovery (*60*).

Materials and methods

This section outlines the operation of our decentralized, cloud-based discovery engine for OSL gain materials discovery. Site- and instrument-specific implementations, along with additional experimental details, are available in the supplementary materials.

Synthesis and end-to-end automated characterization

Throughout the optimization campaign, the cloud hub maintained the up-to-date list of most-informative next experiments (vide infra). From this global catalog, each site was provided with a tailored view based on the available building blocks at the respective site. Human researchers could then flexibly allocate batches of experiments to the respective automated platforms, either semiautomatically through the GUI or in an automated fashion using the API. Once allocated, experiments were immediately restricted from allocation by other users.

Automated two-step one-pot syntheses (second-generation conditions) were performed as follows: The cap building block (1.1 equiv), the bridge building block (1.0 equiv), $P(t\text{-}Bu)_3$ Pd G4 (10 mol %), and TMSOK (1.3 equiv) were added to a reaction vial, which was evacuated and backfilled with N_2. Anhydrous 1,4-dioxane (target concentration: 0.05 M bridge building block) was added, and the reaction mixture was stirred and heated to 60°C for 1 hour. After cooling to room temperature, stock solutions of the core building block (in 1,4-dioxane, 0.40 equiv), XPhos Pd G2 (in 1,4-dioxane, 5 mol %), and Na_2CO_3 (7.5 equiv, in H_2O) were added, and the reaction mixtures were stirred and heated to 100°C for 12 hours. Instrument-specific implementations are detailed in the supplementary materials.

Crude reaction mixtures (after concentration in vacuo and shipping, if applicable) were subjected to the end-to-end automated characterization workflow described by Wu *et al.* (*38*). On the *Chemspeed SWING* platform, the reaction mixtures were filtered through a plug of celite and were injected into the HPLC-MS instrument. If the desired reaction product was detected in the chromatogram (see supplementary materials for details), a second HPLC-MS run was started, and the target compound fraction was collected downstream.

This fraction was then automatically subjected to a sequence of steady-state absorption spectroscopy, steady-state emission spectroscopy, quantum yield determination, and time-resolved emission spectroscopy [see (*38*) and the supplementary materials for a detailed description]. The collected data were uploaded to the cloud-based hub after a rapid, asynchronous visual inspection by a human researcher.

Bayesian optimization

The BO algorithm was designed to maintain the up-to-date ranked list of most-informative next experiments. For this purpose, the optimizer queried the central cloud hub for the most recent data on all completed, ongoing, and pending experiments. A multioutput GP surrogate model was trained on the spectroscopic data from all completed experiments, using the embedding vectors from the pretrained GNN as the molecular featurization. For all ongoing experiments, the GP posterior distribution was conditioned on so-called fantasy values. Concurrently, the optimizer retrieved data on the available cap, bridge, and core building blocks at each site to enumerate the currently accessible active search space (i.e., all molecules that can be synthesized at minimum at one site). Eventually, the updated list of most-informative experiments was generated by evaluating the surrogate model's posterior across the active search space. For this purpose, an ensemble of acquisition functions was iteratively optimized by means of sequential conditioning (for details on the algorithm, see supplementary materials).

Building block preparation

Cap building blocks (as boronic acid pinacol esters) were either purchased from commercial suppliers or synthesized following literature procedures starting from commercially available precursors through (i) esterification of arylboronic acids, (ii) Buchwald-Hartwig–type amination of haloarylboronic acid pinacol esters, or (iii) Miyaura borylation of aryl halides. Details are given in the supplementary materials.

Bridge building blocks (haloarylboronic acid TIDA esters) were prepared by esterification of commercially available haloarylboronic acids. For this, the haloarylboronic acid (1.0 equiv) and TIDA (1.0 equiv) were suspended in an azeotrope solvent mixture [2:1 mixture of benzene or toluene with dimethyl sulfoxide (DMSO)], and the mixture was heated at reflux under Dean-Stark conditions for 3 hours. After washing with water, drying, and concentration in vacuo, the desired product was precipitated from a concentrated acetone solution with hexanes or diethyl ether. This synthesis procedure was either executed in an automated fashion using the *Chemputer* setups in Glasgow and Vancouver (*45*) or was performed manually.

All core building blocks were purchased from commercial suppliers.

On-scale synthesis and purification

For the scale-up of the two-step one-pot synthesis, the cap building block (1.5 equiv), the bridge building block (1.0 equiv), XPhos Pd G2 (10 mol %), and K_3PO_4 (9.0 equiv) were added to a two-necked flask. Under an inert atmosphere, tetrahydrofuran (THF) and water (100 equiv) were added. The reaction was stirred at room temperature and continuously monitored by online HPLC. Upon completion of the first coupling step, a stock solution of the core building block (0.63 equiv) was added, followed by an aqueous NaOH solution (3.0 equiv), and the mixture was stirred overnight. After workup (see supplementary materials for further details), the resulting solid was purified by CPC (*58*, *59*). For this, the crude product was loaded to a dissolver flask followed by the addition of dichloromethane. Continuous circulation of the liquid between the dissolver (30°C) and the crystallizer flask (15°C) over 48 hours led to the crystallization of the product in the latter flask, which was collected by vacuum filtration.

Thin-film device fabrication and characterization

Thin films of the purified materials [3 wt % of emitter in 4,4'-Bis(*N*-carbazolyl)-1,1'-biphenyl] were prepared through spin coating from a chloroform solution. For all films, absorption spectra, emission spectra, and photoluminescence lifetime decay curves were recorded. ASE characteristics were determined by recording the emission intensity and emission line width as a function of excitation intensity. From these data, ASE thresholds were determined by separately fitting the characteristic excitation-emission curves in the subthreshold and above-threshold regimes. For each material, thin-film fabrication and spectroscopic characterization were run in triplicate. All measurements were performed under an inert gas atmosphere. Details of the device fabrication and data analysis are given in the supplementary materials.

REFERENCES AND NOTES

1. A. Zhavoronkov *et al.*, Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nat. Biotechnol.* **37**, 1038–1040 (2019). doi: 10.1038/s41587-019-0224-x; pmid: 31477924
2. H. Uoyama, K. Goushi, K. Shizu, H. Nomura, C. Adachi, Highly efficient organic light-emitting diodes from delayed fluorescence. *Nature* **492**, 234–238 (2012). doi: 10.1038/nature11687; pmid: 23235877
3. D. Larcher, J.-M. Tarascon, Towards greener and more sustainable batteries for electrical energy storage. *Nat. Chem.* **7**, 19–29 (2015). doi: 10.1038/nchem.2085; pmid: 25515886
4. The United Nations, Sustainable Development Goals (2015); https://www.un.org/sustainabledevelopment/.
5. A. Aspuru-Guzik, R. Lindh, M. Reiher, The Matter Simulation (R) evolution. *ACS Cent. Sci.* **4**, 144–152 (2018). doi: 10.1021/acscentsci.7b00550; pmid: 29532014
6. H. Wang *et al.*, Scientific discovery in the age of artificial intelligence. *Nature* **620**, 47–60 (2023). doi: 10.1038/s41586-023-06221-2; pmid: 37532811
7. X. Zhang *et al.*, Artificial Intelligence for Science in Quantum, Atomistic, and Continuum Systems. arXiv:2307.08423 [cs.LG] (2023).
8. R. Macarron *et al.*, Impact of high-throughput screening in biomedical research. *Nat. Rev. Drug Discov.* **10**, 188–195 (2011). doi: 10.1038/nrd3368; pmid: 21358738
9. A. Buitrago Santanilla *et al.*, Nanomole-scale high-throughput chemistry for the synthesis of complex molecules. *Science* **347**, 49–53 (2015). doi: 10.1126/science.1259203; pmid: 25554781
10. F. Häse, L. M. Roch, A. Aspuru-Guzik, Next-Generation Experimentation with Self-Driving Laboratories. *Trends Chem.* **1**, 282–291 (2019). doi: 10.1016/j.trechm.2019.02.007
11. E. Stach *et al.*, Autonomous experimentation systems for materials development: A community perspective. *Matter* **4**, 2702–2726 (2021). doi: 10.1016/j.matt.2021.06.036
12. M. Seifrid *et al.*, Autonomous Chemical Experiments: Challenges and Perspectives on Establishing a Self-Driving Lab. *Acc. Chem. Res.* **55**, 2454–2466 (2022). doi: 10.1021/acs.accounts.2c00220; pmid: 35948428
13. M. Abolhasani, E. Kumacheva, The rise of self-driving labs in chemical and materials sciences. *Nat. Synth.* **2**, 483–492 (2023). doi: 10.1038/s44160-022-00231-0
14. A.-C. Bédard *et al.*, Reconfigurable system for automated optimization of diverse chemical reactions. *Science* **361**, 1220–1225 (2018). doi: 10.1126/science.aat0650; pmid: 30237351
15. M. Christensen *et al.*, Data-science driven autonomous process optimization. *Commun. Chem.* **4**, 112 (2021). doi: 10.1038/s42004-021-00550-x; pmid: 36697524
16. N. H. Angello *et al.*, Closed-loop optimization of general reaction conditions for heteroaryl Suzuki-Miyaura coupling. *Science* **378**, 399–405 (2022). doi: 10.1126/science.adc8743; pmid: 36302014
17. B. P. MacLeod *et al.*, Self-driving laboratory for accelerated discovery of thin-film materials. *Sci. Adv.* **6**, eaaz8867 (2020). doi: 10.1126/sciadv.aaz8867; pmid: 32426501
18. R. A. Skilton *et al.*, Remote-controlled experiments with cloud chemistry. *Nat. Chem.* **7**, 1–5 (2015). doi: 10.1038/nchem.2143; pmid: 25515870
19. D. E. Fitzpatrick, T. Maujean, A. C. Evans, S. V. Ley, Across-the-World Automated Optimization and Continuous-Flow Synthesis of Pharmaceutical Agents Operating Through a Cloud-Based Server. *Angew. Chem. Int. Ed.* **57**, 15128–15132 (2018). doi: 10.1002/anie.201809080; pmid: 30272384
20. B. J. Shields *et al.*, Bayesian reaction optimization as a tool for chemical synthesis. *Nature* **590**, 89–96 (2021). doi: 10.1038/s41586-021-03213-y; pmid: 33536653
21. D. Caramelli *et al.*, Networking chemical robots for reaction multitasking. *Nat. Commun.* **9**, 3406 (2018). doi: 10.1038/s41467-018-05828-8; pmid: 30143646
22. L. M. Roch *et al.*, ChemOS: An orchestration software to democratize autonomous discovery. *PLOS ONE* **15**, e0229862 (2020). doi: 10.1371/journal.pone.0229862; pmid: 32298284
23. J. Bai *et al.*, From Platform to Knowledge Graph: Evolution of Laboratory Automation. *JACS Au* **2**, 292–309 (2022). doi: 10.1021/jacsau.1c00438; pmid: 35252980
24. M. Vogler *et al.*, Brokering between tenants for an international materials acceleration platform. *Matter* **6**, 2647–2665 (2023). doi: 10.1016/j.matt.2023.07.016
25. R. B. Merrifield, Automated synthesis of peptides. *Science* **150**, 178–185 (1965). doi: 10.1126/science.150.3693.178; pmid: 5319951
26. O. J. Plante, E. R. Palmacci, P. H. Seeberger, Automated solid-phase synthesis of oligosaccharides. *Science* **291**, 1523–1527 (2001). doi: 10.1126/science.1057324; pmid: 11222853
27. M. H. Caruthers, Gene synthesis machines: DNA chemistry and its uses. *Science* **230**, 281–285 (1985). doi: 10.1126/science.3863253; pmid: 3863253
28. E. P. Gillis, M. D. Burke, A simple and modular strategy for small molecule synthesis: Iterative Suzuki-Miyaura coupling of B-protected haloboronic acid building blocks. *J. Am. Chem. Soc.* **129**, 6716–6717 (2007). doi: 10.1021/ja0716204; pmid: 17488084
29. J. Li *et al.*, Synthesis of many different types of organic small molecules using one automated process. *Science* **347**, 1221–1226 (2015). doi: 10.1126/science.aaa5414; pmid: 25766227
30. J. Clark, G. Lanzani, Organic photonics for communications. *Nat. Photonics* **4**, 438–446 (2010). doi: 10.1038/nphoton.2010.160

31. A. J. C. Kuehne, M. C. Gather, Organic Lasers: Recent Developments on Materials, Device Geometries, and Fabrication Techniques. *Chem. Rev.* **116**, 12823–12864 (2016). doi: 10.1021/acs.chemrev.6b00172; pmid: 27501192
32. Q. Zhang, W. Tao, J. Huang, R. Xia, J. Cabanillas-Gonzalez, Toward Electrically Pumped Organic Lasers: A Review and Outlook on Material Developments and Resonator Architectures. *Adv. Photon. Res.* **2**, 2000155 (2021). doi: 10.1002/adpr.202000155
33. K. Yoshida *et al*., Electrically driven organic laser using integrated OLED pumping. *Nature* **621**, 746–752 (2023). doi: 10.1038/s41586-023-06488-5; pmid: 37758890
34. T. Aimono *et al*., 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.* **86**, 071110 (2005). doi: 10.1063/1.1867555
35. A. S. D. Sandanayaka *et al*., Indication of current-injection lasing from an organic semiconductor. *Appl. Phys. Express* **12**, 061010 (2019). doi: 10.7567/1882-0786/ab1b90
36. M. Mamada, T. Fukunaga, F. Bencheikh, A. S. D. Sandanayaka, C. Adachi, Low Amplified Spontaneous Emission Threshold from Organic Dyes Based on Bis-stilbene. *Adv. Funct. Mater.* **28**, 1802130 (2018). doi: 10.1002/adfm.201802130
37. Y. Oyama *et al*., Design Strategy for Robust Organic Semiconductor Laser Dyes. *ACS Mater. Lett.* **2**, 161–167 (2020). doi: 10.1021/acsmaterialslett.9b00536
38. T. C. Wu *et al*., A Materials Acceleration Platform for Organic Laser Discovery. *Adv. Mater.* **35**, e2207070 (2023). doi: 10.1002/adma.202370042; pmid: 36373553
39. S. G. Ballmer, E. P. Gillis, M. D. Burke, "B-Protected Haloboronic Acids for Iterative Cross-Coupling" in *Organic Syntheses* (Wiley, 2009), pp. 344–359.doi: 10.1002/0471264229.os086.33
40. D. J. Blair *et al*., Automated iterative Csp^3-C bond formation. *Nature* **604**, 92–97 (2022). doi: 10.1038/s41586-022-04491-w; pmid: 35134814
41. C. P. Delaney, V. M. Kassel, S. E. Denmark, Potassium Trimethylsilanolate Enables Rapid, Homogeneous Suzuki-Miyaura Cross-Coupling of Boronic Esters. *ACS Catal.* **10**, 73–80 (2020). doi: 10.1021/acscatal.9b04353; pmid: 33585070
42. W. Wang *et al*., Rapid Automated Iterative Small Molecule Synthesis. ChemRxiv [Preprint] (2023); https://doi.org/10.26434/chemrxiv-2023-qpf2x
43. S. H. M. Mehr, M. Craven, A. I. Leonov, G. Keenan, L. Cronin, A universal system for digitization and automatic execution of the chemical synthesis literature. *Science* **370**, 101–108 (2020). doi: 10.1126/science.abc2986; pmid: 33004517
44. S. Steiner *et al*., Organic synthesis in a modular robotic system driven by a chemical programming language. *Science* **363**, eaav2211 (2019). doi: 10.1126/science.aav2211; pmid: 30498165
45. R. Rauschen, M. Guy, J. E. Hein, L. Cronin, Universal Chemical Programming Language for Robotic Synthesis Reproducibility. *Nat. Synth.* **3**, 488–496 (2024). doi: 10.1038/s44160-023-00473-6
46. A. V. Deshpande, A. Beidoun, A. Penzkofer, G. Wagenblast, Absorption and emission spectroscopic investigation of cyanovinyldiethylaniline dye vapors. *Chem. Phys.* **142**, 123–131 (1990). doi: 10.1016/0301-0104(90)89075-2
47. H. Nakanotani, C. Adachi, S. Watanabe, R. Katoh, Spectrally narrow emission from organic films under continuous-wave excitation. *Appl. Phys. Lett.* **90**, 231109 (2007). doi: 10.1063/1.2746958
48. H. J. Kushner, A New Method of Locating the Maximum Point of an Arbitrary Multipeak Curve in the Presence of Noise. *J. Basic Eng.* **86**, 97–106 (1964). doi: 10.1115/1.3653121
49. J. Snoek, H. Larochelle, R. P. Adams, "Practical Bayesian Optimization of Machine Learning Algorithms," in *Advances in Neural Information Processing Systems*, F. Pereira, C. J. Burges, L. Bottou, K. Q. Weinberger, Eds. (Curran Associates, Inc., 2012).
50. B. Huang, G. F. von Rudorff, O. A. von Lilienfeld, The central role of density functional theory in the AI age. *Science* **381**, 170–175 (2023). doi: 10.1126/science.abn3445; pmid: 37440654
51. A. Baiardi, J. Bloino, V. Barone, General Time Dependent Approach to Vibronic Spectroscopy Including Franck-Condon, Herzberg-Teller, and Duschinsky Effects. *J. Chem. Theory Comput.* **9**, 4097–4115 (2013). doi: 10.1021/ct400450k; pmid: 26592403
52. M. Aldeghi *et al*., Roughness of Molecular Property Landscapes and Its Impact on Modellability. *J. Chem. Inf. Model.* **62**, 4660–4671 (2022). doi: 10.1021/acs.jcim.2c00903; pmid: 36112568
53. J. Lyu *et al*., Ultra-large library docking for discovering new chemotypes. *Nature* **566**, 224–229 (2019). doi: 10.1038/s41586-019-0917-9; pmid: 30728502
54. E. Castellano-Hernández, P. W. Metz, M. Demesh, C. Kränkel, Efficient directly emitting high-power Tb^{3+}:$LiLuF_4$ laser operating at 587.5 nm in the yellow range. *Opt. Lett.* **43**, 4791–4794 (2018). doi: 10.1364/OL.43.004791; pmid: 30272741
55. Materials 2 and 3 were synthesized manually (see supplementary text, section 3.6, for further details).
56. J. Liu, Y. Sato, F. Yang, A. J. Kukor, J. E. Hein, An Adaptive Auto-Synthesizer using Online PAT Feedback to Flexibly Perform a Multistep Reaction. *Chem. Methods* **2**, e202200009 (2022). doi: 10.1002/cmtd.202200009
57. T. C. Malig *et al*., Development of a telescoped synthesis of 4-(1*H*)-cyanoimidazole core accelerated by orthogonal reaction monitoring. *React. Chem. Eng.* **5**, 1421–1428 (2020). doi: 10.1039/D0RE00234H
58. J. E. Hein, B. H. Cao, M. W. Van Der Meijden, M. Leeman, R. M. Kellogg, Resolution of Omeprazole Using Coupled Preferential Crystallization: Efficient Separation of a Nonracemizable Conglomerate Salt under Near-Equilibrium Conditions. *Org. Process Res. Dev.* **17**, 946–950 (2013). doi: 10.1021/op400081c
59. C. Rougeot, J. E. Hein, Application of Continuous Preferential Crystallization to Efficiently Access Enantiopure Chemicals. *Org. Process Res. Dev.* **19**, 1809–1819 (2015). doi: 10.1021/acs.oprd.5b00141
60. M. Sim *et al*., ChemOS 2.0: An orchestration architecture for chemical self-driving laboratories. ChemRxiv [Preprint] (2023); https://doi.org/10.26434/chemrxiv-2023-v2khf
61. F. Strieth-Kalthoff *et al*., Delocalized, Asynchronous, Closed-Loop Discovery of Organic Laser Emitters, Data set, Zenodo (2024); https://doi.org/10.5281/zenodo.10909257
62. F. Strieth-Kalthoff, M. Seifrid, R. Hickman, Delocalized, Asynchronous, Closed-Loop (ACDC) Discovery of Organic Laser Materials, Github (2023); https://github.com/aspuru-guzik-group/acdc_laser.
63. F. Strieth-Kalthoff *et al*., Delocalized, Asynchronous, Closed-Loop Discovery of Organic Laser Emitters, version 1.0.0, Zenodo (2023); https://doi.org/10.5281/zenodo.8357375

ACKNOWLEDGMENTS

The authors thank A. Fischer for conceiving the DARPA Accelerated Molecular Discovery Program and for numerous fruitful discussions. R. Keunen (University of Toronto) is acknowledged for experimental and administrative assistance. **Funding:** The authors acknowledge the Defense Advanced Research Projects Agency (DARPA) under the Accelerated Molecular Discovery Program under cooperative agreement no. HR00111920027, dated 1 August 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. F.S.-K. is a postdoctoral fellow in the Eric and Wendy Schmidt AI in Science Postdoctoral Fellowship Program, a program by Schmidt Futures. R.P. acknowledges funding through a Postdoc.Mobility fellowship by the Swiss National Science Foundation (SNSF; project no. 191127). C.B.-G. acknowledges funding from a Marie Sklodowska Curie Actions Postdoctoral Fellowship grant (101064374). C.A. thanks the Japan Science and Technology Agency (JST) CREST (grant no. JPMJCR22B3) Specially Promoted Research (grant no. 23H05406). L.C. 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). L.C. thanks Schmidt Futures for an Innovation Fellowship. B.A.G. was supported by the Institute for Basic Science, Korea (project code IBS-R020-D1). M.D.B. acknowledges funding from the Molecule Maker Lab Institute, an AI Institutes program supported by the US National Science Foundation under grant no. 2019897. A.A.-Gu. thanks A. G. Frøseth for his generous support. A.A.-Gu. also acknowledges the generous support of Natural Resources Canada and the Canada 150 Research Chairs program. 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). **Author contributions:** F.S.-K. and H.H. contributed equally to this work, and the order of authorship was determined through a coin flip. 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., and A.A.-Gu. 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.Wo., R.R., C.T.S., L.C., and 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.Wo., R.R., and C.T.S. Funding acquisition: C.A., B.A.G., L.C., J.E.H., M.D.B., and A.A.-Gu. 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., S.Li, A.Wo., R.R., C.T.S., L.C., and 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.C.W., K.H., M.H., A.Wo., R.R., C.T.S., C.B.-G., R.J.H., J.V., A.A.-Gr., E.L.K., R.C.S., W.H., D.G., S.La., A.Wa., O.B., S.R., B.S.-L., C.A., B.A.G., L.C., J.E.H., M.D.B., and A.A.-Gu. 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., and A.A.-Gu. Resources: C.A., B.A.G., L.C., J.E.H., M.D.B., and A.A.-Gu. Software: F.S.-K., H.H., T.G., M.S., C.L., R.P., T.C.W., K.H., M.H., A.Wo., R.R., C.T.S., and D.G. Supervision: C.A., B.A.G., L.C., J.E.H., M.D.B., and A.A.-Gu. 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.Li., C.T.S., and C.B.-G. Visualization: F.S.-K. and T.G. Writing – original draft: F.S.-K., H.H., C.A., B.A.G., L.C., J.E.H., M.D.B., and A.A.-Gu. Writing – review & editing: all authors. **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.-Gu. is chief visionary officer and board member of Kebotix Inc., a company that carries out closed-loop molecular materials discovery. M.D.B., N.H.A., and W.W. are inventors on patents and/or patent applications related to MIDA and/or TIDA boronates held and submitted by the University of Illinois at Urbana-Champaign that cover composition of matter and methods of use. The authors declare no other competing interests. **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 [nuclear magnetic resonance (NMR) spectra of all building blocks and scaled-up materials and raw HPLC-MS data in open-source format] are available at Zenodo (*61*). Synthesis and optical spectroscopy results for all target compounds are given in data S6 and have been deposited at Zenodo (*61*). Results of the high-throughput computational analysis are available at Zenodo (*61*). All training data for ML models are given in data S4 and S5 and are deposited at Zenodo (*61*). All software used in this work is freely available on Github (*62*), and a snapshot of the code has been deposited at Zenodo (*63*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adk9227
Supplementary Text
Figs. S1 to S67
Tables S1 to S19
References (*64*–*109*)
Data S1 to S6

Submitted 18 September 2023; accepted 5 April 2024
10.1126/science.adk9227