Rapid Planning and Analysis of High-Throughput Experiment Arrays for Reaction Discovery

Babak Mahjour[1], Rui Zhang[2], Yuning Shen[1], Andrew McGrath[1], Ruheng Zhao[1], Osama G. Mohamed[3], Yingfu Lin[1], Zirong Zhang[1], James L. Douthwaite[1], Ashootosh Tripathi[1,3], Tim Cernak*[1,2]

[1]Department of Medicinal Chemistry, University of Michigan
[2]Department of Chemistry, University of Michigan
[3]Natural Products Discovery Core, Life Sciences Institute, University of Michigan

ABSTRACT: High-throughput experimentation (HTE) is an increasingly important tool in reaction discovery. While the hardware for running HTE in the chemical laboratory has evolved significantly in recent years, there remains a need for software solutions to navigate data-rich experiments. Here we have developed phactor™, a software that facilitates the performance and analysis of HTE in a chemical laboratory. phactor™ allows experimentalists to rapidly design arrays of chemical reactions or direct-to-biology experiments in 24, 96, 384, or 1,536 wellplates. Users can access online reagent data, such as a chemical inventory, to virtually populate wells with experiments and produce instructions to perform the reaction array manually, or with the assistance of a liquid handling robot. After completion of the reaction array, analytical results can be uploaded for facile evaluation, and to guide the next series of experiments. All chemical data, metadata, and results are stored in machine-readable formats that are readily translatable to various software formats. We also demonstrate the use of phactor™ in the discovery of several chemistries, including a light-enabled reagent-free sp^2–sp^2 deaminative–decarboxylative carbon–carbon coupling and a low micromolar inhibitor of the SARS-CoV-2 main protease. Furthermore, phactor™ has been made available for free academic use in 24 and 96 well formats via an online interface.

INTRODUCTION

Miniaturized high-throughput experimentation (HTE) has emerged as an accessible, reliable, economical, and environmentally friendly technique for the rapid discovery of new reactivities.[1-33] Curated HTE data proves to be increasingly valuable for predictive models,[15-20] While in experimental practice, the community has gravitated towards liquid handling techniques in glass shell microvials with tumble stir dowels, or in plastic 384 or 1,536 wellplates[1,4-14], a standard for HTE data handling has yet to be established. The organizational load required to perform a simple 24-well reaction array is generally manageable by repetitive notebook entries or with spreadsheets, yet managing multiple reaction arrays in a single day, or running ultraHTE in 1,536 wellplates,[1] is challenging without information management software. Moreover, no readily available electronic lab notebook (ELN) can store HTE details in a tractable manner[21,22] or provide a simple interface to extract data and results from multiple experiments simultaneously.[23,24] To continue developing HTE research and position data outputs for machine learning studies, detailed reaction data must be easily accessible for and standardized rapid extraction and analysis.[25-27]

With these issues in mind, we developed the software phactor™ to streamline the collection of HTE reaction data. Our primary objective was to develop a robust yet generalizable HTE workflow solution that captures the nuances of chemical experimentation while reporting data in a standardized, machine-readable format. phactor™ minimizes the time and resources spent between experiment ideation and result interpretation. This enables creativity by freeing up time otherwise used thinking about experiment logistics, facilitates reaction discovery and optimization, and serves as a tool to bolster the amount of available reaction data reported in a standardized format. We have provided phactor™ as a free web service to the academic community, which can be accessed at https://phactor.cernaklab.com.

The workflow of a typical high-throughput experiment involves design of the reaction array, preparation of reagent stock solutions, dosing of stock solutions according to the reaction array recipe (either by hand or with robotics), analysis of reaction outcome, followed by visualization and analysis of data and documentation of results. A standardized reaction template classifies substrates, reagents, and products (Figure 1A).

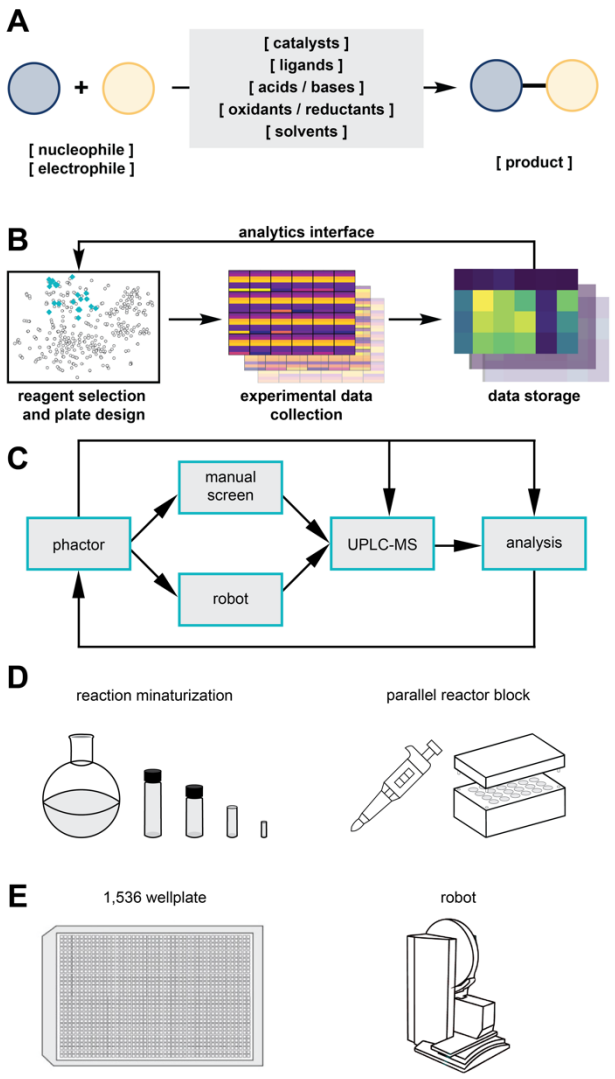


Figure 1. (A) Anatomy of a reaction as encoded by phactor™. (B) High-level software workflow of phactor™. Reaction arrays are designed from chemical inventories and reaction informatics. Resultant data is stored in delimited text (CSV) or in a relational database (SQLite3). phactor™ can convert results to Open Reaction Database (ORD)[20] and Chemical Description Language (XDL)[48] and is readily compatible with optimization programs such as EDBO+[46] and LabMate.ML.[47] (C) Workflow of phactor™. Once the reaction array is designed, phactor™ provides human readable or machine instructions to execute the dosing manually or robotically. (D) phactor™ supports custom volumes allowing for reaction arrays to be performed at any scale. At minimum, the hardware needed to execute a reaction array is an autopipette and an array reactor block. (E) phactor™ facilitates the design and execution of ultraHTE in 1,536 wellplates.

Interconnecting experimental results with online chemical inventories through this shared data format creates a closed-loop workflow for HTE-driven chemical research (Figure 1B) and enables rapid reaction array design and analytics. While developing phactor™, we sought to maximize the automation of data movement and processing. Recognizing the rapidly accelerating chemical research software ecosystem,[28-48] the philosophy behind phactor™'s data structure was to record experimental procedures and results in a machine-readable yet simple, robust, and abstractable format to naturally translate to other system languages. As such, the inputs and outputs of phactor™ can be procedurally generated or modified with basic Excel or Python knowledge to interface with any robot, analytical instrument, software, or custom chemical inventory containing metadata such as reagent location, molecular weight, CAS number or SMILES string (e.g., the organophosphorous ligand platform Kraken).[49] Examples of interfacing phactor™ outputs with ORD,[20] XDL,[48] or EDBO+[46] are given in the Supporting Information.

The event workflow of a typical phactor™ experiment is shown in Figure 1C. The user selects desired reagents from the inventory for automatic field population or enters specific reagent entries manually, such as for a custom substrate. Once all relevant chemicals are selected, the reaction array layout is designed automatically or manually, as the user prefers. Reagent distribution instructions are generated to be executed either manually or by an interfacing liquid handling robot. Last-minute changes in the face of unforeseen circumstances during reaction setups, such as poor chemical solubility, chemical instability, or the need to premix reagents before dosing can be made at any time. Stock solutions are prepared in vials or wellplates and distributed to their respective locations on the reaction wellplate. Once the reactions are complete, they are quenched and analyzed. Any data with a well location map can be uploaded. This allows both data on reaction performance (e.g., UPLC-MS conversion) and biological assay results (e.g., bioactivity data) to be viewed in concert.

The workflows for executing HTE can vary depending on available equipment and desired throughput of the experiment. phactor™ incorporates these parameters into its user interface to ensure a consistent workflow experience agnostic of hardware capabilities (Figure 1D). Examples herein demonstrate phactor™ integrated with the Opentrons OT-2 liquid handling robot for experiments of 384-well throughput or less, and the SPT Labtech mosquito® robot for 1,536-well ultraHTE (Figure 1E). Regardless of instrumentation or throughput, all results are stored in the same format, facilitating analysis of results across multiple experiments. Reaction discovery and library synthesis campaigns utilizing standard 24-, 96-, 384- and 1,536-well experiments are described in detail.[50-55]

RESULTS AND DISCUSSION

Phactor™ facilitates the discovery of new reactivity. Our lab is broadly interested in novel amine-acid coupling reactions[50,51,53-55] and particularly amine-acid C–C coupling reactions[53,55]. As part of our studies, we have been exploring mild benzoic acid decarboxylation methods.[56-60] Using phactor™, we discovered very simple reaction conditions for a decarboxylative–deaminative amine–acid sp^2–sp^2 C–C coupling reaction between acid **1** and *p*-toluidine, activated as its diazonium salt (**2**), to give biaryl product **3** when carried out with UV irradiation in DMSO for 18 hours.

Ultimately 424 experiments were performed over a series of 24 and 96-well reaction arrays to explore this novel catalyst-free reactivity (Figure 2A). For each

reaction array, stock solutions were created manually according to the phactor™ recipe, and the appropriate volume was dispensed manually into glass shell microvials in each solution's designed location. Parylene-coated stir dowels were added to each vial, and the reaction array was sealed in an aluminum block and then stirred at room temperature for 18 hours. Once the reaction was complete, aliquots of each well were added to an analytical plate containing one molar equivalent of caffeine solution in acetonitrile for UPLC-MS analysis. By utilizing high-throughput reaction informatics, the output files provided by phactor™ for each experiment were collated and processed for user visualization. Isoquinolines **1**, **4**, **5**, and **6**, and diazonium salts **2**, **7**, **8**, **9**, **10**, and **11** produced appreciable yields (measured by product/internal standard integrations) in methanol, DMSO, and a 1:1 mixture of methanol and DMSO (Figure 2B). An array performed in the campaign is shown in Figure 2C, where the four acids and six diazonium salts in Figure 2B were irradiated in methanol under UV light for 18 hours. An interactive heatmap that stacks the UPLC-MS chromatograms from each well in a separate 96-well experiment from the campaign is shown in Figure 2D. Timepoints can be selected to identify mass signals, reaction inputs, expected products, and UV intensities by simply uploading the 96 raw UPLC data files. Kernel density estimate plots grouped by reagent and irradiation treatment (Figure 2E) reveal that blue light irradiation modestly outperforms both UV and white light irradiation and that the 1:1 mixture of methanol and DMSO performs better than DMSO or methanol individually when considering average yields, but pure methanol generated the highest yielding products. Scatter plotting shows blue light having a slight edge in reaction performance, although white and UV light do perform well (Supporting Information). The conditions producing the highest yield between acid **1** and diazonium salt **2** were found to be methanol and blue light irradiation, which was scaled up to 0.1 mmol scale and yielded 50% of **3**.

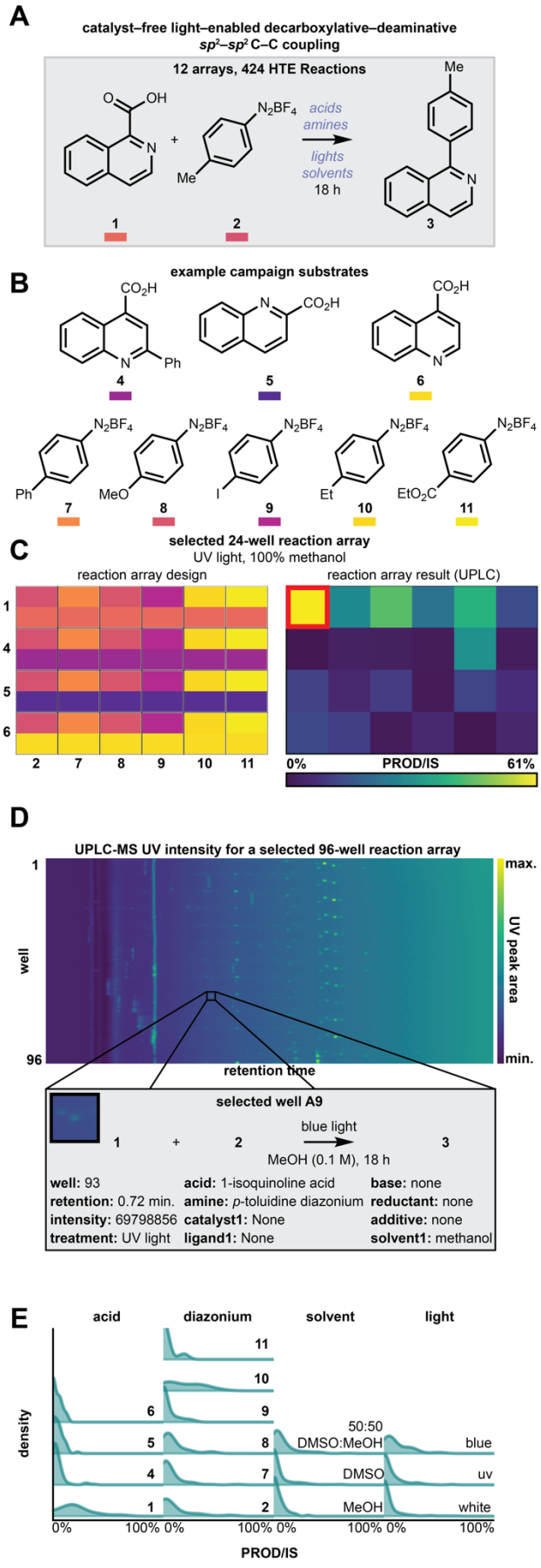


Figure 2. (A) Standard conditions for catalyst-free light-enabled decarboxylative–deaminative sp^2–sp^2 C–C coupling. (B) Sampling of substrates explored. (C) Screen

designed in phactor™ towards the development of the reactivity. (D) Interactive UV heatmap generated from the UPLC-MS data of a 96-well reaction array in the discovery campaign. (E) Kernel density estimate plots generated from phactor™. Model substrates **1** and **2** performed best, followed by other diazoniums with electron-donating substituents.

Experimental Analysis. Diverse chemistries discovered with the aid of phactor™ are shown in Figure 3. Figure 3A shows the discovery of a deaminative aryl esterification.[50,51] In the reaction array design, an amine, activated as its diazonium salt (**2**), a carboxylic acid (**12**), one of three transition metal catalysts, with or without silver nitrate as an additive, and one of four ligands were to be added to each reaction well in acetonitrile, then stirred at 60 ºC for 18 hours. phactor™ automatically designed the reagent distribution recipe by splitting the plate into a simple four-row and six column multiplexed array. At competition, a solution containing one molar equivalent of caffeine was added to each well as an internal standard. An aliquot of each reaction was transferred to a plastic plate, then diluted with acetonitrile for UPLC-MS analysis of the desired ester product (**13**). UPLC-MS output files were analyzed by the commercial software Virscidian Analytical Studio, which provided a csv file containing peak integration values for each of the 24 chromatographic traces. This file was fed into phactor™ to record the experimental outcome and produce the heatmap shown in Figure 3A. Analysis on phactor™ indicated an 18.5% assay yield using 30 mol% CuI and $AgNO_3$, and these specific conditions were triaged for further study.

In the example of Figure 3B, we optimized the penultimate step in our synthesis of umifenovir,[52] an oxidative indolization reaction between **14** and **15** to produce **16**. Inspired by the conditions of Glorius,[61] a reaction array was performed using copper catalysts and ligand/additive combinations. Four copper catalysts: cuprous iodide, cuprous bromide, tetrakis(acetonitrile) copper(I) triflate, or cupric acetate, at 20 mol% were distributed into the four rows while combinations of magnesium sulfate (0.0 equiv or 1.0 equiv) with 2-(1H-tetrazol-1-yl)acetic acid (**L1**), or 2,6-dimethylanilino(oxo)acetic acid (**L2**) at 40 mol% were distributed into the columns as DMSO solutions, with 3.0 equivalents of cesium carbonate added to every well as a suspension in DMSO. The reactions were manually arrayed in a glovebox, sealed, and stirred at 55 ºC for 18 hours. Well B3 (copper bromide with **L1** and no magnesium sulfate) was found to perform best, and a 0.10 mmol scale-up reaction produced desired indole **16** in 66% isolated yield.

In Figure 3C, the allylation of furanone **17** or furan **18** by allylating reagents **19** or **20** was investigated. For each combination of nucleophile and electrophile, one of three ratios of Pd_2dba_3 to (*S*,*S*)-DACH-phenyl Trost ligand **L3** loadings were added, followed by the addition or omission of potassium carbonate as a base. Each reaction was run in toluene for 24 hours at room temperature. The reactions were quenched then analyzed by UPLC for conversion and selectivity. Multiplexed pie charts generated by phactor™ reveal the conditions of well D3, with a 2:1 palladium catalyst to ligand loading and no base generated the desired γ-regioisomer with the greatest selectivity, along with α-allylation and its olefin isomer when **18** was used (see Supporting Information).

In Figure 3D, an organocatalyzed asymmetric Mannich reaction between aldehyde **23**, *p*-anisidine (**24**), and ketone **25** was optimized via solvent and catalyst reaction array and analyzed by TLC. One of six chiral secondary amine catalysts **C1** - **C6** (see Supporting Information) at 20 mol% loading was added to the three starting materials, and the reaction was run in one of four different solvents: DMSO, THF, acetonitrile, or toluene. After 24 hours at room temperature, the reaction array was quenched and analyzed yielding a bright fluorescent spot at Rf value 0.4. Well A3 was scaled up with slight modification to confirm the fluorescent compound as undesired quinoline product **27** in 30% isolated yield (60% to limiting reagent **23**), derived from two equivalents of aldehyde **23**.

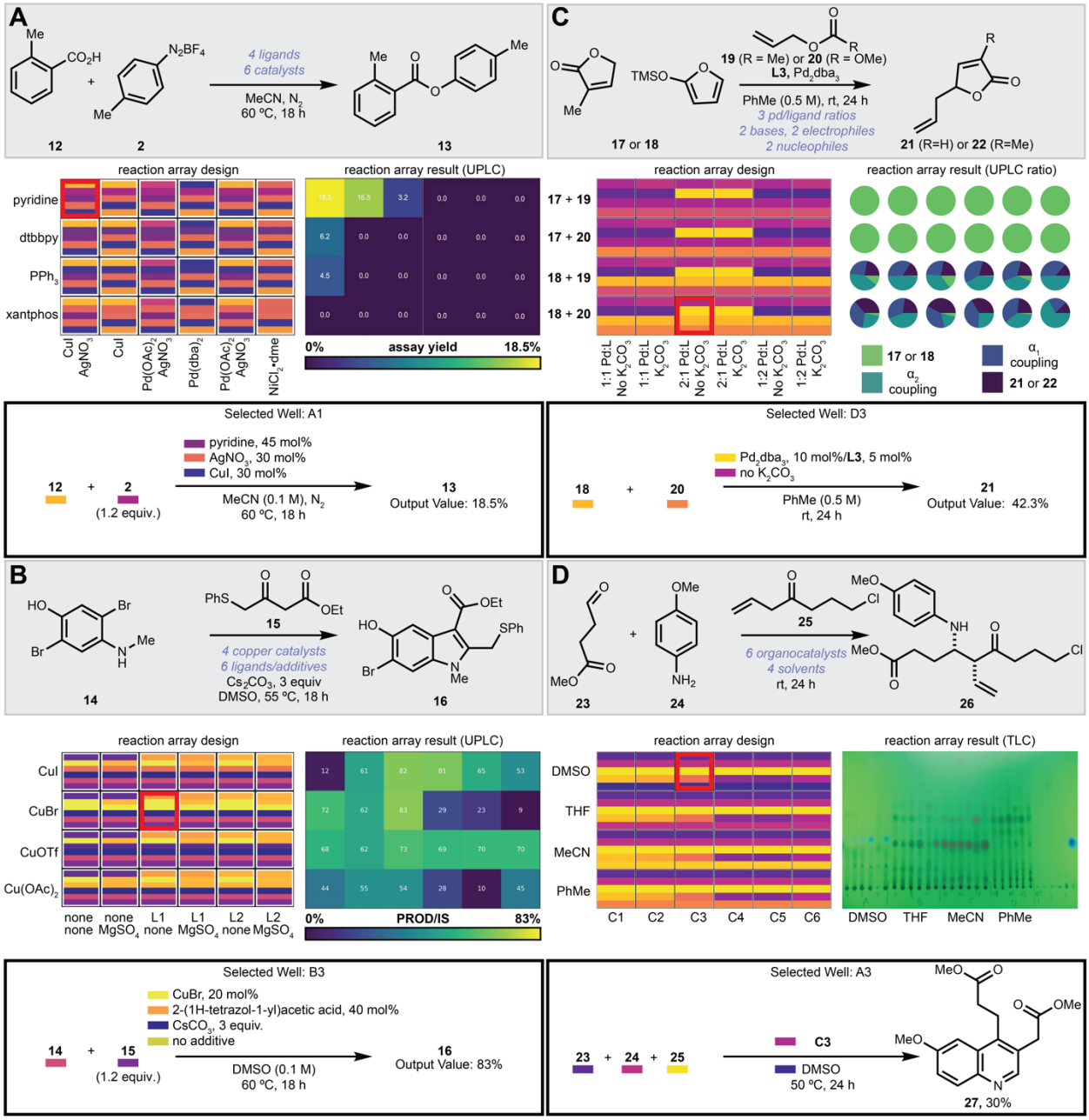


Figure 3. Specific examples of chemistries designed and discovered using phactor™. The reaction array design and results are shown here as displayed on phactor™. Product/internal standard ratios are calculated using the observed UV-derived peak area, while assay yields account for differences in product absorptivity by calibrating to authentic samples of products. See Supporting Information for details. (A) Preliminary esterification hit leading to publication.[51] (B) Optimized oxidative indolization conditions towards the synthesis of umifenovir.[52] (C) Allylation catalyst/ligand concentration ratio and base reaction array analyzed by conversion and selectivity. (D) Mannich reaction optimization with a solvent/catalyst reaction array and analyzed by TLC, yielding in undesired product **27**.

With the software, HTE becomes an exercise in workflow execution, with automation of the organizational aspect of the experiment. This allows chemists to focus on the design and analysis of the reaction array, rather than workflow details (Figure 4). Figures 4A-4C displays three examples of 24-well experiments.

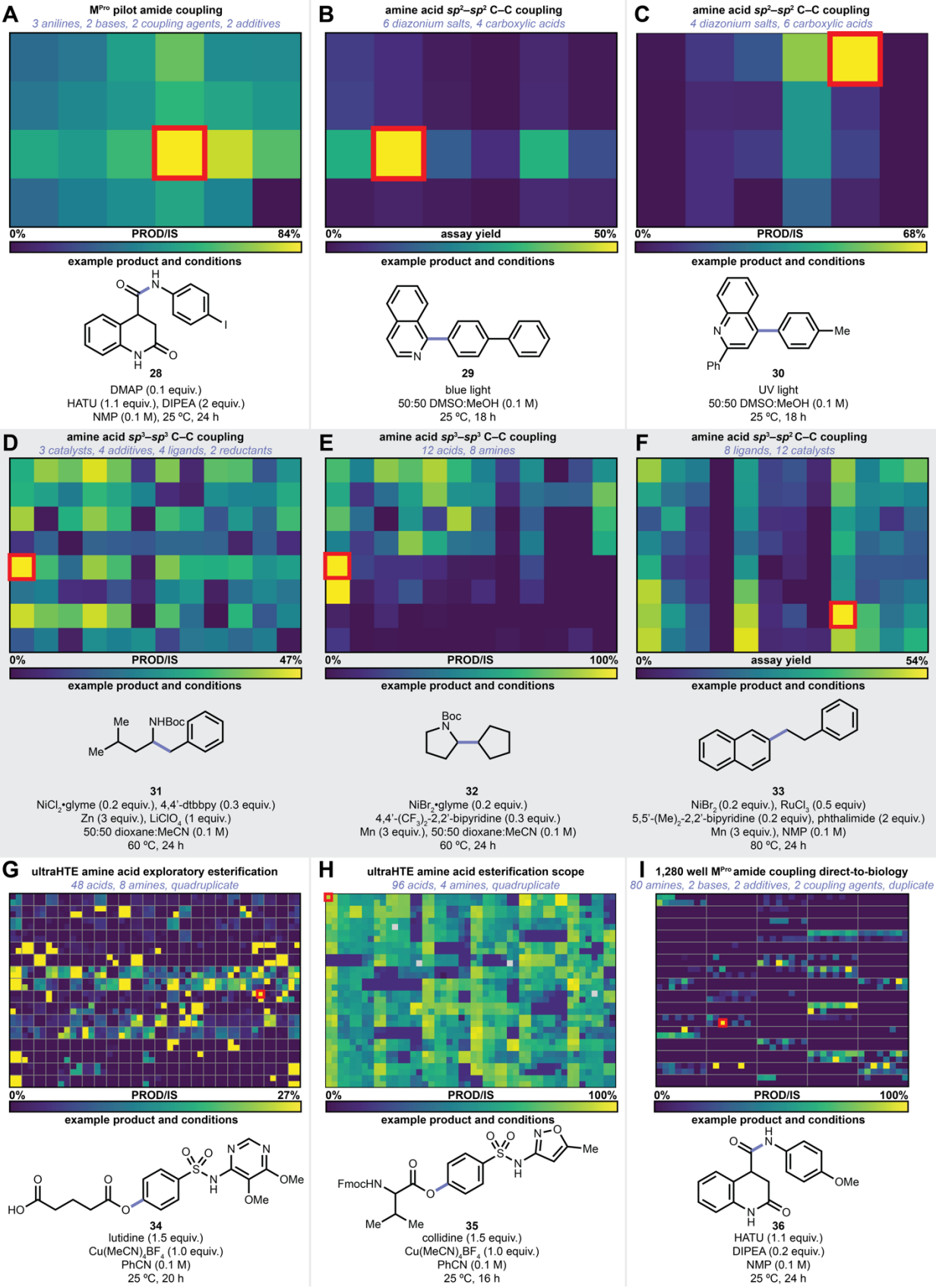

A
M^Pro pilot amide coupling
3 anilines, 2 bases, 2 coupling agents, 2 additives
0%
PROD/IS
84%
example product and conditions
28
DMAP (0.1 equiv.)
HATU (1.1 equiv.), DIPEA (2 equiv.)
NMP (0.1 M), 25 °C, 24 h
B
amine acid sp²–sp² C–C coupling
6 diazonium salts, 4 carboxylic acids
0%
assay yield
50%
example product and conditions
29
blue light
50:50 DMSO:MeOH (0.1 M)
25 °C, 18 h
C
amine acid sp²–sp² C–C coupling
4 diazonium salts, 6 carboxylic acids
0%
PROD/IS
68%
example product and conditions
30
UV light
50:50 DMSO:MeOH (0.1 M)
25 °C, 18 h
D
amine acid sp³–sp³ C–C coupling
3 catalysts, 4 additives, 4 ligands, 2 reductants
0%
PROD/IS
47%
example product and conditions
31
NiCl2•glyme (0.2 equiv.), 4,4'-dtbbpy (0.3 equiv.)
Zn (3 equiv.), LiClO4 (1 equiv.)
50:50 dioxane:MeCN (0.1 M)
60 °C, 24 h
E
amine acid sp³–sp³ C–C coupling
12 acids, 8 amines
0%
PROD/IS
100%
example product and conditions
32
NiBr2•glyme (0.2 equiv.)
4,4'-(CF3)2-2,2'-bipyridine (0.3 equiv.)
Mn (3 equiv.), 50:50 dioxane:MeCN (0.1 M)
60 °C, 24 h
F
amine acid sp³–sp² C–C coupling
8 ligands, 12 catalysts
0%
assay yield
54%
example product and conditions
33
NiBr2 (0.2 equiv.), RuCl3 (0.5 equiv)
5,5'-(Me)2-2,2'-bipyridine (0.2 equiv), phthalimide (2 equiv.)
Mn (3 equiv.), NMP (0.1 M)
80 °C, 24 h
G
ultraHTE amine acid exploratory esterification
48 acids, 8 amines, quadruplicate
0%
PROD/IS
27%
example product and conditions
34
lutidine (1.5 equiv.)
Cu(MeCN)4BF4 (1.0 equiv.)
PhCN (0.1 M)
25 °C, 20 h
H
ultraHTE amine acid esterification scope
96 acids, 4 amines, quadruplicate
0%
PROD/IS
100%
example product and conditions
35
collidine (1.5 equiv.)
Cu(MeCN)4BF4 (1.0 equiv.)
PhCN (0.1 M)
25 °C, 16 h
I
1,280 well M^Pro amide coupling direct-to-biology
80 amines, 2 bases, 2 additives, 2 coupling agents, duplicate
0%
PROD/IS
100%
example product and conditions
36
HATU (1.1 equiv.)
DIPEA (0.2 equiv.)
NMP (0.1 M)
25 °C, 24 h

Figure 4. (A)-(I) Novel chemistries discovered with reaction arrays designed with phactor™. All input and output files used to produce reaction arrays (A)-(F) are provided via online repository in addition to all compiled HTE results in a machine-readable format. Reaction schemes can be found in the Supporting Information.

Figure 4A shows an amide coupling reaction array performed in preparation for an ultra-high throughput direct-to-biology assay (vide infra), with the aim of producing inhibitors of the SARS-CoV-2 main protease (M^{Pro}).[62] A screen of three anilines, two coupling agents HATU or DCC/HOBt and a carboxylic acid, with or without base, produced all desired products. Amide **28** was produced in high conversion using HATU, DMAP, and DIPEA. Figure 4B and Figure 4C follow up on the deaminative decarboxylative sp^2–sp^2 C–C coupling described in Figure 2. Reaction array 4B screened four isoquinoline based acids against six diazonium salts in 1:1 DMSO:MeOH solution under blue light irradiation, generating product **29**. Product **30** was found when the corresponding acid and diazonium were treated with white light in 100% methanol as shown in Figure 4C. Figures 4D-4F contain 96-well experiments designed with phactor™. Figure 4D and Figure 4E show the results from a recently reported deaminative decarboxylative carbon–carbon cross coupling[53] that generate products **31** and **32**. To follow up on our initial report (ref. 55), the catalysts, four additives, four ligands, and two reductants were tested in Figure 4D, and it was found that **31** was afforded at 47% product/internal standard with $NiCl_2$•glyme as the catalyst, 4,4'-di-*tert*-butyl-2,2-bipyridine as ligand, zinc as the reductant, and lithium perchlorate as an additive. After additional optimization, the reaction array in Figure 4E was run to test the reaction's scope. Product **32** was afforded with optimal conditions of $NiBr_2$•glyme, 4,4'-bis-trifluoromethyl-2,2'-bipyridine, and manganese in 1:1 dioxane:acetonitrile. Figure 4F shows the results of a 12 ligand – eight metal catalyst screen expanding on a recently reported sp^2–sp^3 decarboxylative deaminative C–C coupling from an acid activated as a twisted amide and an amine activated as a Katritzky salt, optimizing the yield of model product **33**.[55] Figures 4G-4I show several ultraHTE reaction arrays. Figure 4G and Figure 4H are two reaction arrays produced towards our aryl amine esterification reaction. Both reaction arrays were substrate scope experiments, producing compounds such as **34** and **35.** In Figure 4G, lutidine and copper(I) BF_4 salt tetrakis acetonitrile in benzonitrile showed good scope, and further optimized conditions replacing lutidine with collidine expands the chemistry's scope and reproducibility across quadruplicate measurements as shown in Figure 4F. Finally, Figure 4I shows the chemistry results of a direct-to-biology reaction array towards the creation of novel SARS-CoV-2 M^{Pro} inhibitors, generating amides such as **36**. Files needed for all 24 and 96 well experiments shown in Figure 4 are provided as examples to learn the phactor™ workflow as well as the layout for each of these reaction arrays are displayed in the Supporting Information.

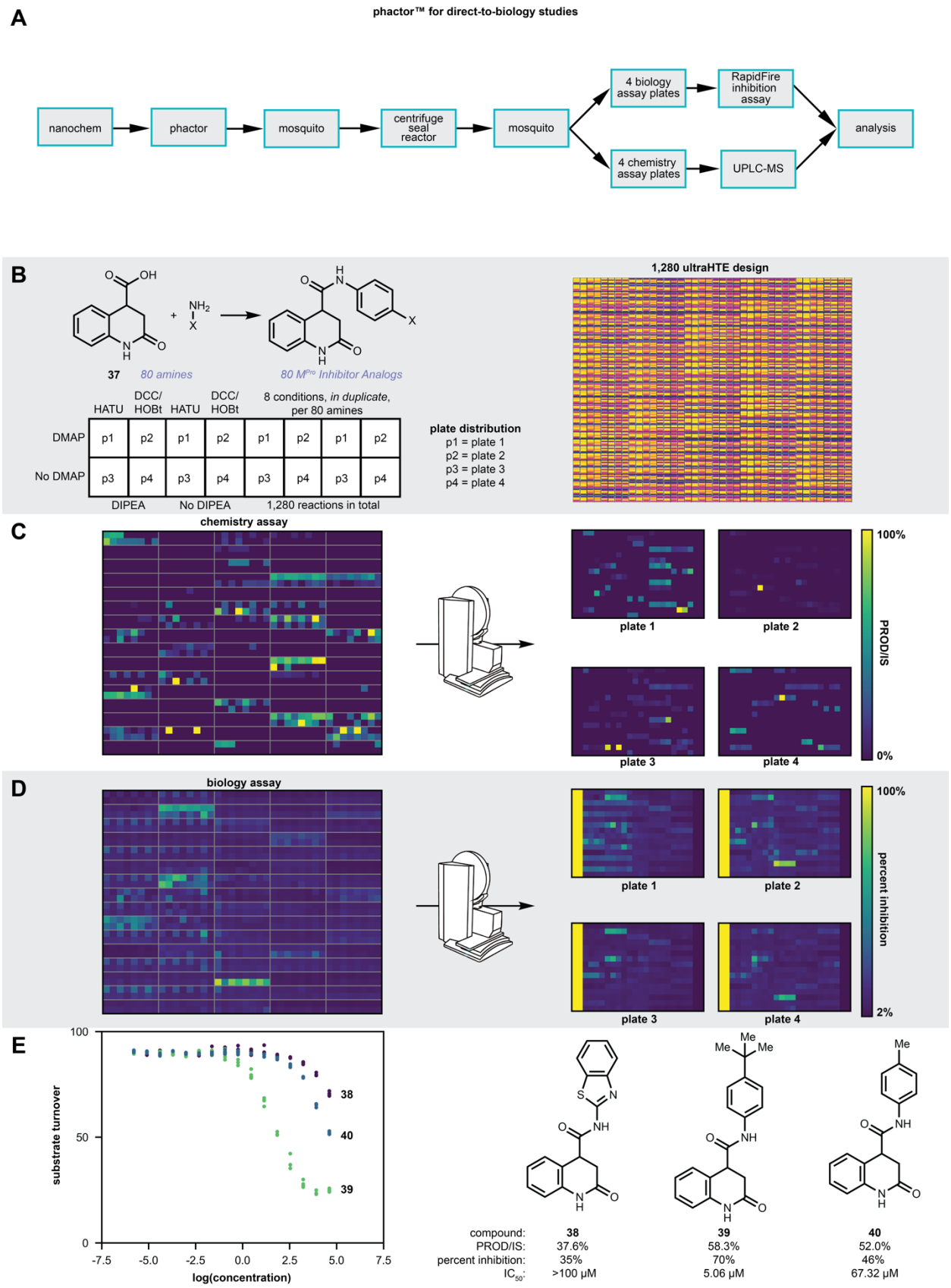


Figure 5. phactor™ facilitates ultraHTE direct-to-biology campaigns. (A) Event workflow for performing ultraHTE using

phactor™ and a Mosquito robot. (B) Design of 1,280 well amide coupling plate. 80 amines were selected to react with carboxylic acid **37.** Eight conditions were run in duplicate for each amine. (C) Results of the amide coupling shown as a product/internal standard integration ratio from a 2-minute LCMS injection of each well. The Mosquito robot is utilized to split the size 1,536 into four size 384 plates for LCMS and bioassay analysis. (D) Percent inhibition of SARS-CoV-2 M^{Pro} when treated with a sample of the reaction mixture from the corresponding well. The 1,280 plate is visually recreated. (E) IC_{50} curves for three scaled up compounds chosen from the reaction array. Compounds **38**-**40** display a range of assay and inhibitory responses.

Discovery of a low micromolar inhibitor of SARS-CoV-2 M^{Pro} via direct-to-biology assay. An amide coupling experiment was planned based on an inventory of amines, largely anilines, and a tetrahydroquinoline carboxylic acid pharmacophore found to be a potent inhibitor of M^{Pro}.[62] A preliminary 24-well amide coupling assay was performed to test the effectiveness of various amide conditions for the acid and anilines (see Figure 4A and Supporting Information). A diluted aliquot from each well was subjected to a RapidFire™ MS binding assay to determine concentration-response curves.[63] Curves were found to correlate with yield and literature reported IC_{50} values were replicated utilizing this assay. Following an optimized direct-to-biology workflow (Figure 5A), amide reactions were executed with the aim of making diverse amides in which were directly tested for activity against M^{Pro} in a single experiment. Eight amide conditions were tested in duplicate for each of the 80 amines, resulting in 1,280 reactions (Figure 5B). A key step in this workflow is the distribution of the 1,536-well reaction plate into four 384-well analysis assay plates suitable for UPLC-MS or RapidFire™ analysis. As such, some of the wells in the 1536-well reaction plate were not utilized to account for the four control columns necessary in each of the four 384-well RapidFire™ assay plates to allow calculation of Z' prime for the assay (0.961).[64] This distribution as well as the chemical and biological assay results are shown in Figure 5C and Figure 5D. Additional data analyses comparing chemical yield to biological response are shown in the Supporting Information. We note that reproducibility is a common concern in HTE and ultraHTE, and analyses of repeat experiments are provided in the Supporting Information. Both chemistry and biology assays are shown to be consistent, with 87% and 93% of data points having less than 10% error in the respective assays. From these analyses, three amides (**38**, **39**, and **40**) were chosen for scale up and IC_{50} determination, two of which (**38** and **39**) were previously unreported in the literature. Compound **39** was found to have a IC_{50} of 5.06 μM (Figure 5E), competitive with the best known M^{Pro} inhibitors in this series. Notably, scaled-up IC_{50} trends perfectly matched the percent inhibition trends obtained on the nanomole scale.

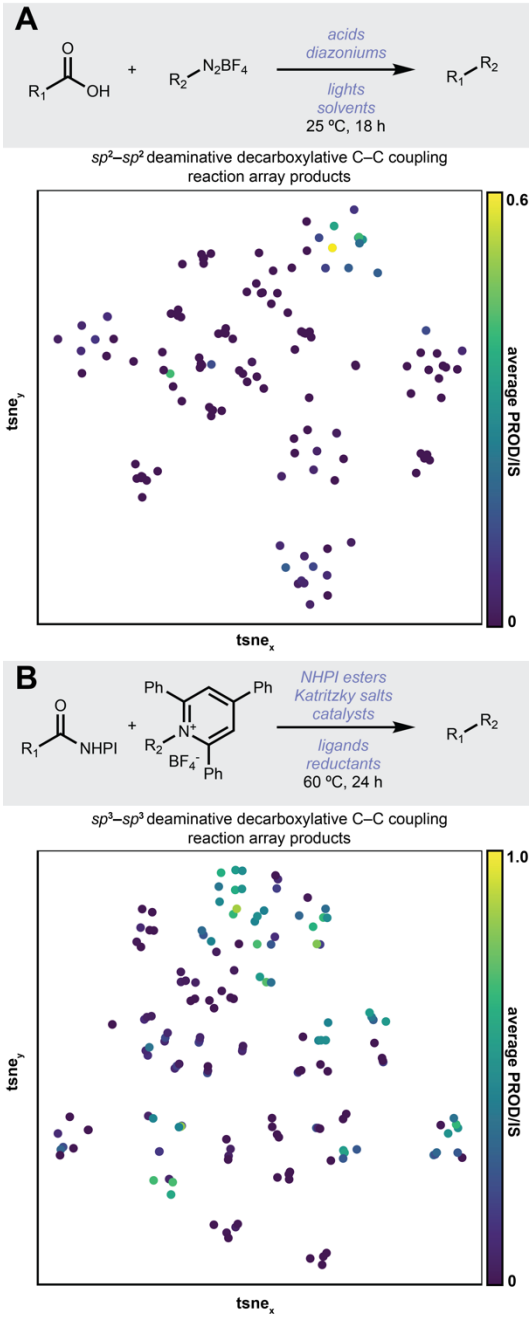


Figure 6. phactor™ enables rapid machine learning analysis of multiple reaction arrays in tandem. Standardized output files can be rapidly merged to create massive datasets. (A) tSNE of all products made in the catalyst-free light-enabled decarboxylative–deaminative sp^2–sp^2 C–C coupling detailed in Figure 2,

colored by average yield. (B) tSNE of all products made in the decarboxylative–deaminative sp^3–sp^3 C–C coupling detailed in ref. 53, colored by average yield.

Conclusion

We present the HTE ELN phactor™, which records all details of the entire experiment to allow for robust reproduction and accelerates discovery. Furthermore, the details are stored in a machine-readable yet tractable and interpretable format using an SQL database to facilitate the use of downstream algorithms. (Figure 6) As all reaction arrays are stored in a centralized database, bulk statistical analysis of multiple reaction arrays can be performed. phactor™ provides an exposed API that can be used to develop interfaces to other robots, assays, and software. Examples of various integrations and code infrastructures are shown in the Supporting Information. We hope that phactor™'s ease of use and further development provides increased accessibility to HTE in the chemistry community. Registration free and non-commercial use of phactor™ in 24- and 96-well formats is available through https://phactor.cernaklab.com/.

Supporting Information

Experimental procedures, characterization data, copies of all spectral data, instructions on using the software, sample files, sample interfacing code, and further discussion is provided free of charge via the Internet at https://pubs.acs.org.

Corresponding Author

* Tim Cernak – Department of Medicinal Chemistry, College of Pharmacy, University of Michigan, Ann Arbor, MI, USA 48104; orcid.org/0000-0001-5407-0643.

Email: tcernak@med.umich.edu

Author Contributions

B.M. wrote and deployed the software and developed, performed, and analyzed chemical and biological experiments. Y.S., Y.L., and R.Zhao. performed chemical experiments. O.G.M. performed the inhibition assay. B.M. and T.C. wrote the manuscript with input from all authors. T.C. conceived and supervised the project. All authors provided feedback on software development. All authors have given approval to the final version of the manuscript.

Funding Sources

We gratefully acknowledge funding from MilliporeSigma (Y.L.), Relay Therapeutics (J.D.), Janssen Therapeutics (A.M.), an ACS MEDI Predoctoral Fellowship (B.M.), as well as generous gifts of liquid handling automation from SPT Labtech and Merck & Co., Inc. Initial studies were funded by startup funds from the University of Michigan College of Pharmacy.

Notes

T.C. holds equity in Scorpion Therapeutics, and is a co-Founder and equity holder of Entos, Inc.

Acknowledgements

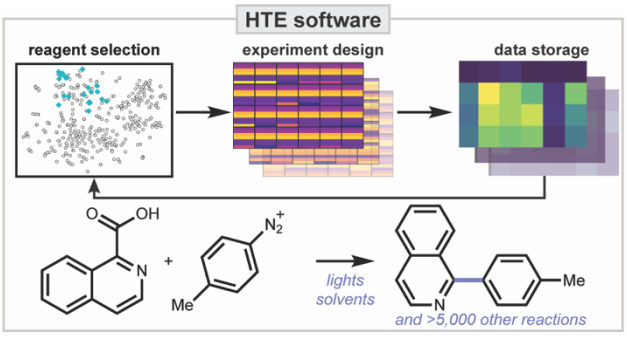
O.G.M and A.T. thank Michigan Drug Discovery for supporting the M^{Pro} inhibition screening assay (MDD-22105S to A.T.). We would like to thank the Center for Structural Biology and the Proteomics and Peptide Synthesis core at the University of Michigan for supplying the M^{Pro} enzyme and the substrate peptide for the M^{Pro} inhibition assay, respectively. We would like to thank the Center for Chemical Genomics at the University of Michigan for providing access to different instruments.

References

1. Mahjour, B.; Shen, Y.; Cernak, T.; Ultrahigh-Throughput Experimentation for Information-Rich Chemical Synthesis. *Acc. Chem. Res.* **2021**, *54*, 2337-2346.
2. Shen, Y.; Borowski, J. E.; Hardy, M. A.; Sarpong, R.; Doyle, A. G.; Cernak, T.; Automation and computer-assisted planning for chemical synthesis. *Nature Reviews Methods Primers* **2021**, *1*, 23.
3. Mennen, S. M.; Alhambra, C.; Allen, C. Liana; Barberis, M.; Berritt, S.; Brandt, T. A.; Campbell, A. D.; Castañón, J.; Cherney, A. H.; Christensen, M.; Damon, D. B.; Eugenio d.; García-Cerrada, S.; García-Losada, P.; Haro, R.; Janey, J.; Leitch, D. C.; Li, L.; Liu, F.; Lobben, P. C.; MacMillan, D.; Magano, J.; McInturff, E.; Monfette, S.; Post, R. J.; Schultz, D.; Sitter, B. J.; Stevens, J. M.; Strambeanu, I. I.; Twilton, J.; Wang, K.; Zajac, M. A.; The Evolution of High-Throughput Experimentation in Pharmaceutical Development and Perspectives on the Future. *Org. Process Res. Dev.* **2019**, *23*, 1213-1242.
4. Wong, H.; Cernak, T.; Reaction miniaturization in eco-friendly solvents. *Current Opinion in Green and Sustainable Chemistry* **2018**, *11*, 91-98.
5. Uehling, M. R.; King, R. P.; Krska, S. W.; Cernak, T.; Buchwald, S. L.; Pharmaceutical diversification via palladium oxidative addition complexes. *Science* **2019**, *363*, 405-408.
6. Lin, S.; Dikler, S.; Blincoe, W. D.; Ferguson, R. D.; Sheridan, R. P.; Peng, Z.; Conway, D. V.; Zawatzky, K.; Wang, H.; Cernak, T.; Davies, I. W.; DiRocco, D. A.; Sheng, H.; Welch, C. J.; Dreher, S. D.; Mapping the dark space of chemical reactions with extended nanomole synthesis and MALDI-TOF MS. *Science* **2018**, *361*, eaar6236.
7. Gesmundo, N. J.; Sauvagnat, B.; Curran, P. J.; Richards, M. P.; Andrews, C. L.; Dandliker, P. J.; Cernak, T.; Nanoscale synthesis and affinity ranking. *Nature* **2018**, *557*, 228-232.
8. Cernak, T.; Gesmundo, N. J.; Dykstra, K.; Yu, Y.; Wu, Z.; Shi, Z.; Vachal, P.; Sperbeck, D.; He, S.; Murphy, B. Ann; Sonatore, L.; Williams, S.; Madeira, M.; Verras, A.; Reiter, M.; Lee, C. Heechoon; Cuff, J.; Sherer, E. C.; Kuethe, J.; Goble, S.; Perrotto, N.; Pinto, S.; Shen, D.; Nargund, R.; Balkovec, J.; DeVita, R. J.; Dreher, S. D.; Microscale High-Throughput Experimentation as an Enabling Technology in Drug Discovery: Application in the Discovery of (Piperidinyl)pyridinyl-1H-benzimidazole Diacylglycerol Acyltransferase 1 Inhibitors. *J. Med. Chem.* **2017**, *60*, 3594-3605.
9. Kutchukian, P. S.; Dropinski, J. F.; Dykstra, K. D.; Li, B.; DiRocco, D. A.; Streckfuss, E. C.; Campeau, L.; Cernak, T.; Vachal, P.; Davies, I. W.; Krska, S. W.; Dreher, S. D.; Chemistry informer libraries: a chemoinformatics enabled approach to evaluate and advance synthetic methods. *Chem. Sci.* **2016**, *7*, 2604-2613.
10. Buitrago S. Alexander; Regalado, E. L.; Pereira, T.; Shevlin, M.; Bateman, K.; Campeau, L.; Schneeweis, J.; Berritt, S.; Shi, Z.; 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*, 49-53.
11. Gesmundo, N.; Dykstra, K.; Douthwaite, J.; Mahjour, B.; Ferguson, R.; Dreher, S.; Sauvagnat, B.; Sauri, J.; Cernak, T.; Miniaturization of Popular Reactions from the Medicinal Chemists' Toolbox for Ultrahigh-Throughput Experimentation. *ChemRxiv* **2022**.
12. Krska, S. W.; DiRocco, D. A.; Dreher, S. D.; Shevlin, M.; The Evolution of Chemical High-Throughput Experimentation To Address Challenging Problems in Pharmaceutical Synthesis. *Acc. Chem. Res.* **2017**, *50*, 2976-2985.
13. Wagen, C. C.; McMinn, S. E.; Kwan, E. E.; Jacobsen, E. N.; Screening for generality in asymmetric catalysis. *Nature* **2022**, *610*, 680-686.
14. Shevlin, M.; Practical High-Throughput Experimentation for Chemists. *ACS Med. Chem. Lett.* **2017**, *8*, 601-607.
15. Żurański, A. M.; Martinez A.; Shields, B. J.; Doyle, A. G.; Predicting Reaction Yields via Supervised Learning. *Acc. Chem. Res.* **2021**, *54*, 1856-1865.
16. Schleinitz, J.; Langevin, M.; Smail, Y.; Wehnert, B.; Grimaud, L.; Vuilleumier, R.; Machine Learning Yield Prediction from NiCOlit, a Small-Size Literature Data Set of Nickel Catalyzed C–O Couplings. *J. Am. Chem. Soc.* **2022**, *144*, 14722-14730.
17. Stevens, J. M.; Li, J.; Simmons, E. M.; Wisniewski, S. R.; DiSomma, S.; Fraunhoffer, K. J.; Geng, P.; Hao, B.; Jackson, E. W.; Advancing Base Metal Catalysis through Data Science: Insight and Predictive Models for Ni-Catalyzed Borylation through Supervised Machine Learning. *Organometallics* **2022**, *41*, 1847-1864.
18. Lexa, K. W.; Belyk, K. M.; Henle, J.; Xiang, B.; Sheridan, R. P.; Denmark, S. E.; Ruck, R. T.; Sherer, E. C.; Application of Machine Learning and Reaction Optimization for the Iterative Improvement of Enantioselectivity of Cinchona-Derived Phase Transfer Catalysts. *Org. Process Res. Dev.* **2022**, *26*, 670-682.
19. Beker, W.; Roszak, R.; Wołos, A.; Angello, N. H.; Rathore, V.; Burke, M. D.; Grzybowski, B. A.; Machine Learning May Sometimes Simply Capture Literature Popularity Trends: A Case Study of Heterocyclic Suzuki–Miyaura Coupling. *J. Am. Chem. Soc.* **2022**, *144*, 4819-4827.
20. Kearnes, S. M.; Maser, M. R.; Wleklinski, M.; Kast, A.; Doyle, A. G.; Dreher, S. D.; Hawkins, J. M.; Jensen, K. F.; Coley, C. W.; The Open Reaction Database. *J. Am. Chem. Soc.* **2021**, *143*, 18820-18826.
21. Kanza, S.; Willoughby, C.; Gibbins, N.; Whitby, R.; Frey, J. Graham; Erjavec, J.; Zupančič, K.; Hren, M.; Kovač, K.; Electronic lab notebooks: can they replace paper?. *Journal of Cheminformatics* **2017**, *9*, 31.
22. Higgins, S. G.; Nogiwa-Valdez, A. A.; Stevens, M. M.; Considerations for implementing electronic laboratory notebooks in an academic research environment. *Nature Protocols* **2022**, *17*, 179-189.
23. Tremouilhac, P.; Nguyen, A.; Huang, Y.; Kotov, S.; Lütjohann, D. Sebastian; Hübsch, F.; Jung, N.; Bräse, S.; Chemotion ELN: an Open Source electronic lab notebook for chemists in academia. *Journal of Cheminformatics* **2017**, *9*, 54.
24. Barillari, C.; Ottoz, D.; Fuentes-Serna, J. Mariano; Ramakrishnan, C.; Rinn, B.; Rudolf, F.; openBIS ELN-LIMS: an open-source database for academic laboratories. *Bioinformatics* **2016**, *32*, 638-640.
25. Cole, J. M.; The chemistry of errors. *Nature Chemistry* **2022**, *14*, 973-975.
26. Christ, C. D.; Zentgraf, M.; Kriegl, J. M.; Mining Electronic Laboratory Notebooks: Analysis, Retrosynthesis, and Reaction Based Enumeration. *J. Chem. Inf. Model.* **2012**, *52*, 1745-1756.
27. Yano, J.; Gaffney, K. J.; Gregoire, J.; Hung, L.; Ourmazd, A.; Schrier, J.; Sethian, J. A.; Toma, F. M.; The case for data science in experimental chemistry: examples and recommendations. *Nature Reviews Chemistry* **2022**, *6*, 357-370.
28. Roch, L. M.; Häse, F.; Kreisbeck, C.; Tamayo-Mendoza, T.; Yunker, L.; Hein, J. E.; Aspuru-Guzik, A.; ChemOS:

Orchestrating autonomous experimentation. *Science Robotics* **2018**, *3*, eaat5559.
29. Christensen, M.; Yunker, L.; Adedeji, F.; Häse, F.; Roch, L. M.; Gensch, T.; dos P.; Zepel, T.; Sigman, M. S.; Aspuru-Guzik, A.; Hein, J. E.; Data-science driven autonomous process optimization. *Communications Chemistry* **2021**, *4*, 112.
30. Häse, F.; Aldeghi, M.; Hickman, R. J; Roch, L. M; Christensen, M.; Liles, E.; Hein, J. E; Aspuru-Guzik, A.; Olympus: a benchmarking framework for noisy optimization and experiment planning. *Machine Learning: Science and Technology* **2021**, *2*, 035021.
31. Shiri, P.; Lai, V.; Zepel, T.; Griffin, D.; Reifman, J.; Clark, S.; Grunert, S.; Yunker, L. P.E.; Steiner, S.; Situ, H.; Yang, F.; Prieto, P. L.; Hein, J. E.; Automated solubility screening platform using computer vision. *iScience* **2021**, *24*, 102176.
32. Shields, B. J.; Stevens, J.; Li, J.; Parasram, M.; Damani, F.; Alvarado, J.; Janey, J. M.; Adams, R. P.; Doyle, A. G.; Bayesian reaction optimization as a tool for chemical synthesis. *Nature* **2021**, *590*, 89-96.
33. Daponte, J. A.; Guo, Y.; Ruck, R. T.; Hein, J. E.; Using an Automated Monitoring Platform for Investigations of Biphasic Reactions. *ACS Catal.* **2019**, *9*, 11484-11491.
34. Lu, J.; Paci, I.; Leitch, D. C.; A broadly applicable quantitative relative reactivity model for nucleophilic aromatic substitution (SNAr) using simple descriptors. *Chem. Sci.* **2022**, *13*, 12681-12695.
35. Lu, J.; Donnecke, S.; Paci, I.; Leitch, D. C.; A reactivity model for oxidative addition to palladium enables quantitative predictions for catalytic cross-coupling reactions. *Chem. Sci.* **2022**, *13*, 3477-3488.
36. Allen, C. Liana; Leitch, D. C.; Anson, M. S.; Zajac, M. A.; The power and accessibility of high-throughput methods for catalysis research. *Nature Catalysis* **2019**, *2*, 2-4.
37. 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*, 399-405.
38. Molga, K.; Szymkuć, S.; Gołębiowska, P.; Popik, O.; Dittwald, P.; Moskal, M.; Roszak, R.; Mlynarski, J.; Grzybowski, B. A.; A computer algorithm to discover iterative sequences of organic reactions. *Nature Synthesis* **2022**, *1*, 49-58.
39. Molga, K.; Szymkuć, S.; Grzybowski, B. A.; Chemist Ex Machina: Advanced Synthesis Planning by Computers. *Acc. Chem. Res.* **2021**, *54*, 1094-1106.
40. Mikulak-Klucznik, B.; Gołębiowska, P.; Bayly, A. A.; Popik, O.; Klucznik, T.; Szymkuć, S.; Gajewska, E. P.; Dittwald, P.; Staszewska-Krajewska, O.; Beker, W.; Badowski, T.; Scheidt, K. A.; Molga, K.; Mlynarski, J.; Mrksich, M.; Grzybowski, B. A.; Computational planning of the synthesis of complex natural products. *Nature* **2020**, *588*, 83-88.
41. Badowski, T.; Gajewska, E. P.; Molga, K.; Grzybowski, B. A.; Synergy Between Expert and Machine-Learning Approaches Allows for Improved Retrosynthetic Planning. *Angewandte Chemie International Edition* **2020**, *59*, 725-730.
42. Gajewska, E. P.; Szymkuć, S.; Dittwald, P.; Startek, M.; Popik, O.; Mlynarski, J.; Grzybowski, B. A.; Algorithmic Discovery of Tactical Combinations for Advanced Organic Syntheses. *Chem* **2020**, *6*, 280-293.
43. Jaworski, W.; Szymkuć, S.; Mikulak-Klucznik, B.; Piecuch, K.; Klucznik, T.; Kaźmierowski, M.; Rydzewski, J.; Gambin, A.; Grzybowski, B. A.; Automatic mapping of atoms across both simple and complex chemical reactions. *Nature Communications* **2019**, *10*, 1434.
44. Grzybowski, B. A; Szymkuć, S.; Gajewska, E. P; Molga, K.; Dittwald, P.; Wołos, A.; Klucznik, T.; Chematica: a story of computer code that started to think like a chemist. *Chem* **2018**, *4*, 390-398.
45. Gallegos, L. C.; Luchini, G.; St. J.; Kim, S.; Paton, R. S.; Importance of Engineered and Learned Molecular Representations in Predicting Organic Reactivity, Selectivity, and Chemical Properties. *Acc. Chem. Res.* **2021**, *54*, 827-836.
46. Torres, J.; Lau, S. Hong; Anchuri, P.; Stevens, J. M.; Tabora, J. E.; Li, J.; Borovika, A.; Adams, R. P.; Doyle, A. G.; A Multi-Objective Active Learning Platform and Web App for Reaction Optimization. *J. Am. Chem. Soc.* **2022**, *144*, 19999-20007.
47. Reker, D.; Hoyt, E. A.; Bernardes, G. J.L.; Rodrigues, T.; Adaptive Optimization of Chemical Reactions with Minimal Experimental Information. *Cell Reports Physical Science* **2020**, *1*, 100247.
48. Hammer, A.; Leonov, A. I.; Bell, N. L.; Cronin, L.; Chemputation and the Standardization of Chemical Informatics. *JACS Au* **2021**, *1*, 1572-1587.
49. Gensch, T.; dos P.; Friederich, P.; Peters, E.; Gaudin, T.; Pollice, R.; Jorner, K.; Nigam, A.; Lindner-D'Addario, M.; Sigman, M. S.; Aspuru-Guzik, A.; A Comprehensive Discovery Platform for Organophosphorus Ligands for Catalysis. *J. Am. Chem. Soc.* **2022**, *144*, 1205-1217.
50. Mahjour, B.; Shen, Y.; Liu, W.; Cernak, T.; A map of the amine–carboxylic acid coupling system. *Nature* **2020**, *580*, 71-75.
51. Shen, Y.; Mahjour, B.; Cernak, T.; Development of copper-catalyzed deaminative esterification using high-throughput experimentation. *Communications Chemistry* **2022**, *5*, 83.
52. Lin, Y.; Zhang, Z.; Mahjour, B.; Wang, D.; Zhang, R.; Shim, E.; McGrath, A.; Shen, Y.; Brugger, N.; Turnbull, R.; Trice, S.; Jasty, S.; Cernak, T.; Reinforcing the supply chain of umifenovir and other antiviral drugs with retrosynthetic software. *Nature Communications* **2021**, *12*, 7327.
53. Zhang, Z.; Cernak, T.; The Formal Cross-Coupling of Amines and Carboxylic Acids to Form sp3–sp3 Carbon–Carbon Bonds. *Angewandte Chemie International Edition* **2021**, *60*, 27293-27298.
54. McGrath, A.; Zhang, R.; Shafiq, K.; Cernak, T.; Repurposing amine and carboxylic acid building blocks with an automatable esterification reaction. *ChemRxiv* **2022**.
55. Douthwaite, J. L; Zhao, R.; Shim, E.; Mahjour, B.; Zimmerman, P.; Cernak, T.; The Formal Cross-Coupling of Amines and Carboxylic Acids to Form sp3–sp2 Carbon–Carbon Bonds. *ChemRxiv* **2022**.
56. Lu, P.; Sanchez, C.; Cornella, J.; Larrosa, I.; Silver-Catalyzed Protodecarboxylation of Heteroaromatic Carboxylic Acids. *Org. Lett.* **2009**, *11*, 5710-5713.
57. Gooßen, L. J.; Linder, C.; Rodríguez, N.; Lange, P. P.; Fromm, A.; Silver-catalysed protodecarboxylation of carboxylic acids. *Chem. Commun.* **2009**. 7173-7175.
58. Dow, N. W.; Pedersen, P. Scott; Chen, T. Q.; Blakemore, D. C.; Dechert-Schmitt, A.; Knauber, T.; MacMillan, D.; Decarboxylative Borylation and Cross-Coupling of (Hetero)aryl Acids Enabled by Copper Charge Transfer Catalysis. *J. Am. Chem. Soc.* **2022**, *144*, 6163-6172.
59. Xu, P.; López-Rojas, P.; Ritter, T.; Radical Decarboxylative Carbometalation of Benzoic Acids: A Solution to Aromatic

Decarboxylative Fluorination. *J. Am. Chem. Soc.* **2021**, *143*, 5349-5354.
60. Witzel, S.; Hoffmann, M.; Rudolph, M.; Kerscher, M.; Comba, P.; Dreuw, A.; Hashmi, A.; Excitation of aryl cations as the key to catalyst-free radical arylations. *Cell Reports Physical Science* **2021**, *2*, 100325.
61. Wang, C.; Rakshit, S.; Glorius, F.; Palladium-Catalyzed Intermolecular Decarboxylative Coupling of 2-Phenylbenzoic Acids with Alkynes via C−H and C−C Bond Activation. *J. Am. Chem. Soc.* **2010**, *132*, 14006-14008.
62. Rossetti, G. G.; Ossorio, M. A.; Rempel, S.; Kratzel, A.; Dionellis, V. S.; Barriot, S.; Tropia, L.; Gorgulla, C.; Arthanari, H.; Thiel, V.; Mohr, P.; Gamboni, R.; Halazonetis, T. D.; Non-covalent SARS-CoV-2 Mpro inhibitors developed from in silico screen hits. *Scientific Reports* **2022**, *12*, 2505.
63. Malla, T. R.; Tumber, A.; John, T.; Brewitz, L.; Strain-Damerell, C.; Owen, C. David; Lukacik, P.; Chan, H.; Maheswaran, P.; Salah, E.; Duarte, F.; Yang, H.; Rao, Z.; Walsh, M. A.; Schofield, C. J.; Mass spectrometry reveals potential of β-lactams as SARS-CoV-2 Mpro inhibitors. *Chem. Commun.* **2021**, *57*, 1430-1433.
64. Zhang, J.; Chung, T. DY; Oldenburg, K. R; A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *Journal of biomolecular screening* **1999**, *4*, 67-73.



For Table of Contents Only