phactor™ – a high throughput experimentation management system

Babak Mahjour[1], Tim Cernak[1]
[1]Department of Medicinal Chemistry, University of Michigan

High throughput experimentation (HTE) is an increasingly important tool in the study of chemical synthesis. While the hardware for running HTE in the synthesis lab has evolved significantly in recent years, there remains a need for software solutions to navigate data-rich experiments. We have developed the software, phactor™, to facilitate the performance and analysis of HTE in a chemical laboratory. phactor™ allows experimentalists to rapidly design arrays of chemical reactions in 24, 96, 384, or 1,536 wellplates. Users can access online reagent data, such as a lab inventory, to populate wells with experiments and produce instructions to perform the screen manually, or with the assistance of a liquid handling robot. After completion of the screen, 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 a machine-readable format.

As chemical synthesis is asked to answer increasingly complex challenges, tactics that increase reaction information density are required. Miniaturized high-throughput experimentation (HTE) has emerged as an accessible and reliable technique for the rapid execution of reaction arrays.[1-8] While the community has gravitated to liquid handling techniques in glass shell microvials with tumble stir dowels, or in plastic 384 or 1,536 wellplates, a standard for data handling has not yet been established. New tools are required to carry data from hypothesis to analysis. The organizational load required to perform a simple 24 well screen is generally manageable by repetitive notebook entries, or with spreadsheets, yet managing multiple screens in a single day, or running ultraHTE in 1,536 wellplates, is challenging without an information management software. Furthermore, no readily available lab notebook can store HTE details in a tractable manner. Available electronic notebooks may allow for custom experiments, but do not provide a simple interface to extract data and results from multiple experiments at once. To continue developing HTE research and position data outputs for machine learning studies, detailed reaction data must be easily accessible for rapid extraction and analysis. Furthermore, minimizing time and resources spent from experiment ideation to understanding the results enables creativity and facilitates reaction discovery and optimization. With these issues in mind, the software phactor™ was developed. phactor™ is a custom in-house HTE solution. Built using a React/Flask framework, phactor™ is a full-stack program that is accessible from any computer with a browser and an internet connection. The software streamlines the collection of manual or automated miniaturized 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 it in a standardized, machine-readable format.

The workflow of a typical HTE experiment involves design of the reaction array, preparation of reagent stock solutions, manual or robotic dosing of stock solutions according to the reaction array recipe, analysis of reaction outcome – in our case by UPLC-MS – followed by visualization of data and documentation of results (Figure 1). With phactor™, we sought to maximize automation of data movement and processing, decreasing the time needed to run a screening experiment, and to facilitate the analysis of the campaigns. While maximal automation was a goal, flexibility was required to accommodate last minute decisions in the face of unforeseen circumstances, such as when a chemical is poorly soluble, when a chemical must be
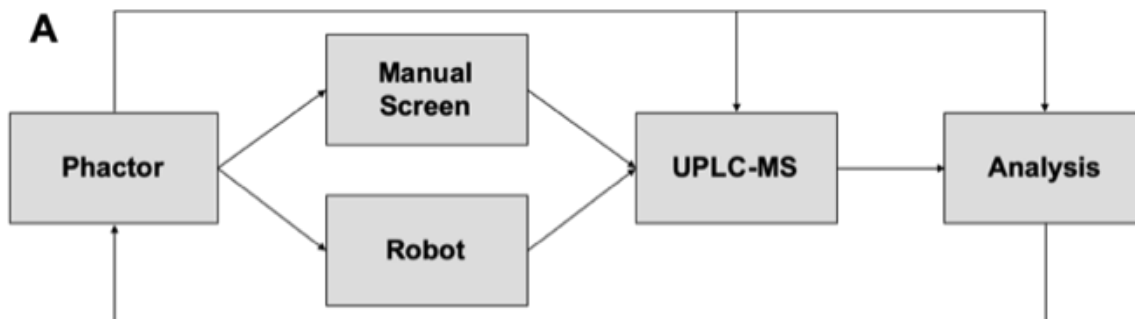
Figure 1. Workflow overview.

replaced, or multiple reagents must be premixed before dosage.

The software was developed to interface with our lab's inventory, which contains metadata such as reagent sub-location in the lab, molecular weight, SMILES string, etc. Any inventory can be connected. The user selects desired reagents from the inventory for automatic field population, or enters specific reagent entries manually, such as for a custom starting substrate. Once all relevant chemicals to the desired screen are selected, the plate layout can be designed automatically or manually, and workflow instructions are generated to be executed either manually, or by an interfacing liquid handling robot. Stock solutions are prepared, and chemicals are distributed to their respective locations on the reaction wellplate, and the reactions are run for the designated amount of time, the reactions can be quenched manually or robotically using an automatically generated quenching script, and prepared for UPLC-MS analysis. Once analysis is complete, UPLC-MS results can be uploaded onto phactor™ for visual analysis and the generation of a final report, which contains information and results from the experiment.

All experiments can be accessed and edited at any point in time, and all screen information and results are stored in an SQL database. The compilation of screen data in a machine-readable format via SQL allows for the bulk analysis and evaluation of screening results, providing ease of access to data for the development of data science tools, algorithms, applications, and visualizations that inform chemical synthesis.

Settings

Users navigate to phactor™ using a web browser. At present, phactor™ has not been made available for public use. The first page the user interacts with is the Settings tab, where the user can access existing experiments or design new screen recipes (Figure 2). The first option within the Settings tab, called create, allows for the creation of a screen: here the user is prompted for a screen name, screen density (in wells, for example 24, 96, etc.), and vial size. The chosen screen name is used to label all relevant files and documents (UPLC input/output files, robotic script files, report and electronic lab notebook name, etc.) generated for the experiment. Options available for screen density are 24 or 96 wells, with separate workflows available for 384 or 1,536 well experiments. Vial sizes are generally 25 μL or 100 μL but can be any volume. Well plates or reactor blocks can be partially filled, for example if the user wishes to run only 53 experiments in a 96 well reactor block. Alternatively, the user can load a previously created screen using the load option. This allows for experiments to be easily sharable, and reviewed at a later date. A duplicate option is also available, which allows the user to create a new screen based on a previous screen. This latter option saves time when experiments can be stored in a common template, for instance when a researcher runs multiple related screens in a single research campaign. Common recipes such as catalyst/ligand screens can be

Figure 2. phactor™ Settings Page. New screens can be created, or loaded from a template, at this interface.

templated and used as a foundation, with customization to suit the individual experiment. Templates can also be based on physical screening 'kits', which have chemicals pre-weighed into wellplates, and are generally directly purchasable from chemical suppliers, or prepared in bulk for storage.

Chemicals

The user clicks "Next" to arrive at the "Chemicals" page (Figure 3a). Here, the chemist specifies the chemicals and solvents being used in the screen. Chemicals can be added from an inventory, or manually by providing a name, SMILES (optionally), molecular weight, density, amount to prepare, and molarity. Name is the identifier for the chemical and is usually the chemical name. Density refers to the expected or known density of the solute, molarity is the desired molarity of the reagent in the final reaction mixture, and amount to prepare is a multiplier to create overage for the stock solution – the default is set to 1.2 for a 20% overage to serve as a dead volume for liquid handling. Chemicals can be inputted manually, but generally the user will add chemicals automatically if the desired reagent exists in the lab's inventory database. Adding a chemical from the inventory database automatically populates

the name, molar mass, density, and SMILES metadata.

To facilitate reaction informatics studies, each chemical must be qualified by a "Type" descriptor specifying its role in the reaction, such as catalyst, solvent, acid, base, electrophile, nucleophile, ligand, salt, or other. As well, an order of addition can be specified to describe in which order the chemicals should be added to the well. Metadata values for temperature, stir rate, and atmosphere can be changed from their defaults set at this stage. In addition to metadata, expected products of each well can be specified by uploading a comma separated values (.csv) file. The .csv contains a well location and a product SMILES, or concatenated list of SMILES, and name expected at that well location. This is used later for analyzing the plate and determining the yield of the desired product.

Factors

The next stage prepares the experimental design of the screen on the "Factors" page (Figure 3b). Depending on the size of the screen, various options are provided to "factorize" your selected chemicals and assign them automatically into appropriate wells. For example, a 24 well screen could be a 2 × 12 screen, where 12 different conditions, for example catalysts,

Figure 3. a, Sample Chemicals Page. Chemicals can be manually entered or selected from a prepopulated chemical inventory database. **b, Sample Factors Page.** Any factorization of the screen size can be selected to automatically generate the experimental design.

are screened in two other different conditions (e.g. solvents). Likewise, a 4 × 6 screen would screen 4 members of one condition against 6 members of another.

Factors are based on Shelvin's[9] description of reaction arrays, which assigns experimental variables as major or minor factors in a wellplate. Two factors account for experimental designs that test two conditions (for example 4 catalysts × 6 ligands, 3 × 8, 12 × 2, etc., in a 24 wellplate, 8 solvents × 12 bases in a 96 wellplate,). Of course, it is often the case that the experiment considers more variables – for example, 2 catalysts versus 2 solvents versus 2 acids versus 3 ligands. Any

factorization can be a valid experimental design as long as it doesn't exceed the wellplate real estate, and indeed phactor™ provides all possibilities and automatically distributes the chemicals into their correct experimental position on the wellplate.

Once the chemist chooses the appropriate factorization of the screen and assigns a chemical to its respective factor group, chemicals are distributed into their associated location specified by the factorization. In this stage, the calculation and distribution of the entire wellplate is reduced to a single button click. Recipes for stock solutions and distribution instructions
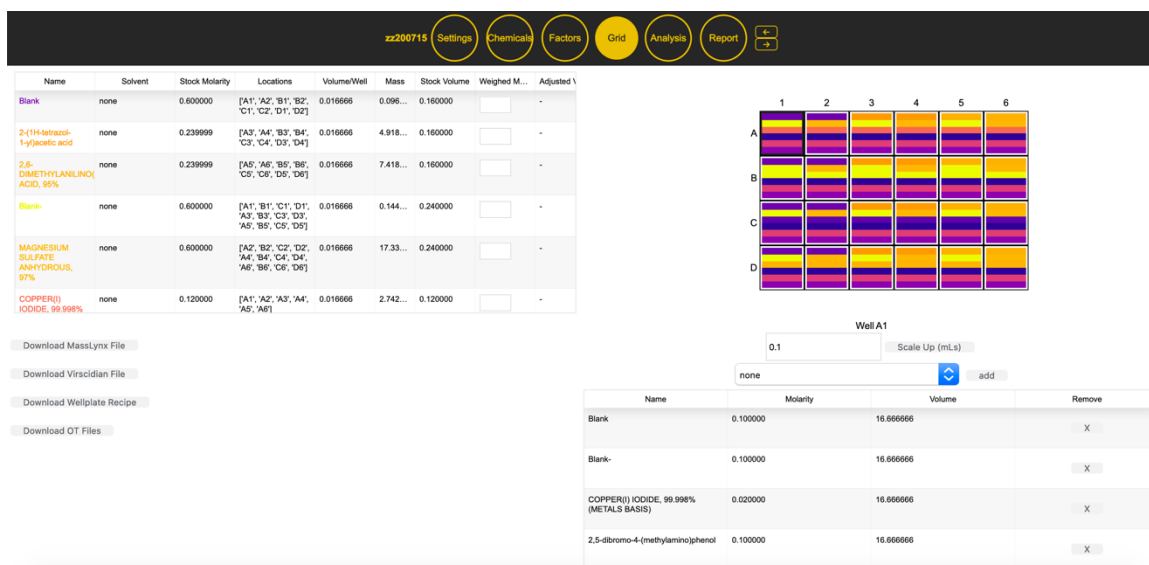
4

Figure 4. Sample Grid Page. Stock solution recipes are displayed alongside a visualization of the wellplate. Individual or bulk changes can be made to the screen recipe by interacting with the grid. Scaled reaction recipes can be generated by selecting single wells. Automatically populated input files for analytical instruments and scripts for dosing robots can be downloaded at this stage.

are prepared, as well as input files for initiating UPLC-MS analysis. Stock solution preparation and distribution instructions can be performed manually or interpreted by liquid handling platforms, and analytical instrumentation can be easily interfaced.

Grid

Once the calculations are complete, typically in <1 second for a 96 well screen, the final plate layout and stock solution recipes are displayed on the grid tab. The layout can be inspected and edited on a well by well basis or in a bulk fashion by dragging the mouse over rows, columns or sections of the plate to select wells for editing. The left side of the grid tab contains a table which lists every stock solution that needs to be prepared. Each row of the table displays the recipe for a stock solution, and contains information such as the final wellplate locations, the mass of reagent to be weighed, and volume of solvent needed to create the stock solution at the necessary concentration to deliver the right dose to each well. Typically, a user sends this tabulated list to a printer to have a physical copy of the recipe to bring to the bench while reagents are weighed.

The concentration of stock solutions is a function of the total volume of reagent in the whole wellplate. The distribution of reagent volumes in each well is proportional to the number of reagents in the well. As such, the volume of each reagent in a five-component reaction would be one-fifth of the total volume of the well (specified in the Settings stage). By summing the volume of each reagent in each well, the stock molarity can be back-calculated using the reaction molarity and the stock mass can then be calculated using the reagent's density and molecular weight.

The wellplate is drawn as a grid on the right half of the screen. For visualization, each well contains a stack of color bars, where each colored bar represents a chemical and the bar's height is proportional to the volume of its respective chemical in that well. Clicking on a well gives a detailed look at the well's chemical contents in a pop-up window. Each chemical in the well is displayed alongside its volume and molarity. Here, edits can be made to the well, such as adding or removing a chemical or modifying the target concentration. In both cases, volumes are recalculated across the whole wellplate to minimize the number of discrete stock solutions. A chemical can be added to

5

multiple wells at once by dragging across the wellplate to select multiple wells for distribution. A "scale-up" button provides a recipe of conditions in a single mouse-click to repeat targeted reactions on larger scale as a follow up on a screen hit, or for confirmation of reaction reproducibility.

Workflow

At this stage the screen design is complete and the chemist enters the wet lab to perform the experiment manually or robotically. Both workflows begin the same: reagents are weighed into appropriate source vials according to the calculated recipe. phactor™ allows the user to input the actual mass that was weighed, for instance if the recipe called for 12.8 mg but in practice the chemist weighed out 13.6 mg, to recalculate the solvent volume needed to achieve the desired stock molarity. Alternatively, if using robotic assistance, placing the vials in the correct location and running the automatically generated script for stock solution creation will initiate robotic solvent dosing into stock solution vials according to the solvent volumes calculated for the desired stock concentration.

A common pitfall of HTE happens at this stage. Poor solubility of the reagent in a solvent can lead to a viscous slurry or a two-phase mixture that may complicate distribution to reaction wells. Frequently, the situation can be resolved by 'slurry loading', which invokes rapid stirring of a suspension of finely divided solids, and use of a wide-bore pipette tip to distribute the stock suspension to designated wells. It is often worthwhile to investigate if combining reagents can improve solubility, for instance if an amine base such as pyridine is to be added to every well, it can often effectively solubilize other requisite reagents, so it may be desirable to premix the amine base into other stock solutions. Alternately, it is often necessary to dilute problematic stock suspensions two-fold to achieve dissolution, and then halve the concentration of complementary stock solutions or accept an overall lower final reaction concentration. Creating alternate conditions on the fly to

replace problematic stock solutions can be essential, and phactor™ is designed to facilitate these last-minute changes. Once the stock solutions are prepared, they are combinatorically dosed into their assigned location in the wellplate. Again, this can be done robotically or manually. After reagent dosing is complete, the wellplate is sealed and transferred to apparatus for stirring, heating, cooling or photoirradiation.

Upon reaction aging for a designated time, the plate is unsealed and a workup procedure is followed. Common workup operations and recipes are encoded as templates and saved for reuse. Next, an aliquot from each well is diluted and transferred to the appropriate plate for analysis, in our case UPLC-MS. Both workup and analysis plate preparations can be done robotically via a script automatically generated by phactor™. Furthermore, the UPLC-MS input files containing each well's relevant metadata such as molecular formula, well location, and sample name are automatically generated,. After the UPLC-MS analysis is complete, phactor™ will accept an output file from an analytical software like MassLynx or Virscidian, which serves as the input data for the Analysis and Report tabs.

Analysis and Report

Once uploaded to phactor™, the results from the analytical output is tied back to the conditions in each well. This data is visualized in an interactive heatmap triptych display that allows the user to evaluate three output results simultaneously. For example, a user may wish to simultaneously observe the measured amounts of desired product, undesired product and residual starting material in each well, and the simultaneous display of three heatmaps allows this. Clicking on a well shows the reaction
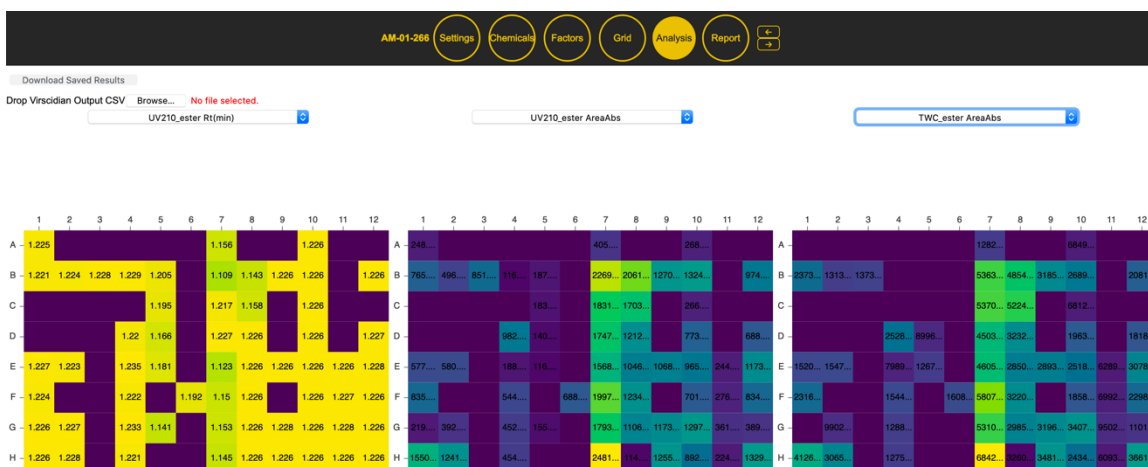
Figure 5. Sample Analysis Page. A variety of analytical outputs are displayed in triptych heatmap format. A cursor hover-over or mouse-click on individual wells displays reaction condition metadata.

conditions used in that well. At this stage, all data regarding the screen has been saved by phactor™ and is ready for bulk or individual screen analysis by accessing the SQL database.

Finally, phactor™ generates a report describing the screen and its results. The report is a printable and savable summary of the experiment that can be shared with peers or appended to a lab notebook. Here, the original metadata is displayed, alongside the experiment name and a scheme of the reaction. Also included is the original recipe for the well plate, with all of the stock solutions and their corresponding mass, volume, molarity, and well location information. Below the recipe list is a picture of the reaction plate from the Grid stage, as well as images of plate locations for the robot, if it was used, and a breakdown of the location of each stock solution and reaction well and their chemical components. Additionally, camera photos can be uploaded here, if used to document the experimental execution, and three heatmaps from the analysis tab can be selected for inclusion in the report. This entire document can be saved and added to an external electronic lab notebook as a .pdf file.

Case Studies

Discovery of a Deaminative Esterification

phactor™ was designed as a tool for discovery in chemical synthesis. As a demonstration, we describe the experimental design generated to achieve a preliminary hit for our recently reported sp^2–sp^2 deaminative esterification reaction.[10] In the screen design, an amine, activated as its diazonium salt (**1**), a carboxylic acid (**2**), 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 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 under nitrogen, then stirred at 60 ºC for 18 hours. Once the reaction was complete, a solution containing one molar equivalent of caffeine was added to each well as an internal standard, and an aliquot of each reaction was transferred to a plastic plate, then diluted with acetonitrile, for UPLC-MS analysis of the desired ester
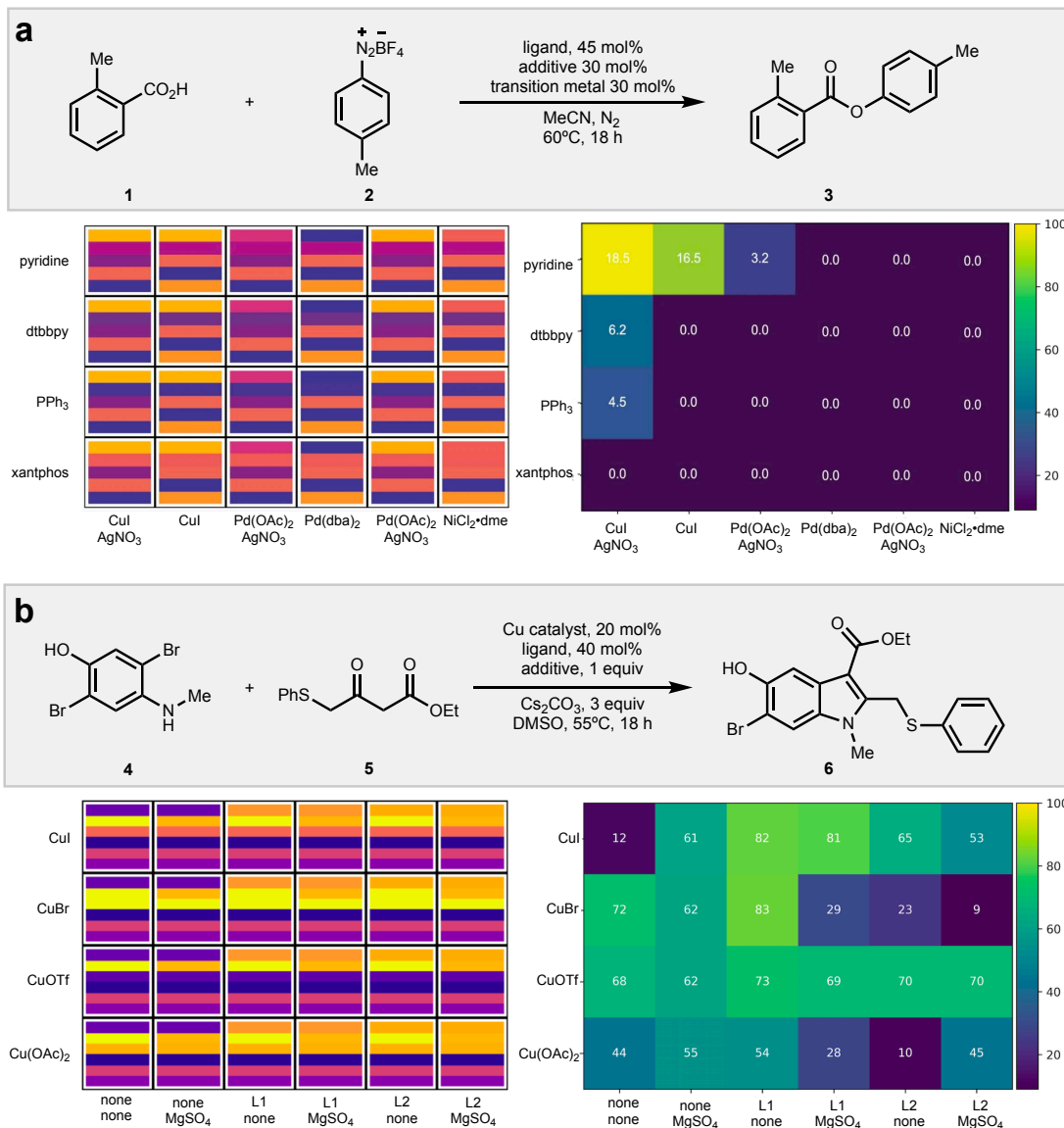
Figure 6. a, Preliminary esterification hit found using phactor^TM. The screen design and results are shown here as displayed on phactor^TM. **b, Optimized oxidative indolization conditions.**

product (**3**). 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^TM to record the experiment outcome and produce the resulting heatmap shown in Figure 6a. Analysis on phactor^TM indicated a 18.5% assay yield using 30 mol% CuI and AgNO$_3$, and these specific conditions were triaged for further study.

Optimization of Copper Catalyzed Coupling Towards the Synthesis of Umifenovir

In a second example, an oxidative indolization reaction screen, inspired by the conditions of Glorius,[11] was performed using copper catalysts and ligand/additive combinations to optimize the penultimate step towards a synthesis of umifenovir via the coupling of **4** and **5** to produce **6**.[12] Four copper catalysts, copper iodide, copper bromide, tetrakis(acetonitrile) copper(I) triflate, and copper 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 equiv. of cesium carbonate added to every well as a suspension in DMSO. The reactions were manually arrayed, sealed, and stirred at 55 ºC for 18 hours. Well A3 (copper bromide with **L1** and no magnesium sulfate) was found to have the best performing conditions, and a confirmation scale-up reaction performed on 0.10 mmol scale produced the desired indole **6** in 66% isolated yield.

When facilitated by phactor™, HTE becomes an exercise in workflow execution, as the organizational aspect of the experiment is handled by the software. This allows chemists to focus on the design and analysis of the screen, rather than details of its preparation. Importantly, phactor™ records all details of the entire experiment to allow for robust reproduction. Furthermore, the details are stored in a machine readable yet tractable and interpretable format using an SQL database to facilitate the development of downstream algorithms. As all screens are stored in a centralized database, bulk statistical analysis of multiple screens can be performed. As such, phactor™ is a framework for the bulk collection of analyzable data. phactor™ provides an API that is abstractable to any robot, and robotic integration will be discussed in future work.

Author Information
Babak Mahjour, Timothy Cernak
University of Michigan, Ann Arbor, MI, USA
Department of Medicinal Chemistry.

Correspondance

tcernak@med.umich.edu

References
1   Uehling, M. R., King, R. P., Krska, S. W., Cernak, T. & Buchwald, S. L. Pharmaceutical diversification via palladium oxidative addition complexes. *Science* **363**, 405, doi:10.1126/science.aac6153 (2019).
2   Lin, S. *et al.* Mapping the dark space of chemical reactions with extended nanomole synthesis and MALDI-TOF MS. *Science* **361**, eaar6236, doi:10.1126/science.aar6236 (2018).
3   Gesmundo, N. J. *et al.* Nanoscale synthesis and affinity ranking. *Nature* **557**, 228-232, doi:10.1038/s41586-018-0056-8 (2018).
4   Cernak, T. *et al.* Microscale High-Throughput Experimentation as an Enabling Technology in Drug Discovery: Application in the Discovery of (Piperidinyl)pyridinyl-1H-benzimidazole Diacylglycerol Acyltransferase 1 Inhibitors. *Journal of Medicinal Chemistry* **60**, 3594-3605, doi:10.1021/acs.jmedchem.6b01543 (2017).
5   Kutchukian, P. S. *et al.* Chemistry informer libraries: a chemoinformatics enabled approach to evaluate and advance synthetic methods. *Chemical Science* **7**, 2604-2613, doi:10.1039/C5SC04751J (2016).
6   Buitrago Santanilla, A. *et al.* Nanomole-scale high-throughput chemistry for the synthesis of complex molecules. *Science* **347**, 49, doi:10.1126/science.1259203 (2015).
7   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* **50**, 2976-2985, doi:10.1021/acs.accounts.7b00428 (2017).
8   Shevlin, M. Practical High-Throughput Experimentation for Chemists. *ACS Med Chem Lett* **8**, 601-607, doi:10.1021/acsmedchemlett.7b00165 (2017).
9   Shevlin, M. Practical High-Throughput Experimentation for Chemists. *ACS Medicinal Chemistry Letters* **8**, 601-607, doi:10.1021/acsmedchemlett.7b00165 (2017)
10   Mahjour, B., Shen, Y., Liu, W. & Cernak, T. A map of the amine–carboxylic acid coupling system. *Nature* **580**, 71-75, doi:10.1038/s41586-020-2142-y (2020).
11   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. *Journal of the American Chemical Society* **132**, 14006-14008, doi:10.1021/ja106130r (2010).

12   Lin, Y., Zirong. Z., Mahjour, B., Wang, D., Zhang, R., Shim, E., McGrath, A., Shen, Y., Brugger N., Turnbull R., Jasty, S., Trice, S., Cernak, T. Reinforcing the Supply Chain of COVID-19 Therapeutics with Expert-Coded Retrosynthetic Software. doi:https://doi.org/10.26434/chemrxiv.12765 410.v1 (2020).