ORGANIC PROCESS RESEARCH & DEVELOPMENT
OPR&D

pubs.acs.org/OPRD Article

Designing Chemical Reaction Arrays Using Phactor and ChatGPT

Babak Mahjour, Jillian Hoffstadt, and Tim Cernak*

Cite This: Org. Process Res. Dev. 2023, 27, 1510–1516 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: High-throughput experimentation is a common practice in the optimization of chemical synthesis. Chemists design reaction arrays to optimize the yield of couplings between building blocks. Popular reactions used in pharmaceutical research include the amide coupling, Suzuki coupling, and Buchwald–Hartwig coupling. We show how the artificial intelligence (AI) language model ChatGPT can automatically formulate reaction arrays for these common reactions based on the literature corpus it was trained on. Critically, we showcase how ChatGPT results can be directly translated into inputs for the management software phactor, which enables automated execution and analysis of assays. This workflow is experimentally demonstrated, with modest to excellent yields of products obtained in each instance on the first attempt.

KEYWORDS: *large language model, high-throughput experimentation, ChatGPT, reaction optimization, literature mining*

■ INTRODUCTION

Chemical synthesis is a primary bottleneck in drug development. High-throughput experimentation (HTE) is a widely practiced method for the discovery and optimization of reaction conditions in medicinal chemistry campaigns.[1−5] Chemists typically design reaction arrays based on conditions found in the literature using search tools such as Google, SciFinder, or Reaxys. The automated generation of reaction arrays to optimize or discover a coupling reaction between two substrates is a contemporary problem.[6−12] Recently, generative transformers, a form of artificial intelligence (AI), have emerged as interactive language models that can interpret and answer scientific questions via verbal human input.[13,14] Herein, we demonstrate how the general-purpose language model ChatGPT can be utilized to generate initial-guess reaction array designs for specific substrate pairs. The output can be directly translated into input files for the HTE management software phactor.[15] We showcase several case studies of using ChatGPT to aid in designing reaction arrays for phactor, specifically for transformations that are most commonly used in pharmaceutical chemistry.[16] With phactor, we execute the arrays designed by ChatGPT experimentally leading to viable first-pass reaction conditions from simple prompts that are easy to devise by non-expert users.

■ EXPERIMENTAL SECTION

To test the effectiveness of reaction arrays designed by ChatGPT, a workflow to automatically generate reagent proposals and execute reaction arrays for popular reactions was developed. A typical workflow can be summarized in three steps (Figure 1):

- First, we should have ChatGPT generate reaction array designs for specific substrates based on simple text prompts.
- Then, we should translate the output from ChatGPT into an input file for the HTE management software phactor. An interfacing script written in python is provided online.
- Finally, we should use phactor to create stock solutions and distribute the chemicals into the reaction array, manually or robotically, and then analyze its results.

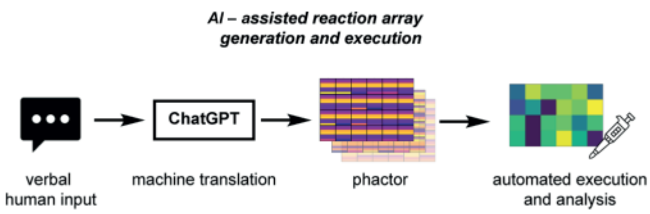

Figure 1. Overview of the ChatGPT to phactor workflow. Verbal input is given by a human to have ChatGPT generate a reaction array design for a particular coupling and substrate pair. The output can be fed directly into phactor, creating an assay recipe to be executed robotically or manually.

ChatGPT can be interrogated during the design step to elaborate on experimental details or reasonings and was asked to clarify experimental details at times. Each product was scaled up using the best conditions identified for its respective reaction array and isolated. We note that all screen designs in this work were derived from GPT-3 model responses provided through the ChatGPT web interface accessed on February 20th, 2023. We also note that the model may provide variable responses over time, which is expected based on the evolution of the large language model and its training data. Experiments

Received: June 6, 2023
Published: August 1, 2023

https://doi.org/10.1021/acs.oprd.3c00186
Org. Process Res. Dev. 2023, 27, 1510–1516

testing variability in model output can be found in the Supporting Information.

■ RESULTS

ChatGPT is a newly released general-purpose AI language model developed by OpenAI.[17] It serves as a conversational model where the user can ask a series of questions and receive text answers based on the context of the conversation. While not directly a model for chemistry, ChatGPT has been trained on a large corpus of scientific literature. As such, in its own words, ChatGPT has "knowledge of basic chemistry concepts, such as the periodic table, chemical reactions, acids and bases, and thermodynamics. It can also provide information on more advanced topics, such as organic chemistry, biochemistry, and physical chemistry". We demonstrate how ChatGPT can be asked to generate reaction arrays of viable reagents and catalysts for common reaction classes for specific substrates. For each class of reactions (amide, Suzuki, and Buchwald–Hartwig couplings), we ask ChatGPT to develop an experimental design for various pairs of substrates. While we hypothesize that the model is exercising little, if any, physical and chemical intuition in its designs, its ability to select popular reagents and catalysts associated with reaction-type keywords leads to viable and interesting proposals for array recipes. Critically, the merger with phactor exploits the strength of ChatGPT to propose several plausible answers and then sample them systematically using HTE as opposed to relying on a single "correct" answer. Ultimately, this merger of ChatGPT and phactor led to successful reaction conditions in every case interrogated.

Amide Coupling. In the first conversation, we requested the generation of a reaction array to optimize an amide coupling between 2-methylbenzoic acid (1) and *p*-toluidine (2) to form amide 3 (Figure 2).

1 2 3

Figure 2. Amide coupling between 1 and 2 ChatGPT was asked to optimize. The reagent classes are specified in the prompt, but the specific species are generated by ChatGPT.

DOIs have been omitted from the shown response as the DOI and citation references did not match the article titles. Furthermore, many of the references are, to the best of our knowledge, not real. This is because ChatGPT is a language model rather than a knowledge model and has been reported to hallucinate citations. Despite this, the array design is reasonable. Further examples of "references" for each subsequent array design are included in the Supporting Information. We emphasize that it is not advised to ask the model for archival data such as DOIs or direct citations at this time. We also note that the model struggled to accurately recreate the SMILES of compounds when asked (see Supporting Information), but structures were derived from the IUPAC name output, which were always reasonable.

The experimental execution of this reaction array precisely followed the suggestions from ChatGPT with resultant reaction metadata such as concentrations, volumes, and well locations designed by phactor (Figure 3). Two-thirds of the

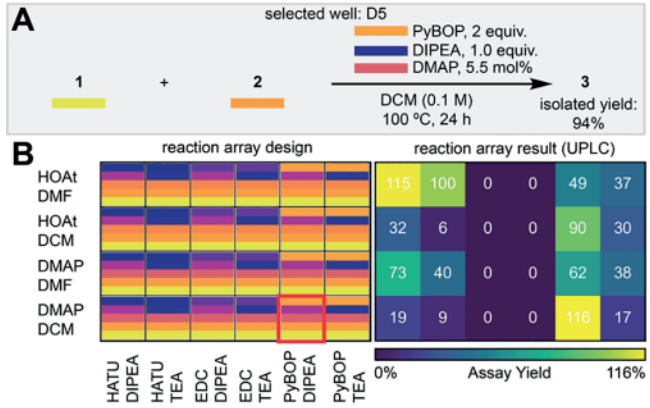

Figure 3. Executed reaction array and UPLC assay results of the screen designed by ChatGPT to perform the amide coupling between 1 and 2. The top-performing reaction condition was repeated on a 0.2 mmol scale to afford 3 in a 94% isolated yield. (A) Best-performing reaction selected from the reaction array and isolated yield of product. Color bars adjacent to reagents correspond to compound color mapping generated on phactor for reaction array visualization. (B) Reaction array design and results as displayed on phactor. Assay yields were calculated through the calibration curve equation derived from the isolated product over internal standard integrations.

array produced hits with a moderate assay yield, while reactions using 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC) as a coupling agent failed entirely. Well D5, the best hit using benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) and *N*,*N*-diisopropylethylamine (DIPEA) with (4-dimethylaminopyridine) DMAP in dichloromethane (DCM), was scaled up and resulted in a 94% isolated yield of product 3.

Amide Coupling on Complex Molecule Sitagliptin. Next, we explored how the conversation can be continued with a more complex substrate for the amide coupling. In the same dialog, we asked ChatGPT to refer to the previous question but to instead optimize the amide coupling between sitagliptin (4) and carboxylic acid 2 to form amide 5 (Figure 4).

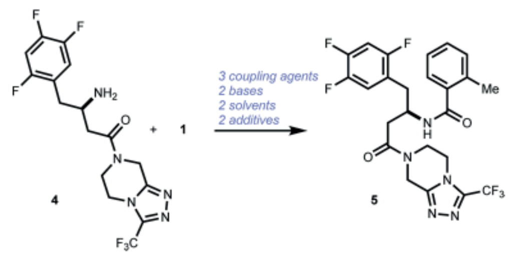

Figure 4. Amide coupling between 4 and 1 ChatGPT was asked to optimize. The resultant design is different from that with the original simpler substrates.

Again, the reaction array generated by ChatGPT was executed using the recipe designed by phactor (Figure 5). Seemingly, the results of this array performed better than the original amide coupling, with only 3 failed hits, but with lower overall yields. The best hit was well A6, which produced 5 in a 62% yield when using (7-azabenzotriazol-1-yloxy)-tripyrrolidinophosphonium hexafluorophosphate (PyAOP), triazabicyclodecene (TBD), and 1-hydroxy-7-azabenzotriazole (HOAt) in dimethylformamide (DMF). When scaled up, the reaction yielded 62% isolated product.

https://doi.org/10.1021/acs.oprd.3c00186
Org. Process Res. Dev. 2023, 27, 1510–1516

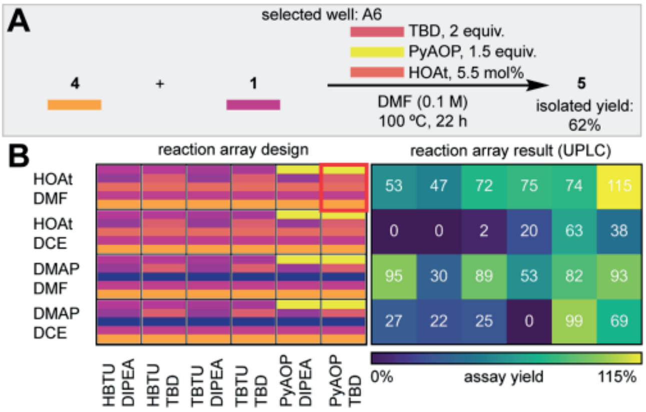

Figure 5. Results of the amide coupling to produce **5** designed by ChatGPT and phactor when executed experimentally. The top-performing reaction condition was repeated on a 0.4 mmol scale to afford **5** in a 62% isolated yield. (A) Best-performing reaction selected from the reaction array and isolated yield of product. Color bars adjacent to reagents correspond to compound color mapping generated on phactor for reaction array visualization. (B) Reaction array design and results as displayed on phactor. Assay yields were calculated through the calibration curve equation derived from the isolated product over internal standard integrations.

Suzuki Coupling. We then attempted to optimize a Suzuki coupling reaction. A new chat was initiated with ChatGPT to propose new inputs for phactor.

Here, we asked for an optimized reaction array to form biaryl **8** from the Suzuki coupling between boronate **6** and chloride **7** (Figure 6). A typical ligand versus catalyst screen is

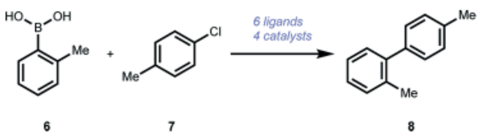

Figure 6. Suzuki coupling between **6** and **7** ChatGPT was asked to optimize.

generated, as requested. The model performed well at pulling ligands and catalysts from the literature corpus, as well as generating additional parameters needed for the assay. These probabilistic proposals for ligands and reagents are the perfect input for phactor, which then enables systematic testing of all combinations.

We note that the model does not discuss the addition of a base, so we directly ask in the same conversation:

The Suzuki coupling between **6** and **7** proved to be more challenging for ChatGPT, with slightly over half the plate failing to produce a hit (Figure 7). Pd(dppf)Cl_2 performed consistently well and DavePhos produced a hit whenever used. The unusual combination of Pd(dppf)Cl_2 and BrettPhos, which has not been previously reported, to our knowledge, generated the best hit with a 25% product/internal standard integration ratio. When scaled up, 66% of the product was isolated. Interestingly, $Pd_2(dba)_3$ has been previously reported to yield over 95% with these substrates with when run with a bisimidazolyl ligand[18] or the WK-Phos ligand.[19] These reports were seemingly missed by the AI.

We noted that ChatGPT did not mention the addition of water as an additive in the Suzuki coupling reaction conditions, which is typically required to activate boronic acids.[20] When

Prompt: "Generate a reaction array with 3 coupling agents, 2 bases, 2 solvents and 2 additives to optimize an amide coupling between 2-methylbenzoic acid and p-toluidine. Add each reagent into a single table, where each row has a column for the reagent class and a column for the reactant name. As well, include specific recommendations and values for other reaction parameters for this screen. Respond succinctly. Finally, cite some literature references from which this screen was designed with DOIs."

ChatGPT Response:

Reagent Class	Reactant Name
Coupling Agent	HATU
Coupling Agent	EDC
Coupling Agent	PyBOP
Base	DIPEA
Base	TEA
Solvent	DMF
Solvent	DCM
Additive	HOAt
Additive	DMAP

For reaction parameters, some specific recommendations for this amide coupling reaction could include:

- Temperature: 0-25°C (refrigeration)
- Time: 2-24 hours
- Stoichiometry: 1.2-2.0 equivalents of coupling agent, 1.0-1.5 equivalents of base, and 1.0-1.2 equivalents of amine relative to carboxylic acid
- Concentration: 0.1-0.5 M
- Inert atmosphere: nitrogen or argon

References:

- Li, H.; Banerjee, S.; Liao, L.; Pei, X.; Chen, X.; Miao, Q.; Huang, W. An efficient one-pot approach for the synthesis of unsymmetrical ureas via domino reaction of isocyanides with amines using EDC as the coupling agent.
- Kamber, D. N.; Jeon, J.; Lee, S. H.; Kim, S. K. Efficient N,N'-dimethylformamide catalyzed synthesis of amides from carboxylic acids using HATU as a coupling reagent.
- Hwang, S. H.; Park, S. H.; Yang, J. E.; Yoon, Y. J.; Kim, D. S.; Lee, D. W. High yield synthesis of peptide-PEGylated sunitinib via EDC-mediated conjugation.

Prompt: "Generate a similar reaction assay that instead optimizes the coupling of Sitagliptin with 2-methylbenzoic acid. Respond in the same format."

ChatGPT Response:

Reagent Class	Reactant Name
Coupling Agent	HBTU
Coupling Agent	TBTU
Coupling Agent	PyAOP
Base	DIEA
Base	TBD
Solvent	DMF
Solvent	DCE
Additive	DMAP
Additive	HOAt

For other reaction parameters, the following values are recommended:

- Temperature: 0-25°C (refrigeration)
- Time: 2-24 hours
- Stoichiometry: 1.2-2.0 equivalents of coupling agent, 1.0-1.5 equivalents of base, and 1.0-1.2 equivalents of amine relative to carboxylic acid
- Concentration: 0.1-0.5 M
- Inert atmosphere: nitrogen or argon

https://doi.org/10.1021/acs.oprd.3c00186
Org. Process Res. Dev. 2023, 27, 1510−1516

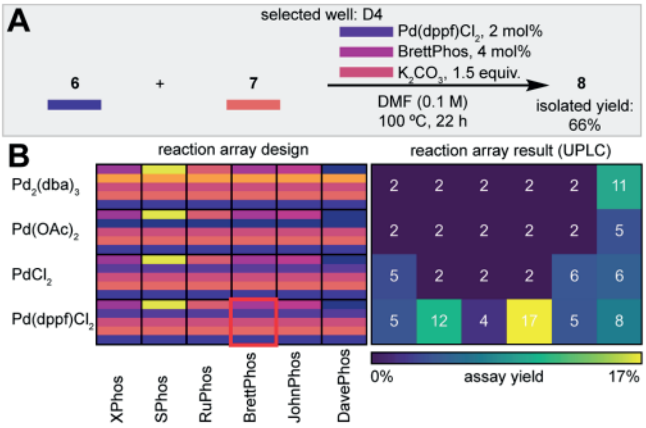

Figure 7. Results of the Suzuki coupling reaction array designed by ChatGPT and executed with phactor. The top-performing reaction condition was repeated on a 0.2 mmol scale to afford 8 in a 66% isolated yield. (A) Best-performing reaction selected from the reaction array and isolated yield of product. Color bars adjacent to reagents correspond to compound color mapping generated on phactor for reaction array visualization. (B) Reaction array design and results as displayed on phactor. Assay yields were calculated through the calibration curve equation derived from the isolated product over internal standard integrations.

asked, the model responds that water may or may not be beneficial to the yield of the reaction (see Supporting Information). We repeated the design above with base added as an aqueous solution to compare the results, finding that the yields are significantly improved across the plate (Figure 8).

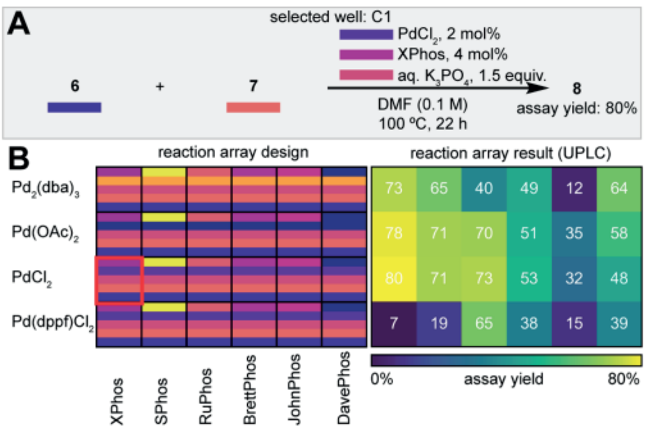

Figure 8. Results of the repeated Suzuki coupling reaction array designed by ChatGPT when 20% water is added. (A) Best-performing reaction selected from the reaction array and assay yield of the product. Color bars adjacent to reagents correspond to compound color mapping generated on phactor for reaction array visualization. (B) Reaction array design and results as displayed on phactor. Assay yields were calculated through the calibration curve equation derived from the isolated product over internal standard integrations.

Buchwald–Hartwig Coupling. Finally, we explored the ability for the model to optimize a Buchwald–Hartwig coupling. With this transformation, the interplay of ligands and base effects is often more complex.

Again, a new conversation was started. In this final example, we asked the model to generate a two factorial screen to optimize the C–N coupling between aryl chloride 9 and aniline 2 to form compound 10 (Figure 9).

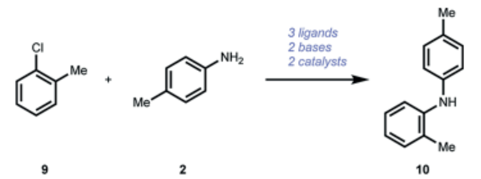

Figure 9. Buchwald–Hartwig coupling ChatGPT is asked to optimize.

Prompt: "Generate a 6 ligand by 4 catalyst reaction array to optimize a Suzuki coupling between o-tolylboronic acid and 1-chloro-4-methylbenzene. Separate each parameter into a single table, where each row has a column for the reagent class and a column for the reactant name. As well, include specific recommendations and values for other reaction parameters for this screen. Respond succinctly. Finally, cite some literature references from which this screen was designed with DOIs."

ChatGPT Response:

Reagent Class	Reactant Name
Ligand	XPhos
Ligand	SPhos
Ligand	RuPhos
Ligand	BrettPhos
Ligand	Buchwald ligand
Ligand	DavePhos
Catalyst	$Pd_2(dba)_3$
Catalyst	$Pd(OAc)_2$
Catalyst	$PdCl_2$
Catalyst	$Pd(dppf)Cl_2$
Substrate	o-tolylboronic acid
Substrate	1-chloro-4-methylbenzene

Conditions: The reaction should be performed under nitrogen atmosphere, and solvents such as DMF, DMSO, or toluene could be used. The reaction should be performed at 100-120°C for 12-24 hours, with a 1:1 ratio of the two substrates, and 2 mol% of Pd catalyst. The ratio of the Pd catalyst to the ligand should be 1:2.

We note the incorrect output "$P(t\text{-}Bu)_3$", which we manually corrected to the palladium complex $Pd(P(t\text{-}Bu)_3)_4$ in our experimental execution. We asked additional questions to clarify the experimental details.

Prompt: "What base should I use?"

The model responds: A base is typically not necessary for Suzuki coupling reactions since the reaction proceeds via a transmetalation step rather than a deprotonation step. However, if a base is needed, a common choice is potassium carbonate (K_2CO_3) or cesium carbonate (Cs_2CO_3) at a loading of 1.2-2 equiv.

We note that the quench of the reaction should be performed after cooling for safety. A follow-up question was asked to interrogate if ChatGPT would identify the single best answer from the array performed with ChatGPT and phactor.

This reaction array was executed experimentally with the exact specifications given by ChatGPT and exact array design given by phactor. The best hit was repeated on a 0.2 mmol scale (Figure 10). The array produced many hits with moderate yield with respect to internal standard. Furthermore, the absence of impurities and byproducts simplified purification. Well B5 was scaled up and resulted in an 84% isolated yield of product 10. The best conditions for coupling this substrate pair reported in the literature yielded 98% of 10

https://doi.org/10.1021/acs.oprd.3c00186
Org. Process Res. Dev. 2023, 27, 1510–1516

Prompt: "Generate a 3 ligand by 4 catalyst by 2 base reaction array to optimize a Buchwald–Hartwig coupling between 1-chloro-2-methylbenzene and p-toluidine. Add each reagent into a single table, where each row has a column for the reagent class and a column for the reactant name. As well, include specific recommendations and values for other reaction parameters for this screen. Respond succinctly. Finally, cite some literature references from which this screen was designed with DOIs."

ChatGPT Response:

Reagent Class	Reactant Name
Ligand	XPhos
Ligand	SPhos
Ligand	RuPhos
Catalyst	$Pd_2(dba)_3$
Catalyst	$Pd(dppf)Cl_2$
Catalyst	$P(t\text{-}Bu)_3$
Catalyst	$Pd(OAc)_2$
Base	K_3PO_4
Base	NaO^tBu

Other recommended reaction parameters:

- Solvent: DMF
- Temperature: 100°C
- Reaction time: 24 hours
- Stoichiometry: 1:1 ratio of 1-chloro-2-methylbenzene and p-toluidine
- Concentration: 0.1 M

Prompt: What catalyst, ligand, and base concentrations should I use for this reaction?

ChatGPT Response: The specific concentrations of catalyst, ligand, and base to use for the reaction depend on various factors such as the desired reaction rate, reaction yield, and the specific reactants being used. However, as a general guideline, a common catalyst concentration range for Buchwald–Hartwig couplings is 0.5-5 mol%. Ligand concentrations can range from 1-10 mol%. Base concentrations can range from 0.5-2 equiv. It is important to note that these concentrations may need to be optimized for specific substrates and ....

Prompt: How about the order of addition for the reagents?

ChatGPT Response: The order of addition for the reagents can also be important for optimizing the reaction conditions. In a typical Buchwald–Hartwig coupling, the base is usually added first to the reaction mixture, followed by the catalyst and the ligand. The organic halide and the amine are then added to initiate the reaction. However, depending on the specific reactants and conditions, the order of addition may need to be adjusted to achieve the desired reaction outcome. Therefore, it is always a good idea to consult the literature or perform some preliminary experiments to determine the optimal order of addition for the specific reaction being

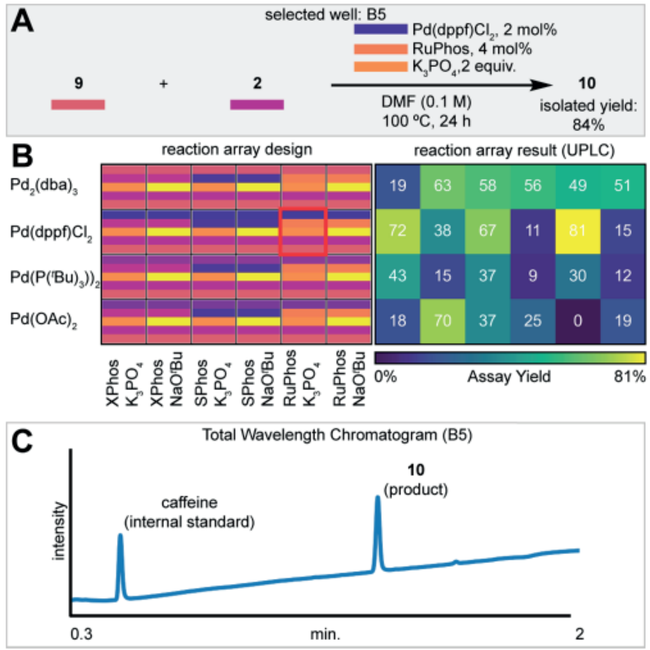

Figure 10. (A) 0.2 mmol scale-up result of the best hit from the reaction array proposed by ChatGPT-phactor. The combination of the palladium catalyst complex $Pd(dppf)Cl_2$ and RuPhos gave the best result. (B) Reaction array design and results. (C) Crude UPLC trace for selected well B5.

when using an N-heterocyclic carbene–palladium(II) chloride complex and potassium *tert*-butoxide as reaction conditions.[21]

The model's output seemed to select catalysts and ligands that have performed well in the literature but may be less aware of how they are used together. Moreover, phactor necessarily combines the answers from ChatGPT, which were not likely given with an understanding of combinatorial enumeration. As such, the resultant array includes unusual combinations of palladium sources and ligands, such as $PdCl_2(dppf)$ and RuPhos wherein two phosphine ligands are present. Furthermore, DMF and DMSO are less typical than ethereal solvents for this reaction. Despite this, the assay still produced positive results.

The combination of $Pd(dppf)Cl_2$ and RuPhos is a surprising result since there is conceivably an equilibrating mixture of $Pd(dppf)L_n$ and $Pd(RuPhos)L_n$ complexes. We suspect the proposal to mix these compounds arises as an artifact of making a combinatorial array (phactor) out of the popular Pd complex and ligand choices from the literature (ChatGPT). Nonetheless, the observation that this "cocktail" of ligands was the most productive result could be supported by related reports of "cocktail" catalysis in the Buchwald–Hartwig coupling.[22−25]

■ DISCUSSION

With several reaction arrays executed and an understanding of effective prompt engineering ChatGPT needs to generate effective arrays, we sought to integrate ChatGPT directly into the interface of phactor. The design goal was to allow chemists to generate relevant reagents for specific substrate pairs and a given transformation, even if they were non-experts. The GPT-3.5 API provided by OpenAI enabled integration on the "chemicals" tab of phactor (Figure 11). Here, the user selects substrates already added to the design stage in phactor and describes the desired reaction. Given the user inputs, the prompt template is automatically filled with ChatGPT's proposals. The user retains the ability to modify the prompt as desired to add or lessen constraints. Once the prompt is submitted, AI responds with a given number of reagents specified by the user. These compounds can be instantly added to phactor with a molecular weight and SMILES by clicking the "add" button, as long as they appear in PubChem.[26]

https://doi.org/10.1021/acs.oprd.3c00186
Org. Process Res. Dev. 2023, 27, 1510−1516

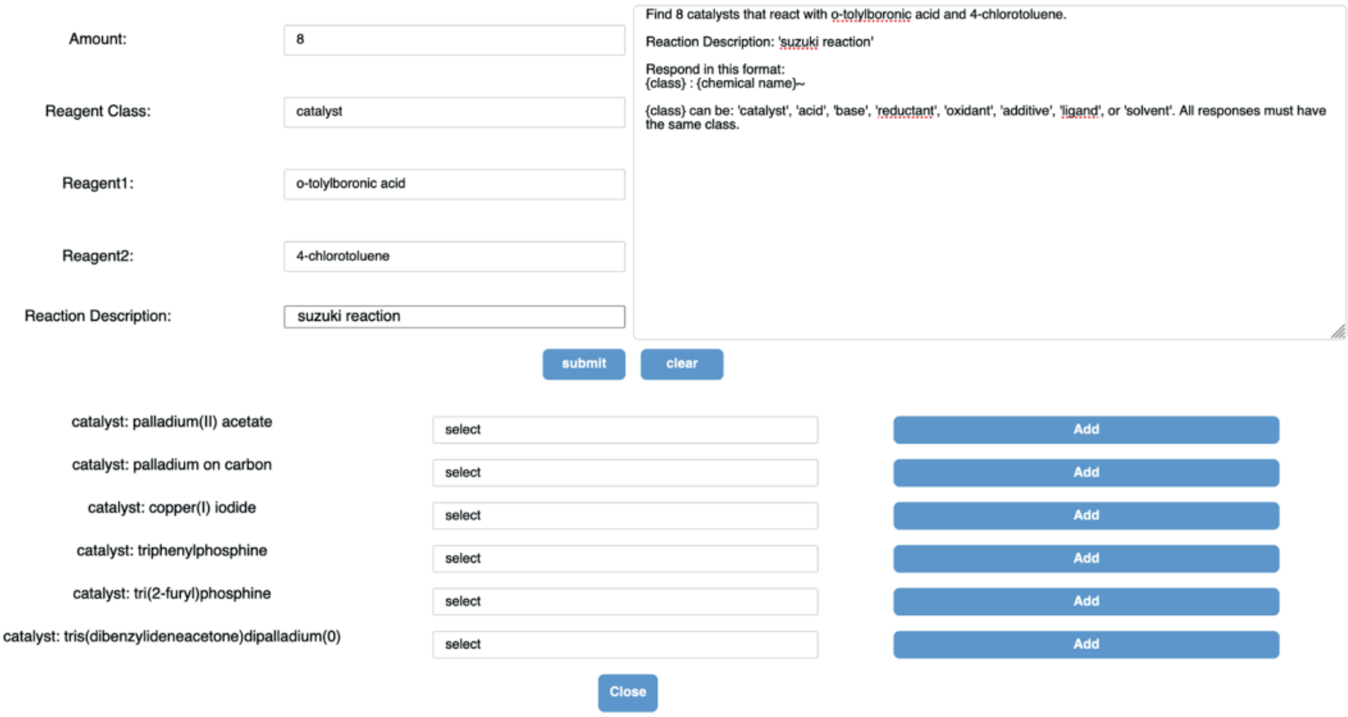

Figure 11. Integrated GPT interface in phactor. Based on the substrates selected by the chemist and a description of the desired transformation, chemicals suggested by AI can be automatically added into the reaction array design by phactor.

CONCLUSIONS

The software ChatGPT was utilized to generate reaction array designs with the HTE design software phactor for several popular reactions. The desired product was observed in every instance tried on the first attempt, and isolated yields ranged from 62 to 94%. This initial study showcases the impact that large language model predictions can have on chemical synthesis when coupled to HTE.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.oprd.3c00186.

Additional model prompts and responses, experimental details, NMR spectra, and calibration curves (PDF)

AUTHOR INFORMATION

Corresponding Author

Tim Cernak − *Department of Medicinal Chemistry and Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109, United States*; Email: tcernak@umich.edu

Authors

Babak Mahjour − *Department of Medicinal Chemistry, University of Michigan, Ann Arbor, Michigan 48109, United States*; orcid.org/0000-0002-8225-6514

Jillian Hoffstadt − *Department of Medicinal Chemistry, University of Michigan, Ann Arbor, Michigan 48109, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.oprd.3c00186

Author Contributions

B.M. engineered ChatGPT prompts, developed the phactor–GPT interface, and performed scale-up chemistries and created calibration curves. J.H. executed reaction arrays. T.C. supervised the work. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding

We gratefully acknowledge funding from the National Science Foundation (CHE-2236215) (T.C.), the Schmidt Futures Foundation, and an American Chemical Society MEDI Predoctoral Fellowship (B.M.).

Notes

The authors declare no competing financial interest.

ABBREVIATIONS

GPT, generative pre-training transformer; HATU, jexafluorophosphate azabenzotriazole tetramethyl uronium; EDC, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide; PyBOP, benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate; DIPEA, *N*,*N*-diisopropylethylamine; TEA, triethylamine; DMF, dimethylformamide; DCM, dichloromethane; HOAt, 1-hydroxy-7-azabenzotriazole; DMAP, 4-dimethylaminopyiridine; HBTU, hexafluorophosphate benzotriazole tetramethyl uronium; TBTU, 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium tetrafluoroborate; PyAOP, (7-azabenzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate; DIEA, *N*,*N*-diisopropylethylamine; TBD, triazabicyclodecene; DCE, 1,2-dichloroethane; $Pd_2(dba)_3$, tris(dibenzylideneacetone)dipalladium(0); $Pd(OAc)_2$, palladium(II) acetate; $Pd(dppf)Cl_2$, [1,1′-bis(diphenylphosphino)ferrocene]dichloropalladium(II)

https://doi.org/10.1021/acs.oprd.3c00186
Org. Process Res. Dev. 2023, 27, 1510−1516

■ REFERENCES

(1) Mahjour, B.; Shen, Y.; Cernak, T. Ultrahigh-throughput experimentation for information-rich chemical synthesis. *Acc. Chem. Res.* 2021, *54*, 2337−2346.

(2) 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.

(3) 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.

(4) 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*, No. eaar6236.

(5) Buitrago Santanilla, A.; Regalado, E. L.; Pereira, T.; Shevlin, M.; Bateman, K.; Campeau, L.-C.; Schneeweis, J.; Berritt, S.; Shi, Z.-C.; Nantermet, P.; Liu, Y.; Helmy, R.; Welch, C. J.; Vachal, P.; Davies, I. W.; Cernak, T.; Dreher, S. D. Nanomole-scale high-throughput chemistry for the synthesis of complex molecules. *Science* 2015, *347*, 49.

(6) Shen, Y.; Borowski, J. E.; Hardy, M. A.; Sarpong, R.; Doyle, A. G.; Cernak, T. Automation and computer-assisted planning for chemical synthesis. *Nat. Rev. Methods Primers* 2021, *1*, 23.

(7) Torres, J. A. G.; Lau, S. H.; 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.

(8) Shields, B. J.; Stevens, J.; Li, J.; Parasram, M.; Damani, F.; Alvarado, J. I. M.; Janey, J. M.; Adams, R. P.; Doyle, A. G. Bayesian reaction optimization as a tool for chemical synthesis. *Nature* 2021, *590*, 89−96.

(9) Williams, W. L.; Zeng, L.; Gensch, T.; Sigman, M. S.; Doyle, A. G.; Anslyn, E. V. The evolution of data-driven modeling in organic chemistry. *ACS Cent. Sci.* 2021, *7*, 1622−1637.

(10) Gao, W.; Raghavan, P.; Coley, C. W. Autonomous platforms for data-driven organic synthesis. *Nat. Commun.* 2022, *13*, 1075.

(11) Tu, Z.; Stuyver, T.; Coley, C. W. Predictive chemistry: Machine learning for reaction deployment, reaction development, and reaction discovery. *Chem. Sci.* 2023, *14*, 226−244.

(12) Shim, E.; Kammeraad, J. A.; Xu, Z.; Tewari, A.; Cernak, T.; Zimmerman, P. M. Predicting reaction conditions from limited data through active transfer learning. *Chem. Sci.* 2022, *13*, 6655−6668.

(13) Luo, R.; Sun, L.; Xia, Y.; Qin, T.; Zhang, S.; Poon, H.; Liu, T.-Y. BioGPT: generative pre-trained transformer for biomedical text generation and mining. *Briefings Bioinformat.* 2022, *23*, No. bbac409.

(14) Ferruz, N.; Schmidt, S.; Höcker, B. ProtGPT2 is a deep unsupervised language model for protein design. *Nat. Commun.* 2022, *13*, 4348.

(15) Mahjour, B.; Zhang, R.; Shen, Y.; McGrath, A.; Zhao, R.; Mohamed, O. G.; Lin, Y.; Zhang, Z.; Douthwaite, J. L.; Tripathi, A.; Cernak, T. Rapid Planning and Analysis of High-Throughput Experiment Arrays for Reaction Discovery. *Nat. Commun.* 2023, *14*, 3924.

(16) Bostrom, J.; Brown, D. G.; Young, R. J.; Keseru, G. M. Expanding the medicinal chemistry synthetic toolbox. *Nat. Rev. Drug Discovery* 2018, *17*, 709−727.

(17) OpenAI. https://openai.com/ (accessed February 2023).

(18) Zhang, C.; Trudell, M. L. Palladium-bisimidazol-2-ylidene complexes as catalysts for general and efficient Suzuki cross-coupling reactions of aryl chlorides with arylboronic acids. *Tetrahedron Lett.* 2000, *41*, 595−598.

(19) Choy, P. Y.; Yuen, O. Y.; Leung, M. P.; Chow, W. K.; Kwong, F. Y. A Highly Efficient Monophosphine Ligand for Parts per Million Levels Pd-Catalyzed Suzuki−Miyaura Coupling of (Hetero) Aryl Chlorides. *Eur. J. Org. Chem.* 2020, *2020*, 2846−2853.

(20) Thomas, A. A.; Denmark, S. E. Pre-transmetalation intermediates in the Suzuki-Miyaura reaction revealed: The missing link. *Science* 2016, *352*, 329−332.

(21) Zhao, X.-Y.; Zhou, Q.; Lu, J.-M. Synthesis and characterization of N-heterocyclic carbene-palladium (ii) chlorides-1-methylindazole and-1-methylpyrazole complexes and their catalytic activity toward C−N coupling of aryl chlorides. *RSC Adv.* 2016, *6*, 24484−24490.

(22) King, A. K.; Brar, A.; Findlater, M. A tertiary phosphine oxide ligand-based recyclable system for the Suzuki−Miyaura and Negishi reactions: evidence for pseudo-homogeneous catalysis. *Catal. Sci. Technol.* 2023, *13*, 301−304.

(23) Prima, D. O.; Madiyeva, M.; Burykina, J. V.; Minyaev, M. E.; Boiko, D. A.; Ananikov, V. P. Evidence for "cocktail"-type catalysis in Buchwald−Hartwig reaction. A mechanistic study. *Catal. Sci. Technol.* 2021, *11*, 7171−7188.

(24) Ondar, E. E.; Burykina, J. V.; Ananikov, V. P. Evidence for the "cocktail" nature of platinum-catalyzed alkyne and alkene hydrosilylation reactions. *Catal. Sci. Technol.* 2022, *12*, 1173−1186.

(25) Fors, B. P.; Buchwald, S. L. A multiligand based Pd catalyst for C− N cross-coupling reactions. *J. Am. Chem. Soc.* 2010, *132*, 15914−15917.

(26) NIH. https://pubchem.ncbi.nlm.nih.gov/ (accessed February 2023).

https://doi.org/10.1021/acs.oprd.3c00186
Org. Process Res. Dev. 2023, 27, 1510−1516