



# A Perspective on Innovating the Chemistry Lab Bench

Alexander G. Godfrey\*, Samuel G. Michael, Gurusingham Sitta Sittampalam and Gergely Zahoránszky-Köhalmi

National Center for Advancing Translational Sciences, National Institutes of Health, Bethesda, MD, United States

Innovating on the design and function of the chemical bench remains a quintessential challenge of the ages. It requires a deep understanding of the important role chemistry plays in scientific discovery as well a first principles approach to addressing the gaps in how work gets done at the bench. This perspective examines how one might explore designing and creating a sustainable new standard for advancing automated chemistry bench itself. We propose how this might be done by leveraging recent advances in laboratory automation whereby integrating the latest synthetic, analytical and information technologies, and AI/ML algorithms within a standardized framework, maximizes the value of the data generated and the broader utility of such systems. Although the context of this perspective focuses on the design of advancing molecule of potential therapeutic value, it would not be a stretch to contemplate how such systems could be applied to other applied disciplines like advanced materials, foodstuffs, or agricultural product development.

## OPEN ACCESS

### Edited by:

Beat Buesser,  
IBM Research, Ireland

### Reviewed by:

David Paul Fries,  
Florida Institute for Human and  
Machine Cognition, United States  
Masahito Ohue,  
Tokyo Institute of Technology, Japan

### \*Correspondence:

Alexander G. Godfrey  
alex.godfrey@nih.gov

### Specialty section:

This article was submitted to  
Computational Intelligence in  
Robotics,  
a section of the journal  
Frontiers in Robotics and AI

Received: 23 May 2019

Accepted: 12 February 2020

Published: 25 February 2020

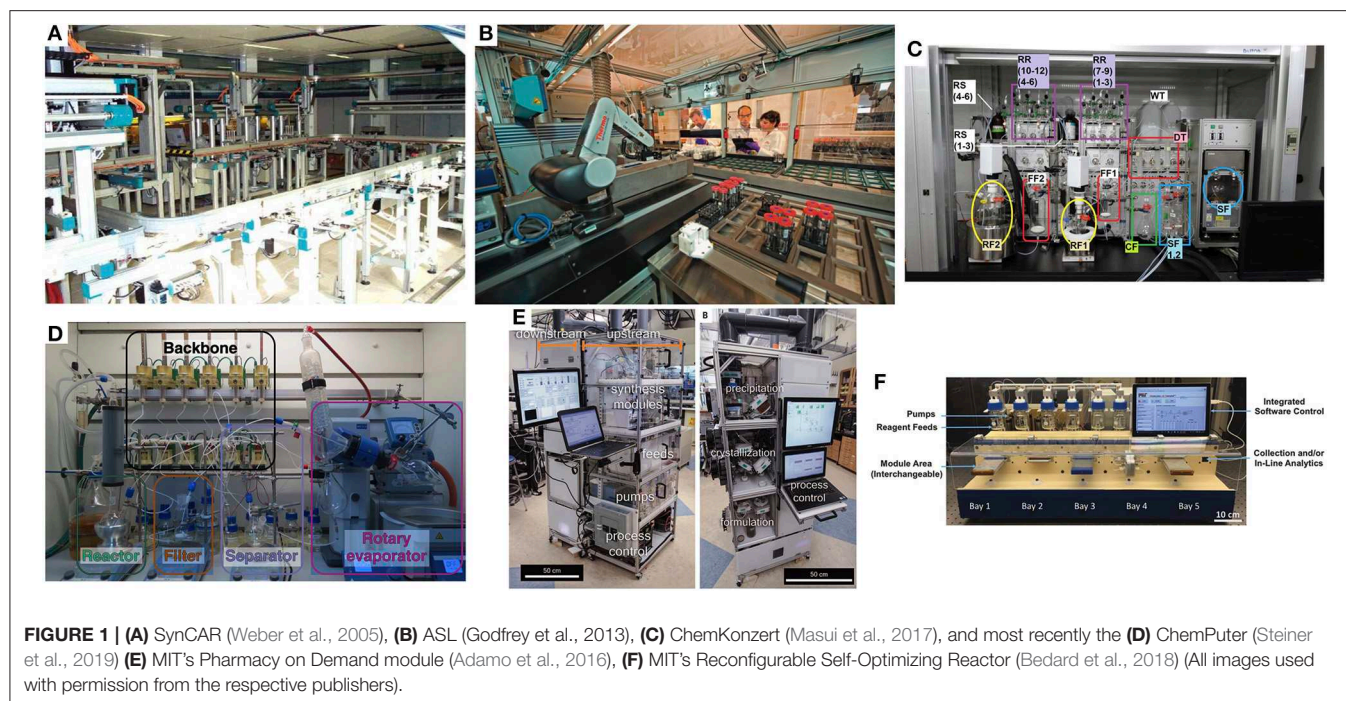
### Citation:

Godfrey AG, Michael SG,  
Sittampalam GS and  
Zahoránszky-Köhalmi G (2020) A  
Perspective on Innovating the  
Chemistry Lab Bench.  
Front. Robot. AI 7:24.  
doi: 10.3389/frobt.2020.00024

**Keywords:** automated chemical synthesis, retrosynthetic analysis, reaction planning, reagent delivery, bench chemistry

## INTRODUCTION

The central thesis posited by this perspective is the following question: What would it take to make the traditional chemistry hood no longer the central element from which chemistry discovery and critical productivity emanates? In other words, what would it take to build a platform that chemists would seek out in their day-to-day work in preference to the traditional chemistry fume hood? Notwithstanding, there are many aspects of work of the synthetic chemist that have enjoyed significant infusions of technological advancements in the form of specialized chemical synthesis platforms designed and standardized around very specific transformations [peptides (Merrifield, 1985), MIDA-boronate couplings (Li et al., 2015), oligosaccharides (Plante et al., 2001)] to more generalized platforms [Figure 1A SynCAR (Weber et al., 2005), Figure 1B ASL (Godfrey et al., 2013), Figure 1C ChemKonzert (Masui et al., 2017), and most recently the Figure 1D ChemPuter (Steiner et al., 2019)]. The ever increasing interest and value of flow chemistry technologies (Bogdan and Dombrowski, 2019) has led to highly integrated systems as well [Figure 1E MIT's Pharmacy on Demand module (Adamo et al., 2016), Figure 1F MIT's Reconfigurable Self-Optimizing Reactor (Bedard et al., 2018)]. High throughput reaction screening platforms have recently seen a resurgence of interest (Buitrago Santanilla et al., 2015; Perera et al., 2018; Isbrandt et al., 2019). But it is interesting to note how often one can find footnotes to highly automated and productive platforms being fed by substrates that have been synthesized manually. This is not directed as a critique and the incredible value provided by such platforms, but to the central question of what it would take to alleviate this manual synthesis support element still so pervasive in modern day labs.



The answer may in part come in the form of more highly integrated modular robotic platforms designed to execute both batch- and flow-mediated chemistries.

The discovery and development of chemistry methodologies remains critical to the advancement of automated synthesis. Automated solid-phase peptide synthesis, despite having initially developed well over 50 years ago (Merrifield, 1963), continues to experience advancements in its chemistry and functional versatility (Isidro-Llobet et al., 2019). Newer experimental techniques mediated through microwave irradiation (Berrino and Supuran, 2018), photo-redox catalysis (Shaw et al., 2016), and electrochemically mediated organic synthesis (Yan et al., 2018) continue to expand the synthetic chemist's toolbox expanding the options for more versatile engineered automation solutions. Continued additions to the chemistry methodology compendium will undoubtedly lead to an expansion of the scope of chemistry that will be inherently automatable and help reduce the need and the expense for more bespoke systems developed for transformations of much narrower scope. In keeping with themes like Idea2Data (Nicolaou et al., 2019) and Design-Make-Test-Analyze (Plowright et al., 2012) artificial intelligence and machine learning or specially tailored algorithms will play ever increasing roles toward more autonomous, routine transformations that can be directly applied without human intervention to more effectively advance knowledge in the life sciences.

## AUTOMATING THE BENCH

One of the biggest challenges an automated synthesis lab is likely to face in a research setting is executing syntheses involving a novel pairing of reactants, reagents, and solvent. Any number of

factors can impede the desired forward progress of the reaction: (a) lack of solubility of any of the reactants or reagents; (b) a mis-match in catalyst selection; (c) unknown reaction kinetics or thermodynamic distribution of products. A means of mitigating some of these issues can be accomplished by running a prescreen of the reaction conditions at a reasonable scale. Results from such a reaction screening experiments to establish key reaction conditions that need to be further optimized for reproducibility, yield and scale-up requirements. Several key factors that may influence these outcomes are addressed in this perspective.

## RETROSYNTHETIC ANALYSIS

Automated synthetic labs being developed now and into the future will likely be assisted if not directly interfaced with computational tools facilitate the route selection process which will likely require the development on novel workflow development interfaces coupled to reaction template databases where automated synthesis execution routines are automatically crafted and executed. Great strides have been made in retrosynthetic analysis largely driven by nearly exponential increasing computational power and refinement of algorithms that are a result of nearly half a century of slow if not steady development (Feng et al., 2018).

Many computer aided synthetic planning programs have come and gone for reasons that are not immediately clear (Szymkuc et al., 2016). In some past anecdotal recounting about such programs the perception by some users is that retrosynthesis algorithms aren't coming up with novel synthetic routes that a skilled chemist wouldn't come up with in about the same amount of time. Compounding that perception are two likely factors: (a)

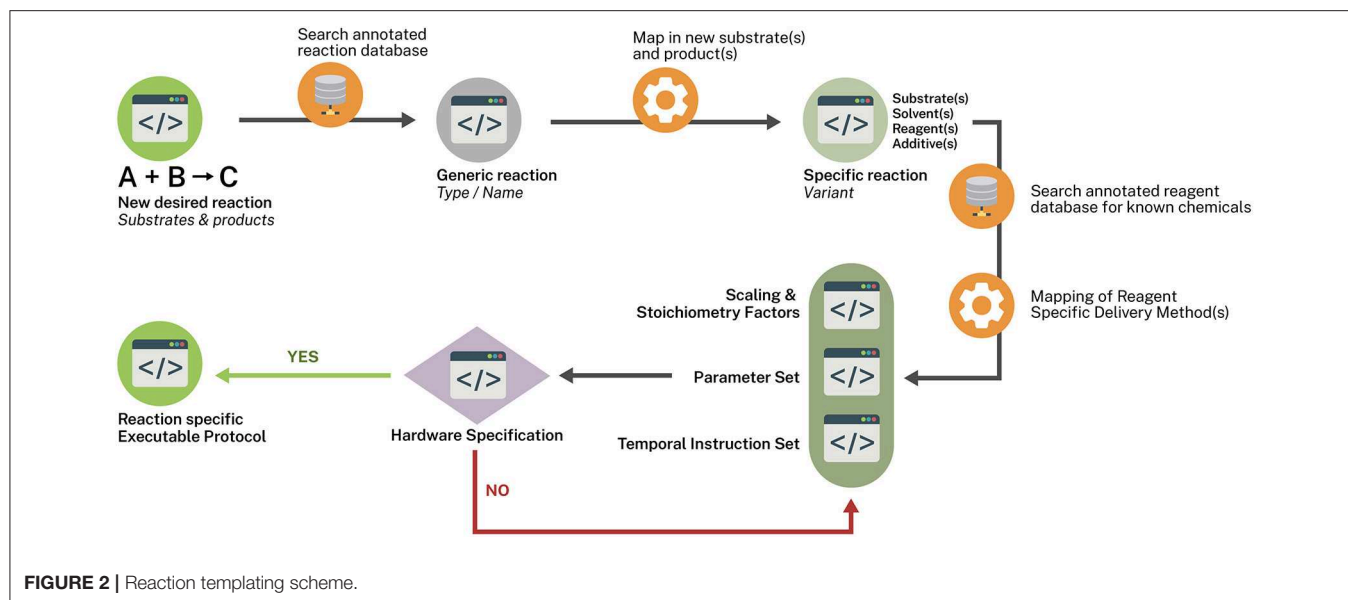


FIGURE 2 | Reaction templating scheme.

the inability of many algorithms to precisely adopt functional group protecting strategies that are applied holistically to a multi-step synthetic proposal; and (b) subjectively weighted scoring functions that rank order solutions from thousands encountered using somewhat arbitrarily ranked factors such as cost of raw materials, number of steps, synthetic complexity scores, etc. that lead to “unimaginative” results from the perspective of the end user. The latter factor has recently been addressed through refinements to the Synthia platform (Badowski et al., 2019).

More recently, a resurgent interest in automated synthesis and high throughput experimentation (Mennen et al., 2019) coupled with dramatic increases in computational performance and advances in synthetic technology provides a unique opportunity particularly when applied to molecules of lower complexity. The possibility that retrosynthetic algorithms can now readily identify disconnections considered trivial by the synthetic practitioner actually makes an interesting case for integrating retrosynthetic analysis with automated chemical synthesis (Nicolaou et al., 2016). It could be possible for such a system to computationally evaluate and triage hundreds if not thousands of possible synthetic targets, identifying those that can be assembled via simple automated transformations whilst also sorting out the availability of the raw materials involved and spawning such executions automatically on an appropriately configured platform.

An important deficiency to note in modern retrosynthetic algorithms is the lack of their ability to identify optimal reaction conditions or additives (e.g., catalyst-ligand combinations, solvents etc.) for executing the synthesis of a particular target. In the case of more complex targets, current algorithms also fail to identify specific protecting group strategies that are compatible with multistep sequences. Commercially available algorithms (Feng et al., 2018) often do a good job of identifying close or exact matches of transformation in the literature from references linked to the best fitting reaction rule. However,

such examples may not incorporate the most effective catalyst-ligand combination or solvent that could be possibly substituted consistent with the most current state-of-the-art knowledge on the transformation.

## REACTION TEMPLATING

Having identified a potential protocol for automatic execution the next hurdle is establishing a mechanism for appropriately parameterizing its execution on an automated platform. We refer to this process as Reaction Templating. A proof of concept for this approach has recently been demonstrated in the development of the ChemPuter (Li et al., 2015). The goal of such a process needs to achieve an abstraction of the general process for chemistry execution into a master schema—which may be considered a compendium of schemas of highly structured information in an extensible format (e.g., JSON, XML) that can be computationally interpreted and directly applied to machine operation and control.

A means of coupling these schemas to execution protocols may come from linking an ontologically annotated reaction database to schemas that generically represent the reaction and various key chemical participants. Such a schema (Figure 2) could then be mapped to a more specific variant of the reactant less generically represented followed by a mapping-in of novel substrates and a schema representing the parameters and scaling factors. This schema is in turn mapped to a schema of temporal instructions which in turn is mapped to hardware schema capable of executing the chemistry within the defined parameter set and instructions. A mapping exercise of the known chemicals to be involved with a reagent database annotated with reagent specific delivery descriptors (e.g., a solid-supported reagent could be delivered by a powder dispensing robot) which then map to the Temporal Instruction Set schema. Any given reaction type may in fact give rise to a variety of possible reaction

vectors such as batch or flow modes. Within these modes various temporal instruction sets may exist along with stoichiometry and hardware specification considerations. A schema such as the one proposed would allow for a variety of one to many relationships to be created without sacrificing specificity or the flexibility of introducing new methodologies as long as matching hardware specifications can be successfully mapped to the implementation being queried.

## REAGENT DELIVERY STRATEGIES

The need to achieve the reproducibility of experimental results has been currently recognized as an important challenge across a number of scientific disciplines (Baker, 2016). One of the common features sought in the development of automated synthesis equipment is addressing the reproducibility of results. As the development of conventional platforms for liquid and solid delivery continue to advance it is likely that both the scope and precision of such equipment will improve. A recent review of a variety of solid dispensing technologies surveys their strengths and weaknesses (Bahr et al., 2018).

Another means of dealing with this challenge may be to cardinalize the parameter space in such a way as to only allow delivery of pre-packaged or encapsulated units of reagent. Such pre-packaged units organized in a monetarized structure would thus allow more facile delivery of reagents and likely a much more reproducible means of precise control of variables involved in the delivery of such reagents in repeated experiments. It has recently been shown that it is possible to introduce encapsulated reagents whereby reagents can be released once delivered to the target reactor under the appropriate set of conditions that release the reagent from its encapsulation (Sather et al., 2015). This form of “synthesis discretization” may permit one to rein in the exponentially high permutations that arise from a non-bounded parameter space and reduce the cost of engineered solutions. More recently a group at ETH Zurich (Jiang et al., 2019) have shown how innovative reagent design, custom hardware engineering, and a precisely executed workflow can lead to complete and repeatable workflows that lead to cleanly isolated products. In all likelihood reagent formulation will play an ever-increasing role in the development and expansion of practical automated synthetic solutions.

## PRODUCT ISOLATION

Product isolation refers to the process involved in separating the desired synthesis product from the reaction milieu. This milieu could include a solvent that is incompatible with downstream processing and/or byproducts resulting from spent reagents or undesired side-products involving the reactants. Product isolation often possesses the greatest challenge in providing robust automated processes. As shown in apparatus developed by scientists at ETH Zurich mentioned earlier (Jiang et al., 2019), special considerations for product isolation are likely just as important as the reaction setup itself if improved automated outcomes are desired. Confounding factors include the largely unpredictable physical characteristics of the raw reaction

mixtures. Any one or combination of these characteristics may be present: (a) Solid precipitates; (b) Unexpected or unidentified byproducts; (c) changes in material viscosity (d) unconsumed reactant material requiring additional quenching.

## PURIFICATION AND CHARACTERIZATION

More often than not, automated purification processes refer to chromatographic processes, although in process development settings, economic and scaling relegate ultimate purification efforts to re-crystallization processes. However, material constraints and lack of practical engineering solutions make the consideration of re-crystallization processes prohibitively difficult to pursue in a smaller scale laboratory setting, although there are documented methods for automated removal of solids from automated workstreams (Fitzpatrick and Ley, 2018).

Ideally, reaction isolates emerging from the product isolation process will have their overall purity significantly enhanced such that additional purification is unnecessary or sufficiently pure enough to be re-introduced for multistep synthetic applications. In the event the material does not meet purity thresholds set by business or research guidelines, it would then enter the chromatographic purification process. The process usually begins with a screening of chromatographic methods to establish the best method to achieve the appropriate recovery and purity of the material. Quite often this involves exploring the redissolution of the material in select solvents or solvent mixtures to facilitate material transfers. The lack of ability to anticipate difficulties in achieving full dissolutions presents challenges that make automating the process more difficult. Quite often manual processes are implemented to remove solids either by centrifugation or filtration then further analyses are performed to determine the makeup of these isolated solids. Given the relevance of the water solubility as key property to be optimized in the context of drug development, a great deal of effort has been spent on the *in silico* prediction (Bergstrom and Larsson, 2018). However, there is a dearth in the advancement of general solubility prediction tools in solvents other than water but is not without some recent precedent (Piccione et al., 2019).

## ROLE OF AI/ML

The role of AI/ML will undoubtedly play a critical role in development of the lab bench of the future. Critical to the application of such techniques will require highly regularized or structured data. The proposed Reaction Template scheme (Figure 2) allows for a structured means of creating a network of discrete paths for executing chemical synthesis in a potentially highly reproducible manner. Pattern analysis of such networks by machine learning techniques might lead to the discovery of new, viable, immediately executable paths. In particular, if such reaction databases can be further annotated with results of reaction screening results run within the same platform, quantitative results—both positive and negative yield results—can be used to refine the algorithm.

Further, a well-annotated reaction network has the potential to transform the way we plan synthetic routes today. That is,

instead of exploring potentially untested routes to specific targets, it might be possible to identify hubs (Sudarshan et al., 2012), i.e., easily accessible intermediates or scaffolds from which the target molecules can be reached via a few number of low-risk synthetic steps. This approach can also help to reduce overall operation costs by identifying intermediates that needs to be synthesized at larger quantity in order to serve as stepping stone toward other targets. It is also possible, that new generation AI algorithms will be able to systematically reveal chemical transformations that are unknown to date (Segler and Waller, 2017). In summary, AI-based technology in chemistry not only holds the promise to facilitate the navigation of chemical transformation-space more efficiently, but it might lead to reshaping this space via the discovery of “synthetic wormholes,” i.e., novel chemistry.

## REFERENCES

- Adamo, A., Beingsner, R. L., Behnam, M., Chen, J., Jamison, T. F., Jensen, K. F., et al. (2016). On-demand continuous-flow production of pharmaceuticals in a compact, reconfigurable system. *Science* 352, 61–67. doi: 10.1126/science.aaf1337
- Badowski, T., Molga, K., and Grzybowski, B. A. (2019). Selection of cost-effective yet chemically diverse pathways from the networks of computer-generated retrosynthetic plans. *Chem. Sci.* 10, 4640–4651. doi: 10.1039/c8sc05611k
- Bahr, M. N., Damon, D. B., Yates, S. D., Chin, A. S., Christopher, J. D., Cromer, S., et al. (2018). Collaborative evaluation of commercially available automated powder dispensing platforms for high-throughput experimentation in pharmaceutical applications. *Org. Process Res. Dev.* 22, 1500–1508. doi: 10.1021/acs.oprd.8b00259
- Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature* 533, 452–454. doi: 10.1038/533452a
- Bedard, A.-C., Adamo, A., Aroh, K. C., Russell, M. G., Bedermann, A. A., Torosian, J., et al. (2018). Reconfigurable system for automated optimization of diverse chemical reactions. *Science* 361, 1220–1225. doi: 10.1126/science.aat0650
- Bergstrom, C. A. S., and Larsson, P. (2018). Computational prediction of drug solubility in water-based systems: qualitative and quantitative approaches used in the current drug discovery and development setting. *Int. J. Pharm.* 540, 185–193. doi: 10.1016/j.ijpharm.2018.01.044
- Berrino, E., and Supuran, C. T. (2018). Advances in microwave-assisted synthesis and the impact of novel drug discovery. *Expert Opin. Drug Discov.* 13, 861–873. doi: 10.1080/17460441.2018.1494721
- Bogdan, A. R., and Dombrowski, A. W. (2019). Emerging trends in flow chemistry and applications to the pharmaceutical industry. *J. Med. Chem.* 63, 6422–6488. doi: 10.1021/acs.jmedchem.8b01760
- Buitrago Santanilla, A., Regalado, E. L., Pereira, T., Shevlin, M., Bateman, K., Campeau, L.-C., et al. (2015). Nanomole-scale high-throughput chemistry for the synthesis of complex molecules. *Science* 347, 49–53. doi: 10.1126/science.1259203
- Feng, F., Lai, L., Lai, L., Pei, J., and Lai, L. (2018). Computational chemical synthesis analysis and pathway design. *Front. Chem.* 6:199. doi: 10.3389/fchem.2018.00199
- Fitzpatrick, D. E., and Ley, S. V. (2018). Engineering chemistry for the future of chemical synthesis. *Tetrahedron* 74, 3087–3100. doi: 10.1016/j.tet.2017.08.050
- Godfrey, A. G., Masquelin, T., and Hemmerle, H. (2013). A remote-controlled adaptive medchem lab: an innovative approach to enable drug discovery in the 21st Century. *Drug Discov. Today* 18, 795–802. doi: 10.1016/j.drudis.2013.03.001
- Isbrandt, E. S., Sullivan, R. J., and Newman, S. G. (2019). High throughput strategies for the discovery and optimization of catalytic reactions. *Angew. Chem. Int. Ed.* 58, 7180–7191. doi: 10.1002/anie.201812534
- Isidro-Llobet, A., Kenworthy, M. N., Mukherjee, S., Kopach, M. E., Wegner, K., Gallou, F., et al. (2019). sustainability challenges in peptide synthesis and purification: from R&D to production. *J. Org. Chem.* 84, 4615–4628. doi: 10.1021/acs.joc.8b03001
- Jiang, T., Bordi, S., McMillan, A., Chen, K.-Y., Saito, F., Nichols, P., et al. (2019). An integrated console for capsule-based, fully automated organic synthesis. *ChemRxiv* 1–11. doi: 10.26434/chemrxiv.7882799.v1
- Li, J., Ballmer, S. G., Gillis, E. P., Fujii, S., Schmidt, M. J., Palazzolo, A. M., et al. (2015). Synthesis of many different types of organic small molecules using one automated process. *Science* 347, 1221–1226. doi: 10.1126/science.aaa5414
- Masui, H., Yosugi, S., Fuse, S., and Takahashi, T. (2017). Solution-phase automated synthesis of an alpha-amino aldehyde as a versatile intermediate. *Beilstein J. Org. Chem.* 13, 106–110. doi: 10.3762/bjoc.13.13
- Mennen, S. M., Alhambra, C., Allen, C. L., Barberis, M., Berritt, S., Brandt, T. A., et al. (2019). The evolution of high-throughput experimentation in pharmaceutical development, and perspectives on the future. *Org. Process Res. Dev.* 23, 1213–1242. doi: 10.1021/acs.oprd.9b00140
- Merrifield, R. B. (1963). Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* 85, 2149–2154. doi: 10.1021/ja00897a025
- Merrifield, R. B. (1985). Solid-phase syntheses (Nobel lecture). *Angew. Chem.* 97, 801–812.
- Nicolaou, C. A., Humblet, C., Hu, H., Martin de la Nava, E.-M., Dorsey, F. C., Castle, T. M., et al. (2019). Idea2Data: towards a new paradigm for drug discovery. *ACS Med. Chem. Lett.* 10, 278–286. doi: 10.1021/acsmedchemlett.8b00488
- Nicolaou, C. A., Watson, I. A., Hu, H., and Wang, J. (2016). The proximal lilly collection: mapping, exploring and exploiting feasible chemical space. *J. Chem. Inf. Model.* 56, 1253–1266. doi: 10.1021/acs.jcim.6b00173
- Perera, D., Tucker, J. W., Brahmabhatt, S., Helal, C. J., Chong, A., Farrell, W., et al. (2018). A platform for automated nanomole-scale reaction screening and micromole-scale synthesis in flow. *Science* 359, 429–434. doi: 10.1126/science.aap9112
- Piccione, P. M., Baumeister, J., Salvesen, T., Grosjean, C., Flores, Y., Groelly, E., et al. (2019). Solvent selection methods and tool. *Org. Process Res. Dev.* 23, 998–1016. doi: 10.1021/acs.oprd.9b00065
- Plante, O. J., Palmacci, E. R., and Seeberger, P. H. (2001). Automated solid-phase synthesis of oligosaccharides. *Science* 291, 1523–1527. doi: 10.1126/science.1057324
- Plowright, A. T., Johnstone, C., Kihlberg, J., Pettersson, J., Robb, G., and Thompson, R. A. (2012). Hypothesis driven drug design: improving quality and effectiveness of the design-make-test-analyse cycle. *Drug Discov. Today* 17, 56–62. doi: 10.1016/j.drudis.2011.09.012
- Sather, A. C., Lee, H. G., Colombe, J. R., Zhang, A., and Buchwald, S. L. (2015). Dosage delivery of sensitive reagents enables glove-box-free synthesis. *Nature* 524, 208–211. doi: 10.1038/nature14654
- Segler, M. H. S., and Waller, M. P. (2017). Modelling chemical reasoning to predict and invent reactions. *Chem. Eur. J.* 23, 6118–6128. doi: 10.1002/chem.201604556

## AUTHOR CONTRIBUTIONS

AG conceived and wrote the first version of the paper. GS and SM discussed, edited the manuscript and supplied key refinements to the manuscript. GS reviewed the overall structure and content of the paper, provided critical feedback and obtained copyright permissions for all the images presented. GZ-K contributed substantially to the Role of AI/ML section.

## ACKNOWLEDGMENTS

The authors would like to gratefully acknowledge and thank Kyle Brimacombe for producing the graphical layout on Reaction Templating (Figure 2).

- Shaw, M. H., Twilton, J., and MacMillan, D. W. C. (2016). Photoredox catalysis in organic chemistry. *J. Org. Chem.* 81, 6898–6926. doi: 10.1021/acs.joc.6b01449
- Steiner, S., Wolf, J., Glatzel, S., Andreou, A., Granda, J. M., Keenan, G., et al. (2019). Organic synthesis in a modular robotic system driven by a chemical programming language. *Science* 363:eaav2211. doi: 10.1126/science.aav2211
- Sudarshan, I. S. R., Veni, M. C. E., Zweig, K. A., and Natarajan, A. (2012). Understanding human navigation using network analysis. *Top. Cogn. Sci.* 4, 121–134. doi: 10.1111/j.1756-8765.2011.01178.x
- Szymkuc, S., Gajewska, E. P., Klucznik, T., Molga, K., Dittwald, P., Startek, M., et al. (2016). Computer-assisted synthetic planning: the end of the beginning. *Angew. Chem. Int. Ed.* 55, 5904–5937. doi: 10.1002/anie.201506101
- Weber, A., von Roedern, E., and Stilz, H. U. (2005). SynCar: an approach to automated synthesis. *J. Comb. Chem.* 7, 178–184. doi: 10.1021/cc049838z
- Yan, M., Kawamata, Y., and Baran, P. S. (2018). Synthetic organic electrochemistry: calling all engineers. *Angew. Chem. Int. Ed.* 57, 4149–4155. doi: 10.1002/anie.201707584

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Godfrey, Michael, Sittampalam and Zahoránszky-Köhalmi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.