

Machine intelligence for chemical reaction space

Philippe Schwaller^{1,2}  | Alain C. Vaucher^{1,2}  | Ruben Laplaza^{2,3}  |
 Charlotte Bunne^{2,4}  | Andreas Krause^{2,4}  | Clemence Corminboeuf^{2,3}  |
 Teodoro Laino^{1,2} 

¹IBM Research Europe, Rüschlikon, Switzerland

²National Center for Competence in Research-Catalysis (NCCR-Catalysis), Zurich, Switzerland

³Institute of Chemical Sciences and Engineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

⁴Department of Computer Science, ETH Zurich, Zurich, Switzerland

Correspondence

Philippe Schwaller, IBM Research Europe, Rüschlikon 8803, Switzerland.

Email: philippe.schwaller@epfl.ch

Funding information

Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: 180544

Edited by: Raghavan Sunoj, Associate Editor

Abstract

Discovering new reactions, optimizing their performance, and extending the synthetically accessible chemical space are critical drivers for major technological advances and more sustainable processes. The current wave of machine intelligence is revolutionizing all data-rich disciplines. Machine intelligence has emerged as a potential game-changer for chemical reaction space exploration and the synthesis of novel molecules and materials. Herein, we will address the recent development of data-driven technologies for chemical reaction tasks, including forward reaction prediction, retrosynthesis, reaction optimization, catalysts design, inference of experimental procedures, and reaction classification. Accurate predictions of chemical reactivity are changing the R&D processes and, at the same time, promoting an accelerated discovery scheme both in academia and across chemical and pharmaceutical industries. This work will help to clarify the key contributions in the fields and the open challenges that remain to be addressed.

This article is categorized under:

- Data Science > Artificial Intelligence/Machine Learning
- Data Science > Computer Algorithms and Programming
- Data Science > Chemoinformatics

KEY WORDS

artificial intelligence, chemical reactions, computer-assisted synthesis planning, data-driven approaches, machine intelligence

1 | INTRODUCTION

In 1950, Alan Turing envisioned intelligent machines as digital computers that can replicate human behavior¹ to the extent of exhibiting an intelligence indistinguishable from that of a human being (imitation game or Turing test). From that moment on, machine intelligence has steadily integrated into numerous technologies such as search engines, translation, and navigation tools, assisting us in processing large amounts of data more quickly. Machine intelligence also shaped a new way to accelerate the scientific discovery, including the development of technologies to help chemists accessing and exploring the chemical reaction space.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs License](#), which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

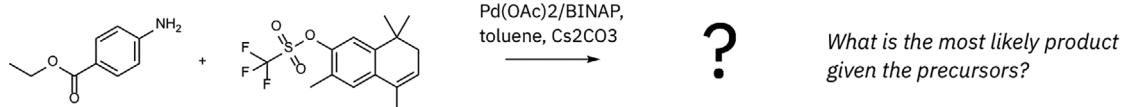
© 2022 IBM Research Europe. *WIREs Computational Molecular Science* published by Wiley Periodicals LLC.

Chemical reactions describe how reactants transform into products, and their prediction is crucial for guiding the exploration of the chemical reaction space toward more easily made molecules. Machine intelligence for chemical reaction space covers a wide range of technologies to increase the synthesis success rate. With Corey's^{2,3} pioneering work in codifying retrosynthetic rules ushering in the era of intelligent machines assisting in synthesis planning, interest in such technologies developed in tandem with the improved performance of computing resources. In the following decades, multiple computer-assisted synthesis planning projects emerged primarily focusing on the construction of knowledge bases.^{4–8} In 1970, Hendrickson⁹ developed a mathematical framework for the classification of reactions based on the carbon skeleton and its substituents. The use of these or subsequent classification scheme allowed the seamless exploration of similar chemical reactions, recommending similar transformation for chemical reaction planning. Dugundji and Ugi¹⁰ were the first to encode chemical reactions in terms of bond-electron and reaction matrices, followed by other expert systems for computer-assisted synthesis design^{11–14} inspired by the formal reaction logic. Satoh and Funatsu developed a reaction-type independent outcome prediction approach named Sophia.¹⁵ Chematica/Synthia^{16–18} later built on the knowledge base efforts of the 1970s,^{4,5,7,8} expanding the technique to include over 100 k rules extracted over a 20-year period.

The most recent progress in chemical reaction space are driven by machine learning, where interpolating functions or patterns are learned from data. None of this would have been possible without the availability of open datasets, which allowed benchmarks and validation across different approaches. Among few, Lowe¹⁹ should be acknowledged for his efforts in text-mining reaction dataset from US Patents, with updates in the recent years.²⁰ These works^{19,20} allowed the development of several deep learning approaches for computer-assisted synthesis planning.^{21–30}

The field of digital chemistry is evolving at an unprecedented speed. Here we plan to provide few insights on some of the most recent advances in the chemical reaction space. Figure 1 shows an overview of the topics discussed in this

(a) forward reaction prediction task



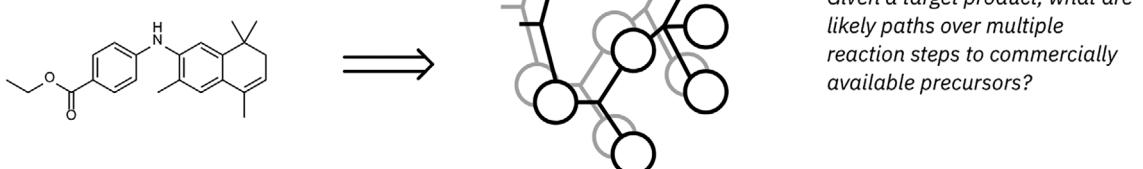
What is the most likely product given the precursors?

(b) single-step retrosynthesis prediction task



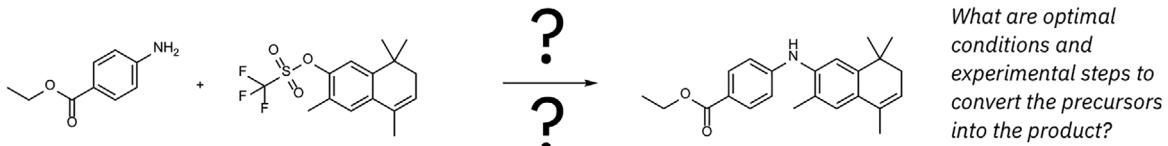
What are likely precursors or reactants combinations given a product?

(c) multi-step synthesis planning task



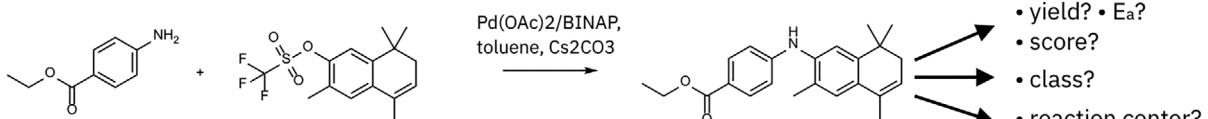
Given a target product, what are likely paths over multiple reaction steps to commercially available precursors?

(d) reaction condition and procedure predictions

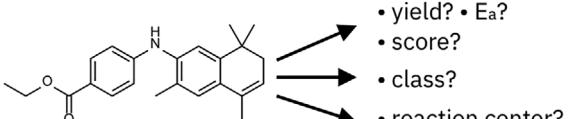


What are optimal conditions and experimental steps to convert the precursors into the product?

(e) reaction regression and performance optimization tasks



(f) reaction classification



(g) atom-mapping

- yield?
- E_a?
- score?
- class?
- reaction center?

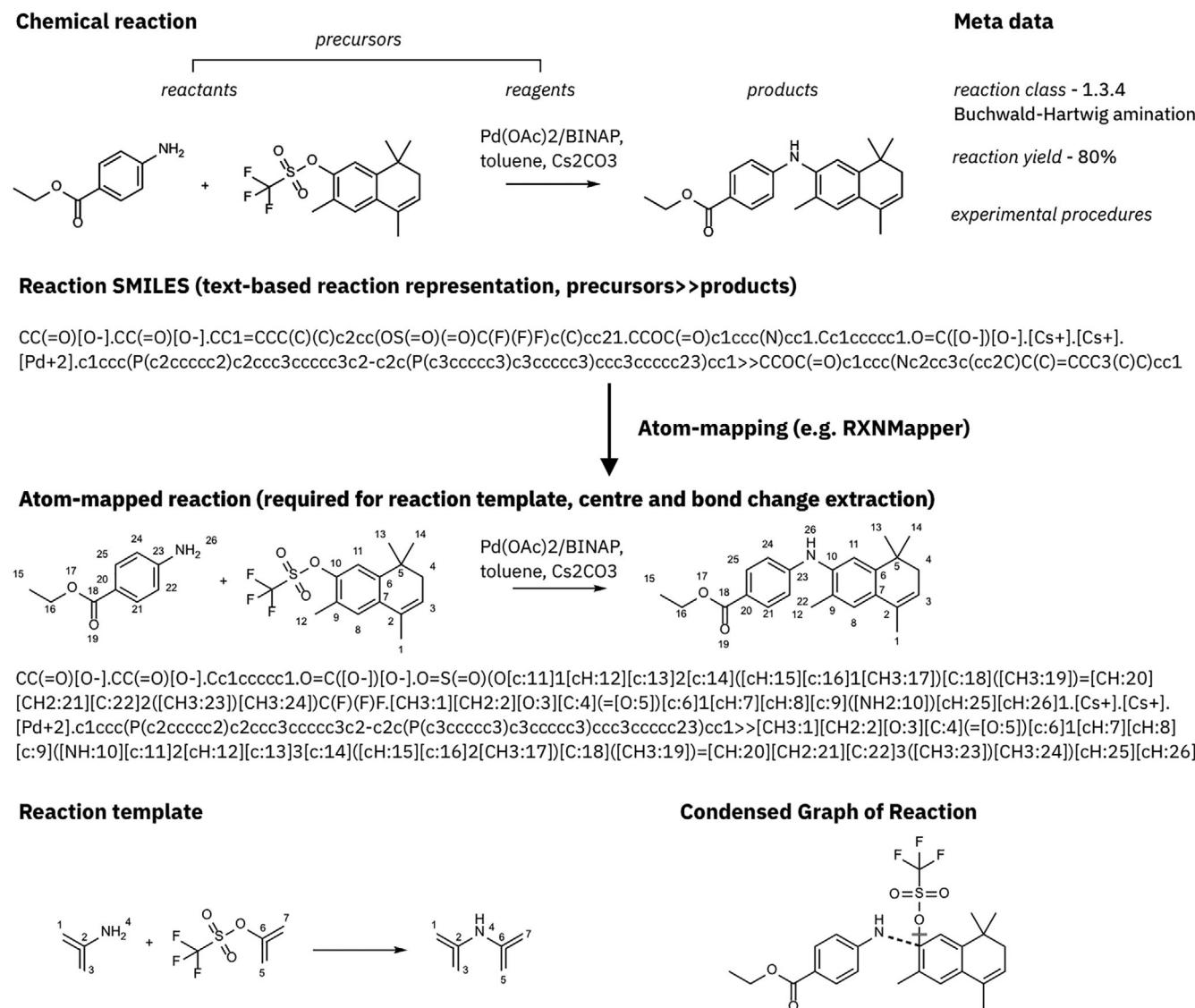
FIGURE 1 Overview of chemical reaction space tasks

manuscript. We will start with an analysis of chemical reaction datasets and their representation to finalize with the description of the diverse techniques for chemical reactions tasks, including outcome prediction, multi-step retrosynthesis, catalyst design, and reaction optimization.

In the different sections, we will touch upon several success stories where machine intelligence for chemical reaction space made an impact with experimental results, such as predicting the outcome of carbohydrate reactions,³¹ planning synthesis routes for natural products,³² optimizing catalytic reactions,³³ and the de novo design of drugs.³⁴

2 | DATA AND REPRESENTATION

Chemists use chemical equations to abstract the complex processes that transform mixtures of precursors under specific conditions into products. Following a never defined practice, reactants are placed to the left of the arrow, the reagents and conditions around the arrow, and the products to the right (see Figure 2). Because the difference between reactants and reagents can be ambiguous and multiple conventions exists how to make this distinction, we define *reactants* as the molecules that contribute atoms to the product, and *reagents* as the ones that take part in the reaction but do not contribute atoms to the reported product (e.g., catalysts and solvents). The knowledge on how atoms reconfigure in reactions, known as atom-mapping, is often used to automate this labeling. Atom-mapping also makes it possible to



locate the reaction center, determine bond changes, and extract reaction templates. Better atom-mapping schemes lead to better labeled data. After decades dominated by maximum common substructure and optimization-based atom-mapping approaches,³⁵ Schwaller et al.³⁶ have shown that neural networks capture atom-mapping patterns from unmapped reactions in unsupervised training. The resulting open-source RXNMapper is currently one of the best performing atom-mapping tools available to the chemoinformatic community.³⁷ However, in order to limit the dependence from atom-mapping performance,²⁷ it is a common practice to combine reactants and reagents as *precursors* and make no distinction between them.

There are several ways to express the molecular structures involved in a chemical reaction in computer-readable formats, such as line notations,^{38–41} MOLfiles, and SDfiles,⁴² that can be used to provide inputs for a machine learning model. Graphs are a natural way to represent most molecules, using nodes to represent atoms and edges for bonds. However, the use of low dimensional graphs (two dimensional [2D]) comes with serious limitations in the description of stereocenters. For atomistic modeling tasks, molecular structures are often represented using Cartesian coordinates for the atomic positions. In chemoinformatic tasks, the most commonly used representation for storing chemical reactions is the Simplified Molecular-Input Line-Entry System (SMILES) notation,^{39,40,43} developed in the 1980s. In contrast to 2D graphs, stereochemistry can be encoded to a certain extent, resembling SMILES to something akin to linearized 2.5-dimensional graphs. Thanks to advances in chemical language models,^{27,36,44} SMILES recently became a popular machine input representation. In SMILES, all configurational isomers of a molecular structure are mapped to a single string. Their alpha-numerical string nature makes SMILES a versatile representation also for database and search queries. As shown in Figure 2, in reaction SMILES, “.” characters separate the molecules in reactions and “>” the reactants, reagents, and products. If the reactions are atom-mapped, the underlying transformations can be represented either as Condensed Graph of Reaction (CGR),^{45,46} reaction templates using reaction SMARTS,⁴⁷ SMIRKS,⁴⁸ or CHMTRN/PATRAN,⁴⁹ precursors plus bond edits²² or ReactionCode.⁵⁰ Alternative string-based representations exist, which can be converted from and to SMILES. An example being the self-referencing embedded strings (SELFIES),⁵¹ designed to ensure always syntactically valid molecules in generative tasks. However, despite the availability of more recent molecular representation formats, the simplicity of the SMILES makes them the primary choice in different public datasets of chemical reaction records.

Focusing on data availability, there are multiple datasets extracted from US Patents: the USPTO_MIT subset (0.5 M, without stereochemistry),^{21,22} USPTO_STEREO^{24,31} (1 M, with stereochemistry), USPTO_full (1 M, with stereochemistry, without reagents)⁵² to benchmark different reaction prediction and single-step retrosynthesis tasks. Beyond the general-purpose databases, there are a few more specialized ones. A dataset with 1000 reactions from 50 name reaction classes, generated with NameRXN,⁵³ was made available by Schneider et al.⁵⁴ In another study, Schneider et al.⁵⁵ published a dataset with reaction superclasses and high-quality atom-mapping, from which Liu et al.⁵⁶ derived the USPTO_50k (with stereochemistry, without reagents) benchmark set. Recently, Schwaller et al.⁵⁷ generated a reaction classification dataset called USPTO 1 k TPL using the 1000 most frequent reaction templates in the USPTO_STEREO as reaction classes. A more extensive and continually improved patent dataset called Pistachio is commercially available from the NextMove Software.⁵⁸ Other sources for chemical reaction data are commercial literature databases such as Reaxys,⁵⁹ SciFinder,⁶⁰ and Science of Synthesis⁶¹ with different levels of data quality and curation. However, bulk access to the reaction data in those databases is only given to selected research groups. A literature reaction dataset was recently made available by Jiang et al.⁶² In all the literature-extracted datasets, reagents, solvents, and catalysts are described as text, and the conversion to a structural representation is often challenging mainly due to the usage of non-IUPAC names. Although patent and literature datasets cover a significant chemical reaction space, the datasets tend to be biased toward positive, high-yield reactions.^{63,64} A challenge in literature and patent reaction databases alike are erroneously reported or mined reactions. To address this problem, Toniato et al.⁶⁵ developed an approach to automatically clean reaction datasets by looking at forgetting events during the training of reaction prediction models.

An alternative to comprehensive reaction datasets are the narrow reaction datasets that come from high-throughput experimentation (HTE) platforms, in which the influence of various conditions on the yield or the selectivity of a certain reaction is thoroughly investigated. Examples are the Buchwald–Hardwig reaction by Ahneman et al.,⁶⁶ the defluorination reactions by Nielsen et al.,⁶⁷ and the Suzuki–Miyaura reactions by Perera et al.⁶⁸

Recently, Kearnes et al.⁶⁹ introduced the Open Reaction Database (ORD), which provides a centralized platform to collect and access standardized reaction data from diverse sources. The USPTO patent, single-step batch, and many HTE experimentation datasets have already been added to the ORD. If well adopted by the community, such that reporting reaction data in a structured format (including negative data) becomes the new standard, it could have an immense impact on the performance of machine intelligence models for chemical reactions.

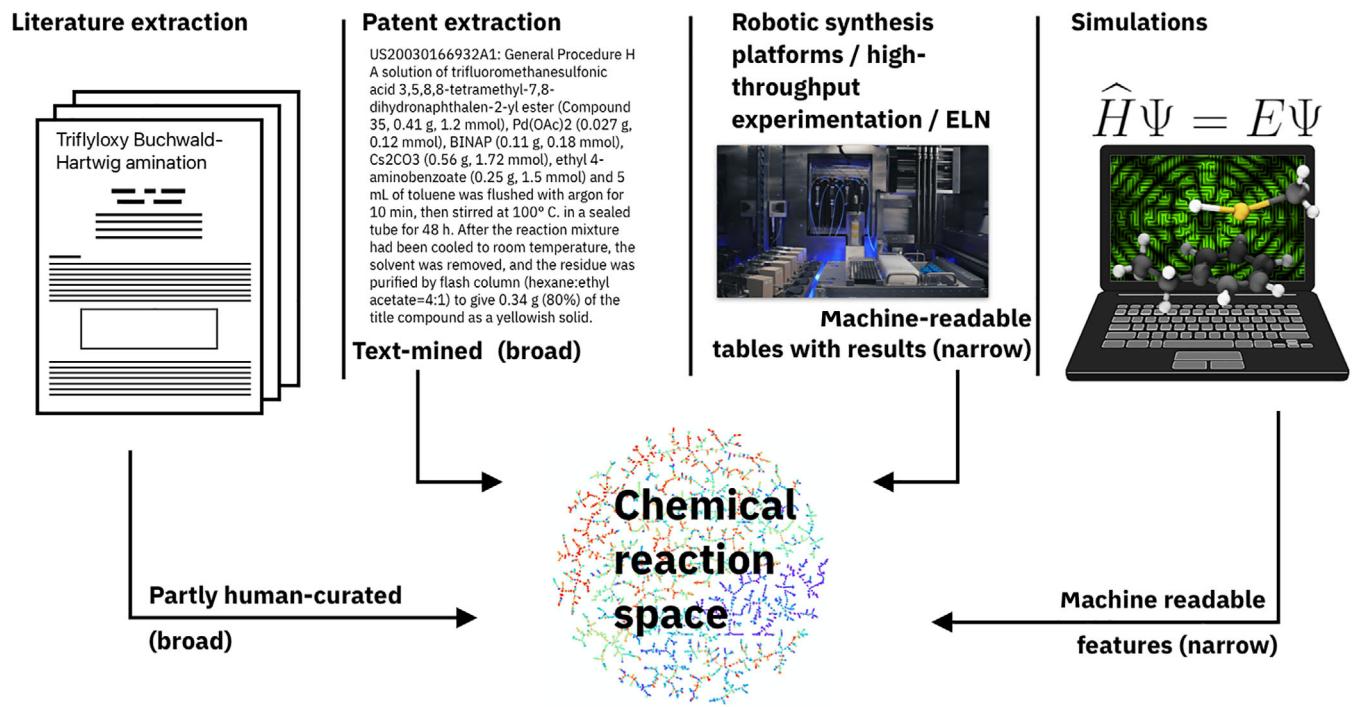


FIGURE 3 Principal sources of chemical reaction data including literature, patents, robotic platforms, and simulations

An overview of the different reaction dataset extraction pipelines and a typical reaction presentation is shown in Figure 3. Based on the chemical reaction datasets described above, numerous machine intelligence models tackling diverse chemical reaction tasks were developed. We review those chemical reaction tasks, the corresponding models and how machine intelligence can help accelerate chemical synthesis and molecular discovery in the next section.

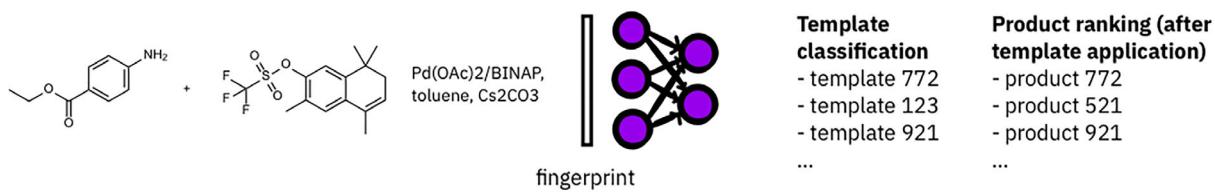
3 | CHEMICAL REACTION TASKS

3.1 | Reaction outcome prediction

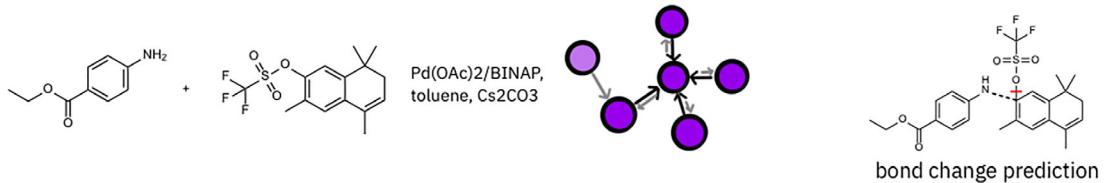
Predicting chemical reactivity and the outcome of chemical reactions is a fundamental task in organic chemistry. Validating the outcome of chemical reactions before performing the experiment in the laboratory enables chemists to save time and resources. In reaction prediction, the aim is to predict the most likely products given a set of precursors and possibly conditions. Techniques for reaction prediction before the deep learning era have been extensively discussed in previous reviews.^{70–73} Here we focus primarily on recent neural network-based approaches, which can be broadly categorized into: template-based, graph-edit-based, and sequence-based approaches (Figure 4).

Kayala et al.^{74,75} constructed an in-house dataset of elementary reactions. They developed a neural network model to identify electron sources and sinks and predict the reactions. The most recent version was built using 11 k elementary reactions.⁷⁶ Wei et al.⁷⁷ formulated the reaction prediction task as a classification problem and trained concatenated molecular fingerprints, that is, encodings of the structure of molecules, to predict the most likely out of 16 reaction templates. Similar to the work of Wei et al.,⁷⁷ but on a larger scale, Segler et al.⁷⁸ ranked the most favorable reaction rules out of 8820 templates extracted from the Reaxys database.⁵⁹ Instead of learning the precursors' fingerprints, the authors featurized the reactants by summing the individual Extended-Connectivity Fingerprints (ECFP4).⁷⁹ Because templates can lead to multiple matches, each of those generating more than one product, Coley et al.²¹ introduced the ranking of the products generated by the different templates. The predictions of template-based approaches are limited to the scope of reactions described by the underlying template set. These sets are built by automatically extracting the reaction center and neighboring atoms from atom-mapped reactions. During the extraction, there is a trade-off between the specificity of the templates, including the effect of long-range functional groups, and the total

Template-based approaches (atom-mapping dependent)



Graph edit-based approach (atom-mapping dependent)



Sequence-based approach (atom-mapping independent)

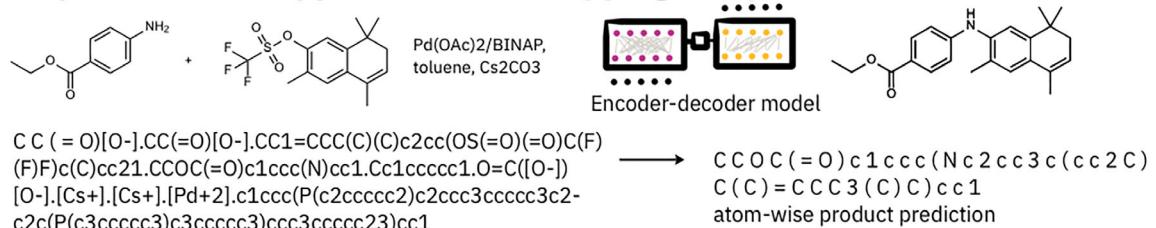


FIGURE 4 Main reaction prediction approaches

number of extracted templates. Taking into account more atoms around the reaction center leads to more specific reaction templates but also larger sets.

In 2017, Jin et al.²² introduced a graph convolutional neural network-based approach predicting bond changes in the precursors. Similar to templates, those graph edits were extracted from atom-mapped chemical reactions but did not require to specify a size around the reaction center. Bradshaw et al.²⁵ developed a gated graph neural network (GNN), Do et al.⁸⁰ a graph transformation policy network, and Nikitin et al.⁸¹ a Relational Graph Convolutional Neural Network to predict bond changes. Coley et al.²⁶ tweaked the graph-convolutional neural networks by Jin et al.²² allowing it to predict up to five bond changes in order to achieve better results on the USPTO_MIT dataset without stereochemical information.

The last category of reaction prediction approaches are sequence-based approaches, for which the precursors and products of the chemical reactions are represented as text, typically using SMILES. Models originally developed for translating between natural languages are trained to convert precursors into products, as first described by Nam and Kim.⁸² To generate input sequences, reaction SMILES are split into tokens in a process called tokenization. Schwaller et al.²⁴ tokenized reactants atom-wise and reagents molecule-wise to match the results of Jin et al.²² on the USPTO_MIT benchmark and demonstrated the potential of sequence-based approaches in reaction prediction. One advantage of sequence-based approaches compared to graph-based ones is that they can easily handle stereochemical information as long as it can be encoded in the underlying string notation. To date, the Molecular Transformer introduced by Schwaller et al.²⁷ is still the best performing reaction prediction approach on standard benchmark datasets with and without stereochemical information. One of the contributions of the work was not to make the distinction between reactants and reagents and predict reaction outcomes from a set of precursors. Making the reactants-reagents distinction requires knowing the product beforehand, information that should not be available in a realistic test set. This reaction representation makes it possible to train models on any reaction dataset without requiring atom-mapping. On the USPTO_MIT dataset, slight improvements over the original Molecular Transformer work were achieved using the same architecture with massive data augmentation and a computationally costly testing procedure.⁸³

Alternative reaction prediction methods have recently been developed. Qian et al.⁸⁴ studied the integration of symbolic rules to improve the performance of GNN-based approaches. Sacha et al.³⁰ used a graph representation of the

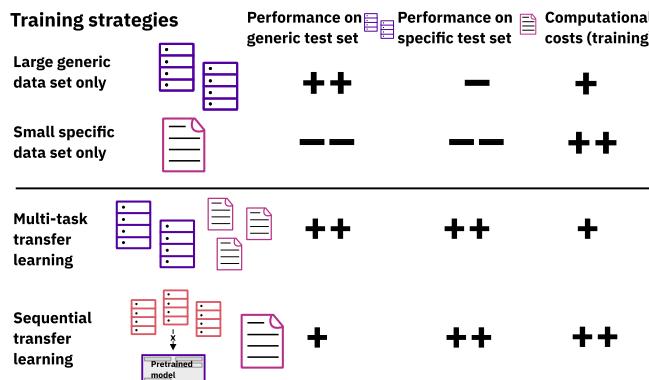


FIGURE 5 Training strategies including multi-task and sequential transfer learning. The “++,” “+,” and “–” are qualitative measures of the effects of the transfer learning strategies, as observed in References 31,88

precursors as input but then developed a graph sequence decoder that combined the advantages of GNN and sequence-based approaches by defining a canonical order for generating the output graph. Mann et al.⁸⁵ used a grammar-ontology-based representation to guarantee the validity of the predicted string. However, none of them reached the same performance as the Molecular Transformer.^{27,83} Transformer models capture well the SMILES grammar,³⁶ and even on a time-split test set containing stereochemistry, the fraction of syntactically invalid SMILES predicted by the Molecular Transformer amounted to only 0.5%.²⁷ Bi et al.⁸⁶ described a nonregressive electron redistribution model and evaluated it on USPTO/MIT reactions with reactant-reagent separation. Moreover, Kovacs et al.⁸⁷ demonstrated that many mistakes by the Molecular Transformer models can be explained by bias in the training data.

Going toward more challenging reactions and smaller training sets, Pesciullesi et al.³¹ studied transfer learning approaches to extend the predictive performance of the Molecular Transformer to regio- and stereoselective carbohydrate reactions using a small specific dataset. In transfer learning, the performance of a model on a specific task is improved by training the model first or simultaneously on a related task, from which knowledge can be transferred (Figure 5). Pesciullesi et al.³¹ tested the model prediction hypothesis using an experimentally accomplished lipid-linked oligosaccharide synthesis. Similarly, Wang et al.⁸⁹ studied transfer learning strategies on Heck reactions using a Transformer model, and Wu et al.⁹⁰ Baeyer–Villiger oxidations using a graph-convolutional neural network without stereochemical information. Seidl et al.⁹¹ investigated how modern Hopfield networks performed in template-based zero and few-shot learning.

Recent trends in using biocatalyzed reactions for sustainable chemistry have motivated Kreutter et al.⁸⁸ to extend the Molecular Transformer for enzymatic reactions. To do so, the reaction representation was modified to include the enzyme names. Concurrently, Probst et al.⁹² described the enzymes catalyzing the reactions by their Enzyme Commission numbers to mitigate the limited data volume for few enzymes by clustering them according to an expert classification Scheme.

3.2 | Synthesis planning

Forward reaction prediction models are crucial to evaluate the ability of data-driven schemes in learning reactivity patterns and signals. When scientists create compounds in silico, however, the synthesis of the target is frequently what determines a valuable design hypothesis. In synthesis planning, the aim is to find a route from a target molecule over multiple reaction steps back to commercially available or synthetically known building blocks. Retrosynthesis can be separated into the prediction of the individual steps and the search algorithm to complete the routes.²³

3.2.1 | Single-step retrosynthesis

Most machine learning-based synthesis planning approaches focus on single-step retrosynthesis and predict reactants leading to the target molecule. Segler et al.⁷⁸ investigated the same template-based method as for reaction prediction,

where a neural network selected the templates to apply. Concurrently, Coley et al.⁹³ studied template ranking based on molecular similarity. Both studies used the molecular fingerprint of the product as input to their models. Concurrently, Liu et al.⁵⁶ applied the first sequence-to-sequence neural networks to the task of single-step retrosynthesis. With the models originally developed to translate from one language to another language,⁹⁴ they transform the SMILES of the target into the SMILES of the most likely reactants. Since then, the USPTO-50k dataset⁵⁵ that Liu et al.⁵⁶ processed and split into well-defined train/validation/test sets became the main single-step retrosynthesis benchmark dataset. The reagents were removed from the reactions as part of the curation procedure, and the top-N accuracy was suggested as the evaluation metric. In retrospect, both choices led to an oversimplification of the single-step retrosynthesis task. In fact, the top-N accuracy is computed by checking whether the top-N suggestions contain the reaction suggested in the ground-truth (e.g., the reported patent reaction). However, numerous unreported reactions might exist to form the target compound. The single-step retrosynthesis benchmarks are currently heavily biased toward specific reaction classes. High top-N accuracy on those benchmarks does not necessarily translate to a good model performance when used for multi-step synthesis planning, as other criteria, like the diversity of the suggested reactions, play an important role.²⁸ Further simplifying the task, Tetko et al.⁸³ introduced the MaxFrag accuracy, where only the largest molecule of the precursors is checked.

Similarly to the work of Liu et al.⁵⁶ and inspired by the Molecular Transformer on reaction prediction,²⁷ Lee et al.,⁹⁵ Karpov et al.,⁹⁶ and Duan et al.⁹⁷ applied the Transformer model for the single-step retrosynthesis task. Later, multiple groups modified the Transformer-based approach to achieve better reactants' top-N accuracy. Zheng et al.⁹⁸ added a syntax correction model on top. Tetko et al. used data augmentation techniques at training, and test-time.⁸³ Kim et al.⁹⁹ coupled a retrosynthesis Transformer with a forward prediction Transformer and used a multinomial latent variable to generate more diverse reactant candidates. Ucak et al.¹⁰⁰ replaced the atom-wise tokenization with substructure tokenization, still relying on a Transformer-based model.

Instead of using a molecular fingerprint as an input to a neural network like Segler et al.,⁷⁸ Dai et al.⁵² inputted the molecular graph of the product to predict the most likely templates to apply using a conditional GNN. Chen et al.¹⁰¹ added a global reactivity attention layer to their GNN to better encode nonlocal reactivity dependencies. However, stereocenters remain a challenge and are neglected.

Shi et al.¹⁰² developed a graph-to-graph approach, in which a first model predicts synthons (or the reaction center) from the product and the second model to predict reactants starting from the synthons. Somnath et al.¹⁰³ framed the task along the same lines and allowed their synthon completion model to rank the best leaving group out of a fixed set.

Recently, several studies have begun to examine the duality between SMILES and molecular graphs and brought sequence-based and graph-based approaches closer together. Seo et al.,¹⁰⁴ for example, introduced a new attention algorithm called Graph Truncated Attention, which makes use of the molecular graph to define attention masks in the Transformer. Wang et al.¹⁰⁵ investigated a Transformer-based approach closely related to graph-based approaches, which separate the single-step retrosynthesis into synthon prediction and synthon completion. Sacha et al.,³⁰ in contrast, developed a graph-based approach to predict the reactant graphs as a canonical sequence of graph-edits.

Other studies focused on the improvement of fingerprint-based approaches. Thakkar et al.¹⁰⁶ developed the "Ring Breaker" to enhance the prediction of ring-forming reactions. Fortunato et al.¹⁰⁷ increased the diversity of the reactants suggested by the template ranking neural network by data augmentation. Seidl et al. investigated zero-shot and few-shot prediction capabilities of Modern Hopfield networks.⁹¹

All single-step retrosynthesis studies discussed above aim to predict the reactants or largest reactant given a product. In some sense, they are all atom-mapping dependent, as it was used to remove the reagents from the reactions in the underlying dataset. To date, Schwaller et al.²⁸ presented the only approach to predict all precursors simultaneously (Figure 6, top). The advantage of such an approach is not only the independence from atom-mapping, but also that the model predictions provide chemists with more information on how to perform the reactions.

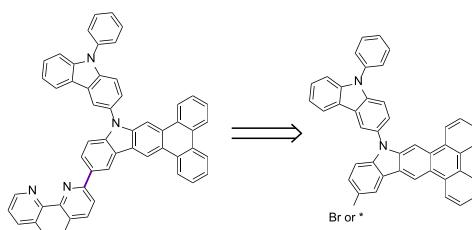
How to best evaluate single-step retrosynthetic models remains an open question. Evaluation metrics based complementary models have been proposed,^{28,65} but ideally, the predictions would be checked by human experts or performed in the lab.

3.2.2 | Multi-step approaches

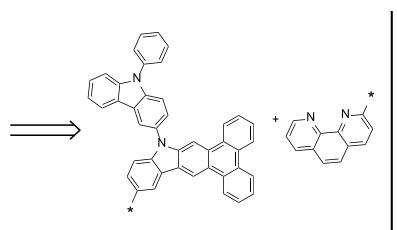
Single-step retrosynthesis methods are restricted in their use as the majority of synthesis pathways require many reaction steps to resolve the route and find commercially accessible compounds. The primary challenge is multi-step

Single-step retrosynthesis tasks

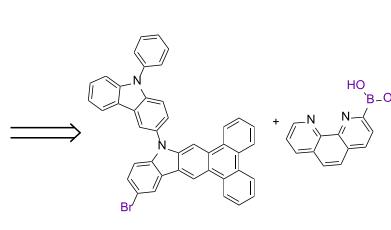
Largest fragment/reactant/synthon prediction



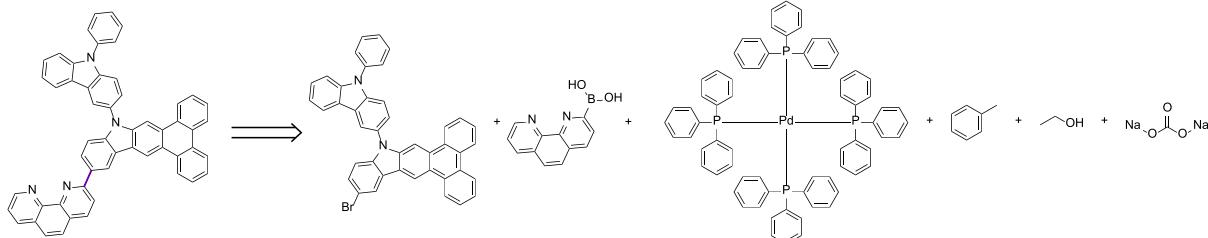
Synthons prediction



Reactants prediction / synthons completion



Precursors (including reagents, catalysts, solvents) prediction



Multi-step synthesis planning tools

2016

2017

2018

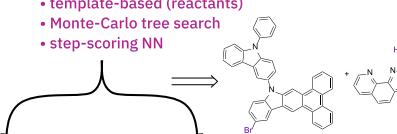
...

Closed

- expert rule-based
- "wide & deep" search
- simulations, ML scoring

Segler & Waller
Nature, 2018

- template-based (reactants)
- Monte-Carlo tree search
- step-scoring NN



Other proprietary tools

- MoleculeOne
- ChemicalAI
- PostEra
- Iktos
- DeepMatter
- SciFinder
- Reaxys
- ...

Freely accessible

- LillyMol**
Watson et al.
J. Cheminf., 2019

- template-based
- template-freq score

- ASKCOS**
Coley et al.,
Science, 2019

- template-based (reactants)
- Monte-Carlo tree search
- step-scoring NN

- AiZynthFinder**
Thakkar et al., *Chem. Sci.*, 2020
Gehneden et al., *J. Cheminf.*, 2020

- template-based (reactants)
- Monte-Carlo tree search
- step-scoring NN
- interactive mode

- IBM RXN for Chemistry**
Schwaller et al.
Chem. Sci. 2020

- template-free (precursors)
- hyperbeam search
- step-scoring NN
- interactive mode
- stereochemistry

...

2019

2020

2021

FIGURE 6 Top: single-step retrosynthesis prediction tasks of varying difficulties: largest reactants, synthons, synthons plus completion, reactants, and precursors. Bottom: development of machine intelligence-driven synthesis planning tools

retrosynthesis. It is uncertain how evaluation metrics such as the top-N accuracy, primary objective in all current single-step studies, transfer to the multi-step setting.

In 2018, Segler et al.²³ pioneered the use of a template-based single-step approach in conjunction with a Monte-Carlo tree search (MCTS) to plan multi-step synthesis pathways. They scored individual reactions and pruned the tree using a neural network. Due to the template-based methodology, the suggested reaction steps contained only reactants. The performance was assessed with a chemical Turing test: 45 graduate-level organic chemists had to choose their preferred route among computer designed and literature pathways for nine compounds in a blind experiment. This test demonstrated that the predicted routes were comparable to the literature routes. There was no clear preference for one of the two sources over the other.

Based on the ideas of the LHASA knowledge base,⁴ human experts in the Chematica/Synthia project manually collected more than 100 k reaction rules over the last two decades.^{16,17,32} Recent work added diverse filters to the rules, including, for example, ML-based regioselectivity predictions in Diels-Alder reactions,¹⁰⁸ and made the closed-source platform a hybrid expert–AI system.

Similar to the work of the Chematica/Synthia,^{16–18} which successfully predicted natural product pathways in a recent study,³² Segler et al.'s work²³ was not made open-source or available through an application programming interface, which prevented the comparison to other approaches. Coley et al.¹⁰⁹ with ASKCOS, and Thakkar et al.^{29,110} with AIZynthfinder have built computer-assisted synthesis tools inspired by the work of Segler et al.²³ and made them available as open-source. Another open-source template-based approach was described by Watson et al.¹¹¹ To date, only two studies have used Transformed-based models to suggest the reaction steps. While Lin et al.¹¹² focused on reactants using an MCTS algorithm, Schwaller et al.²⁸ directly predicted precursors sets and explored the multi-step hyper-graph with a beam search. In the work of Schwaller et al.,²⁸ a forward prediction Molecular Transformer model²⁷ is coupled with the retrosynthesis model to validate and rank the single steps. Recently, Probst et al.⁹² extended the same approach to enzymatic reactions. The two latter approaches are integrated into the RXN for Chemistry platform,^{28,113} a synthesis planning platform comprising trained machine learning models (Figure 6, bottom). To facilitate the adoption by synthetic chemists, the models can be used without coding experience. Other efforts to enhance the accessibility to synthesis planning models were made by AiZynthFinder^{29,110} and ASKCOS.¹⁰⁹

Regarding search algorithms, MCTS, as suggested by Segler et al.,²³ is the most frequently used, as can be seen from Table 1. Wang et al.¹¹⁹ compared the different MCTS variants and incorporated a solvent greenness score into the search. Alternative search strategies exist. Schreck et al.,¹¹⁴ for example, investigated learned strategies for template ranking using reinforcement learning. Shibukawa et al.¹²⁰ studied the full enumeration of all possible synthesis routes given a set of templates. In their experiments on the 20 molecules benchmark by Heifets et al.,¹²¹ only half were successfully resolved. Guo et al.¹¹⁷ combined a graph logic network-based template suggestion approach¹¹⁷ with a Molecular Transformer²⁷ to then use a sequential Monte Carlo (SMC) search. Recently, Chen et al.¹¹⁵ have demonstrated that an A* search algorithm performed better than others in template-based retrosynthesis planning. Their algorithm was recently extended by Kim et al.¹¹⁶ to simultaneously optimize the search and the single-step model in an end-to-end training.

Template-based synthesis route approaches can be used to suggest thousands of routes leading to the same molecule. As the ranking may not be ideal, it can be overwhelming to select the best one, and which is the best route may also depend on the target molecule. Gao et al.¹²² applied mixed-integer optimization for reducing the amount of

TABLE 1 Different retrosynthetic prediction approaches

Input (product)	Output	Single-step approach	Multi-step algorithm
<i>Sequence-based</i>			
SMILES	Reactants SMILES	Seq-2-seq LSTM	None: 56
SMILES	Largest reactant SMILES	Transformer	None: 83
SMILES	Reactants SMILES	Transformer	None: 83,95–99, MCTS: 112
SMILES	1: Synthons prediction, 2: Synthons completion	Transformer	None: 105
MACCS keys	Reactants MACCS keys	Transformer	None: 100
SMILES	Precursors SMILES	Transformer	Beam search: 28
<i>Fingerprint-based</i>			
Fingerprint	Reaction template	Similarity	None: 93
Fingerprint	Reaction template	Feed-forward NN	None: 78,106,107, MCTS: 23,29,109,110, RL: 114, A*: 115,116
Fingerprint	Reaction template	Modern Hopfield network	None: 91
<i>Graph-based</i>			
Molecular graph	Reaction template	Graph neural network	None: 52,101, SMC: 117
Molecular graph	1: Synthons prediction, 2: Synthons completion	Graph neural network	None: 102,103,118
Molecular graph	Sequence of graph-edits	Graph neural network	None: 30

commercially available precursors required to synthesize World Health Organization (WHO) essential medicines. In another work, Gao et al.¹²³ used similar techniques to find overlapping routes for multiple target molecules. Along the same line, Weber et al.¹²⁴ identified strategic molecules in large reaction networks. Mo et al.¹²⁵ recently developed another approach to cluster and rank synthesis pathways. To do so, they extracted pathways from the Pistachio database by matching reactants with products in reactions from the same patent. They then trained a Tree-LSTM model, that is, a generalization of a long short-term memory (LSTM) network to tree-structured network topologies, to recognize whether routes were predicted by their ASKCOS tool or originated from the patents.

As the aim of multi-step retrosynthesis is to find routes starting from commercially available molecules, it is essential to have the same set of commercially available molecules when comparing different approaches. Moreover, as not all possible disconnections and functional group interconversions are reported in the literature or patents, a fair evaluation of synthesis routes still requires the help of human experts.^{23,28}

3.3 | Reaction conditions and experimental procedures

A retrosynthesis offers very useful insight about the accessibility of a molecular template from existing building blocks. However, the retrosynthetic design (Section 3.2) lacks essential details for the execution of the synthesis in wet-lab experiments. Details like reaction temperatures and duration, operations specific to given classes of compounds or reactions, and, often, which solvents and reagents to use are decisions under the responsibility of the domain experts. Traditionally, chemists know from experience how to conduct reactions, or they look up closely related transformations from the literature or reaction databases. Even then, chemists may need to adapt the procedures depending on the nature of the functional groups characterizing the compound under investigation.

With the development toward automated synthesis and the desire to increase synthesis throughput, different groups recently developed machine learning models to suggest experimental conditions in an automated manner. Unlike in reaction optimization (Section 3.5), these approaches are noniterative in nature and may learn from different regions of chemical reaction space.

In 2018, Gao et al.¹²⁶ trained a neural network on several million organic reactions extracted from Reaxys for the consecutive prediction of the reaction catalyst, up to two solvents, up to two reagents, and the temperature. The model is independent of the reaction class and predicts the conditions starting from the product and reaction fingerprints. Walker et al.¹²⁷ concentrated on the prediction of the solvent for five selected reaction classes. They explored three models for this task, focusing on the predictability for this restricted chemical space. Maser et al. formulated the prediction of reaction conditions as a multilabel classification model.¹²⁸ They considered four reaction classes, for which they predicted between 5 and 8 labels each, including compound identity (metal, ligand, base, additive, and solvent), intervals for temperature and pressure, as well as the presence or absence of gaseous carbon monoxide. Vaucher et al.,¹²⁹ instead, defined the task as a prediction of the experimental steps that a chemist would perform in the lab. In addition to predicting reaction parameters such as the temperature or reaction duration, this formalism enables the machine learning model to predict when specific operations are necessary, such as filtration, phase separation, and extraction, or the dropwise addition of compounds depending on the feature of the precursors and target molecules. The model is trained on experimental procedures extracted from patent data.¹³⁰ Unlike the other approaches, this model expects the input reaction to specify all the species involved in the reaction (including solvents and catalysts).

3.4 | Reaction performance prediction

While the outcome of a chemical reaction may be predictable, changes in the conditions, catalysts, or solvent might result in significantly diverse results. This is frequently the case for common catalyzed processes of substantial industrial interest, where even little improvements could result in significant benefits. In such circumstances, the goal is to capture the functional dependence between the inputs parameters and the reaction outcomes. Compared to other chemical reaction tasks, reaction performance models tend to be highly specific for a single reaction or a family of reactions, capturing the effects of quantitative variations in recipe ingredients on the overall performance.

A single scalar value (e.g., the reaction yield) can generally describe the performance of a reaction, and thus a reaction performance task can be framed as a regression problem. Historically, several linear models stemming from physical organic chemistry have successfully unveiled trends and qualitative predictions. More sophisticated multi-variate



and nonlinear models have been developed in the last decade with ever-increasing accuracy and applicability. Depending on the application of interest and the availability of data, different performance metrics can be targeted. Coarsely, reaction performance prediction models can be divided into four groups depending on the target: reaction yield, activity, selectivity, and activation energy.

3.4.1 | Yield prediction

The yield of a chemical reaction is perhaps the most direct way to evaluate the reaction performance, and, for this reason, it has been one of the primary targets of recent ML works on experimental data. Ahneman et al.⁶⁶ used lab automation to generate a yield database (in the 0% to 100% range) of over 4000 single step C–N cross coupling reactions, which then served to train different ML models exploiting 120 selected descriptors. They computed a plethora of molecular (e.g., E_{HOMO} , E_{LUMO} , dipole moment), atomic (e.g., nuclear magnetic resonance [NMR] shifts), and vibrational descriptors characterizing the reaction components (aryl halide, additive, Pd catalysts, and base).

A point of controversy¹³¹ was the type of featurization (i.e., how and which descriptors were selected) potentially lacking transferability and not being superior to alternative fingerprints without chemical information. Comparisons with noninformative models revealed that using representations containing chemical descriptors led to improved models for predicting the range of the yield distribution, hence providing more robust extrapolation in unseen samples. Nonetheless, the importance of random-control tests to assess the model generality was pointed out and acknowledged by Doyle and coworkers.¹³²

Since then, several groups attempted to derive improved models trained on the same cross-coupling yield database.⁶⁶ Instead of using manual featurization to generate the representation, Sandfort et al.¹³³ proposed automated molecular fingerprints from RDKit (the Multiple Fingerprint Features method) and achieved a comparable accuracy. Schwaller et al.⁶⁴ took an alternative approach and exploited the reaction SMILES as a representation together with the molecular BERT Transformer architecture. This strategy was shown to outperform both previous approaches, with the implication that reaction SMILES might be a very complete representation. This model combined with data augmentation techniques showed improved robustness in the low-data regime and allowed the estimation of the prediction uncertainty.¹³⁴ Haywood et al.¹³⁵ further analyzed the different representations and compared molecular fingerprints with computed descriptors using support vector machine (SVM) and found fingerprints to give better results.

Using the same strategy as initially applied to the cross-coupling database, Nielsen et al. exploited a much smaller yield database of ~700 deoxyfluorination reactions.⁶⁷ Due to the sparsity of data, the yield prediction was less accurate, but the model was shown to remain useful for reaction optimization.

As a result of the great interest on the use of HTE frameworks, a larger number of experimental reaction yields will be available in the future, which will also catalyze the development of more general and sophisticated machine learning models. In that direction, active learning approaches to train more data-efficient models have been developed by Eyke et al.¹³⁶ from available high-throughput experimental data.^{68,137} Accurate yield predictions for the yield of Pd-catalyzed Suzuki–Miyaura cross-coupling were obtained with a reduced number of training instances using active learning and a neural network-based on fingerprints. Similarly, Granda et al.¹³⁸ described an approach for searching for new reactivity and evaluated their models on the same Suzuki–Miyaura reaction dataset.

3.4.2 | Activity

In addition to yield, the increase of the reaction rate (activity) with the use of a catalyst is another appealing target, accessible both experimentally and computationally. The availability of activity databases is limited by the need of performing time-resolved experiments or modulating reaction conditions to determine reaction rates (e.g., via an Arrhenius plot) while carefully tackling masking effects (e.g., per site activity versus the number of exposed sites in heterogeneous catalysts).

Within this context, Smith et al.¹³⁹ trained a neural network to predict rate constants of the water gas shift reaction from a database of 2228 reactions extracted from the literature. They used 27 experimental descriptors, including reaction conditions (e.g., temperature and calcination time) and chemical information (e.g., catalyst loading and electronegativity) to build representations, obtaining remarkable learning curves. However, their findings suggested that the chemical diversity in the dataset was suboptimal, with only 187 different formulations.

Relying on a molecular volcano description of the computed turn over frequency,¹⁴⁰ Cordova et al.¹⁴¹ trained Kernel-based machine learning models to predict the turnover frequency (TOF) of Ni catalysts featuring phosphine and carbene ligands for aryl-ether cleavage reactions. Having rapid access to such a large number of energy-based values correlating with the TOFs demonstrated that only a handful number of ligands possess particular characteristics to align with the top (i.e., optimum) region of the volcano. In other words, this work further stresses that identifying the best-suited metal–ligand combination for this reaction remains a challenge. The same ML-Volcano approach has also been employed to extract the C–C cross-coupling catalysts with the best thermodynamic profiles from a database of 18,062 ligand/metal combinations.¹⁴²

Alternative examples exist in the field of electrocatalysis with the work of Zhong et al.,¹⁴³ who developed models to predict the CO binding energy activity for CO₂ electrocatalysts, which leads to the TOF through a volcano plot. An automated ML model was used in combination with vectorized atomic environment representations of copper-containing intermetallic surfaces, starting from a database of 19,644 density functional theory (DFT)-computed values. The model was used as an active learner, iteratively selecting promising surfaces for a further computational investigation that were added to the training set. After several cycles and over 4000 additional DFT computations, they identified promising experimental candidates, which were then synthesized. In 2015, Yang et al.¹⁴⁴ proposed a pioneering work in the same direction with a database of ~250 datapoints, which has since then been revisited for alloy catalysts for carbon dioxide reduction reaction. Gradient Boosting Regression was shown to outperform several other methods for the prediction of the binding energy. The possibility of linking binding or adsorption energies with the catalytic activity of surfaces has sparked further interest in such models.^{145,146}

3.4.3 | Selectivity

Once a catalyst is identified as active, its selectivity (chemo-, regio-, site-, diastereo-, enantio-) toward a given product is often the most desirable property in catalyzed reactions. In contrast to yield and conversion, the desired pathway often requires a kinetic advantage over closely related competing alternatives. Selectivity can be expressed by the ratio between the desired product and the entire set of possible outcomes.

The RegioSQM¹⁴⁷ is the first example of an automated semi-empirical protocol predicting the regioselectivity of electrophilic aromatic substitutions (i.e., on what atom the substitution will take place). More recently a GNN architecture¹⁴⁸ was presented by Struble et al. The ML model uses a representation based on SMILES graphs of the reactants and cheminformatics atom descriptors from RDKit (e.g., Gasteiger charges). The GNN was shown to be similar in quality but significantly faster than RegioSQM. Both approaches, being more specific, were shown to be better than a completely general approach.²⁷ The GNN architecture of Struble et al.¹⁴⁸ has been further extended to more general regioselectivity problems using a multitask architecture. The connectivity graph information from SMILES is used to generate a set of atomic descriptors (e.g., Fukui indices, atomic charges) as a first task, trained on quantum chemical data; then uses both sets of information to predict regioselectivity.¹⁴⁹

Site selectivity was also the focus of Li et al.,¹⁵⁰ who trained several random forest (RF) models using different representations to predict the carbon atom to which the radical will attach in the context of radical C–H functionalization. The performance of the models using either physical organic selected descriptors (RF model) or structure-based descriptors (SOAP) plus molecular fingerprints (XGBoost model) was excellent; although the RF with selected descriptors showed better generality.

Another example of a classification task is the model by Sunoj et al.,¹⁵¹ which predicts the product outcome of a difluorination reaction on alkenes (i.e., either the 1,1 or the 1,2 difluorinated product as a major outcome). The substrate was represented by expert-crafted computed descriptors for a database of fewer than 100 points. Despite the size limitations, the model showed >90% accuracy.

A larger database of 6355 Diels–Alder reactions taken from Reaxys was exploited by Beker et al.,¹⁰⁸ who built several ML models (a RF model performing the best) to predict the regio-, site-, and diastereoselectivity. Each of these models relied upon different descriptors (e.g., Hammett indices and steric indices, and fingerprints) depending on the selectivity target. Interestingly, the out-of-sample predictions were more accurate when using physically meaningful descriptors.

In addition to predicting the correct output structures, additional efforts were placed on defining and predicting scales that correlate with the product outcome. Tavakoli et al.,¹⁵² for instance, proposed a general approach to define atomic reactivity through the methyl cation and methyl anion activity scales. These scales aim at measuring the acid or

base-like behavior of atoms in functional groups and indirectly the chemoselectivity of a reaction. These two indices, which were computed for a diverse set of 2421 atoms in different environments, were shown to be well reproduced by a GNN architecture using SMILES inputs.

Regression tasks have also been relevant to the prediction of selectivity. Moon et al.¹⁵³ recently quantified the stereoselectivity as the percentage of alpha product using an RF model trained on a dataset of 261 experimental data points from glycosylation reactions, combined with selected features describing steric and electronic properties of all reaction participants, including solvent.

Finally, in the context of enantioselectivity, Zahrt et al.¹⁵⁴ developed a pipeline for the prediction of $\Delta\Delta G$ of enantioselective reactions. They introduced three dimensional (3D) features encompassing conformational and symmetrical information (e.g., the average steric occupancy) in addition to a plethora of selected features for training their ML support vector regression model on 1075 datapoints corresponding to N,S-acetal formation reactions. Their $\Delta\Delta G$ predictions were found to be remarkably accurate (mean MAD = 0.15 kcal/mol). The relevance of chemically meaningful features was shown by attempting extrapolative predictions using randomized labels, which led to worsened performance.¹⁵⁵

Huang et al.¹⁵⁶ focused on improving the featurization of the Spectral London Axilrod–Teller–Muto (SLATM) structure-based representation in order to develop and improve the performance of kernel ridge regression models for the enantioselectivity prediction. They introduced reaction-based representations and exploited metric learning and supervised feature selection techniques to filter the information contained in the molecular representations. This featurization was shown to dramatically improves the performance of the machine learning model as illustrated by the ML prediction of the DFT-computed enantioselective excess of a Lewis base-catalyzed propargylation reaction.¹⁵⁷ High-quality out-of-sample predictions also confirmed the transferability of reaction-based representations.

3.4.4 | Activation energy and transition states

The computational access to mechanistic pathways leads to an ensemble of properties and intermediates that could be correlated with experimental outputs and exploited with a data-driven approach. In this respect, activation energies, especially those of the rate-determining transition states (TSs), are especially relevant. However, activation energy reference datasets, which are most often computed with quantum chemistry methods, can be fairly computationally demanding. In 2018, Kim et al.¹⁵⁸ tried to bypass TS computations by directly predicting the activation energies using reaction-based representations obtained from fingerprints, thermodynamic quantities, and topological data of reactants and products for 6078 reactions from the RMG-py database.¹⁵⁸ Using a gradient boosting model, they were able to learn activation energies with an root mean squared error (RMSE) of 4.49 kcal/mol. Within the dataset, degrading accuracy was noted as the number of bonds broken or formed increased.

Heinen et al.¹⁵⁹ and von Rudorf et al.¹⁶⁰ constructed models for activation energy prediction using a high-quality database of E2 and S_N2 reactions. Starting from 3D structure-based representations (SLATM, FCHL19)^{156,161} of the reactant complex, an mean absolute error (MAE) as low as 2.5 kcal/mol could be obtained but using random one-hot labeling as a representation was shown to worsen the out-of-sample performance. Alternatively, the activation energy of a dataset of 2574 dihydrogenation reactions of Ir-containing Vaska's complex was the target of a Gaussian process regression (GPR) model trained by Friederich et al.¹⁶² Autocorrelated functions of selected atom and bond descriptors served as representations, which were supplemented with fingerprints.

A related effort on hydrogenation reactions has also been placed in the field of heterogeneous catalysis. Singh et al.¹⁶³ developed a neural network, stemming from 315 computed data points for dehydrogenation, N₂ and O₂ dissociation steps. Using as few as seven selected descriptors, they were able to predict activation energies with an MAE of 4.5 kcal/mol, without much improvement upon simpler linear models in this case.

While all the above models are highly reaction-specific, Grambow et al.¹⁶⁴ covered a much broader range of chemical reactions using their own quantum chemical database of over 5×10^4 computed reactions and activation (zero point energy [ZPE]-corrected, electronic) energies, spawning from 0 to 200 kcal/mol. They trained an message-passing neural network model based on atom-mapped SMILES of reactants and products as a representation.¹⁶⁵ Their predictions of activation energies¹⁶⁶ lead to a satisfactory 3.4 kcal/mol RMSE.

Instead of targeting the prediction of computed activation energies, Jorner et al.¹⁶⁷ collected 443 existing experimental activation energies associated with nucleophilic aromatic substitution reactions to develop a hybrid model combining quantum chemical computation and ML, attempting to bypass possible underlying deficiencies in DFT

computations. Starting from SMILES, their workflow first performs an automated reaction mechanism exploration in semiempirical + DFT prior to building the GPR model to correct the DFT activation energies to match the experimental values. Conceptually, their work established a hierarchy of methods for the *in silico* prediction of activation energies, ranging from fully *ab initio* quantum chemistry (no assumptions) to entirely data-driven ML (requires extensive data and assumptions), in which the intermediary alternatives offer promising compromises.

Locating the TS in the potential energy surface is usually a challenging task that requires extensive human intervention. The previous models thus aim to avoid using the TS information for the prediction of activation energies. Still, other models attempt to identify the geometry of the TS using ML instead of the usual quantum chemistry approach. Pattanaik et al.¹⁶⁸ trained a GNN to predict TS geometries of isomerization reactions starting from the reactant and product geometries. This approach uses the GNN to interpolate the distance matrix of the TS, which is then reconstructed back to Cartesian coordinates. Along the same line, Lemm et al.¹⁶⁹ have put forward a general geometry prediction protocol, called graph to structure, which is capable of tackling TS geometries as well if trained appropriately.

In related work, machine learning has been applied to identify reaction coordinates instead of directly determining TSs. This is usually achieved by extracting such reaction coordinates from simulation data, or by enhancing sampling schemes.^{170–172}

3.5 | Reaction performance optimization and discovery

Given that reaction performance can be accurately predicted, using predictive models to improve reaction performance is an immediate next step. Instead of starting with a fixed dataset of reaction conditions, it is natural to use the predictive model to iteratively propose new experiments with new conditions that yield promising performance. One central challenge here is the so-called *exploration-exploitation dilemma*. While machine learning approaches typically generalize well to reaction conditions similar to those in the training set, they often fail for very different conditions to those represented by the data. Thus, it is necessary to, on the one hand, *explore*, by carrying out experiments with uncertain outcomes in order to obtain relevant data, while at the same time *exploiting* the collected data by focusing on reactions with plausibly high performance (yield, activity).

3.5.1 | Bayesian optimization

One approach that has gained considerable popularity is Bayesian optimization (BO).¹⁷³ Here, one places a prior (often a Gaussian process¹⁷⁴) on the unknown performance function. Outcomes of reaction experiments are incorporated via Bayes' rule, and the uncertainty captured by the posterior distribution over the unknown objective is used to guide the exploration-exploitation tradeoff. The latter is done by solving a surrogate optimization problem, which optimizes an *acquisition function* (such as the expected improvement¹⁷⁵ and upper confidence bound¹⁷⁶).

In an early application to the chemical domain, Romero et al.¹⁷⁷ used BO to optimize functional properties of macromolecules.

In the context of reaction optimization, Hase et al.¹⁷⁸ presented Phoenics, an approach building on ideas from BO and kernel density estimation, and demonstrated it on reverse engineering initial conditions and reaction constants for the Oregonator reaction. Burger et al.¹⁷⁹ used BO to optimize 10 parameters of photocatalysts mixtures. They integrated BO into a robotic platform that autonomously carried out 688 experiments over 8 days. Shields et al.³³ compared BO approaches to classical (noniterative) experimental design techniques, as well as to experiments proposed by human expert chemists. For the latter, they implemented an online game where 50 experts from academia and industry were asked to design experiments to identify best-performing configurations. On the reported reaction, BO outperformed human experts on average. Felton et al.¹⁸⁰ reported a benchmark study comparing BO to other experimental design strategies in terms of their ability to identify effective trade-offs when simultaneously optimizing multiple performance objectives.

3.5.2 | Reaction-guided molecular structure optimization

Next to optimizing the performance of chemical reactions by identifying optimal reaction conditions, machine intelligence has been applied to discover and generate new molecule structures. A key challenge here is that the search space



is vast and unstructured, and appropriate techniques are needed to characterize and navigate the search space. The field of de-novo design of molecules is too extensive to survey exhaustively.¹⁸¹ We will focus on examples of the primary molecule generation techniques and studies that used chemical reactions to ensure that the generated structures are synthetically accessible.

An approach toward exploring the space of molecules uses *generative models*. Starting from an existing dataset, generative models aim to infer the data-generating process. They seek to learn a (typically probabilistic) model that can generate novel data items similar to but different from those in the original dataset. While generative modeling is a classical approach widely employed in statistics, in recent years, approaches based on deep neural network models have led to remarkable advances in generative modeling of complex high-dimensional such as images. The key idea is to learn a neural network that takes a known, simple distribution (e.g., a multivariate Gaussian distribution) and transforms it nonlinearly to resemble the distribution of the data. A key challenge is that the resulting probability density cannot efficiently be evaluated, and as a consequence, classical techniques like maximum likelihood estimation do not apply. Instead, alternative training criteria have been developed. Variational autoencoders (VAEs)^{182,183} for example instead maximize a variational lower bound on the model evidence. Generative adversarial networks (GANs)¹⁸⁴ treat the estimation problem as a game between a generator and a discriminator, a second neural network that aims to distinguish the generated “fake” examples from real ones.

In chemistry, GANs and VAEs recently found applications to propose potential drug candidates.^{185,186} Gómez-Bombarelli et al.,¹⁸⁷ for example, use VAEs to map the SMILES representation of molecules into a continuous latent space, which is additionally constructed to be predictive of the target property. BO is then used on the latent space, and concrete molecules are proposed by decoding into the space of SMILES strings, filtering out those leading to invalid molecules. In order to remedy the issue of generating invalid molecules, Boitreaud et al.¹⁸⁸ replaced previously used SMILES representations with recently introduced SELFIES strings,⁵¹ resulting in a graph-to-SELFIES VAE module called OptiMol. While molecular properties can be optimized in silico, the aim should be to experimentally synthesize and test the molecules. However, synthesizability is neglected in most de novo design studies.¹⁸⁹ In some exceptions, chemical reaction information is used to improve the synthetic accessibility of generated molecules. Button et al.,¹⁹⁰ for example, applied expert-selected reaction templates and Bradshaw et al.¹⁹¹ reaction prediction models to design the molecules. Later, Bradshaw et al.¹⁹² extended the search for new molecules to multi-step forward synthesis steps, where the molecules are constructed through a sequence of building block addition and reaction steps. With ChemBO, Korovina et al.¹⁹³ implemented a BO framework for generating and optimizing organic molecules for desired molecular properties, incorporating additional criteria such as the synthesizability of molecules into their sample-efficient search procedure. Recently, Grisoni et al.³⁴ designed and synthesized potent liver X receptor agonists biasing the generative model toward synthesizable molecules using virtual reaction filters.

A downside of generative models is that they need a representative dataset to begin with and can only generate data items similar to those in the dataset. A promising alternative is *reinforcement learning* (RL),¹⁹⁴ which provides systems the ability to learn from experience by trading exploration and exploitation of an environment without direct supervision. RL considers an agent that aims to maximize reward in an environment described by a Markov Decision Process.¹⁹⁵ In recent years, fueled by advances in deep learning, RL has seen remarkable successes, in particular for games and robotic tasks.^{196,197} Simm et al.¹⁹⁸ described a general framework for modeling molecule design as an RL problem. The state models the positions of atoms placed in a 3D conformation. The agent's actions modified an existing conformation by placing new atoms. The reward function was designed to capture physical principles related to the energy, which is approximated by fast quantum chemical calculations. Gottipati et al.¹⁹⁹ described an RL framework for constructing molecules via forward synthesis. Here, the state identified the current molecule, and the agent's actions corresponded to the choice of reactant. As reward functions, Gottipati et al.¹⁹⁹ optimized the Quantitative Estimate of Druglikeness Bickerton et al.²⁰⁰ and the calculated partition coefficient (clogP)²⁰¹ of the constructed molecule.

3.6 | Reaction featurization and classification

Reaction classification has been of interest to chemists for many decades.²⁰² Many reaction classification schemes, especially early ones, relied on an inspection of the reaction center based on predefined patterns, and will not be detailed further in the present work. For data-driven classifiers, data representation does not only play a role in terms of how the considered reactions are reported and stored (Section 2), but also under which format the classification algorithms ingest them. For this, it is typically necessary to provide a set of features common to all chemical reactions and hence

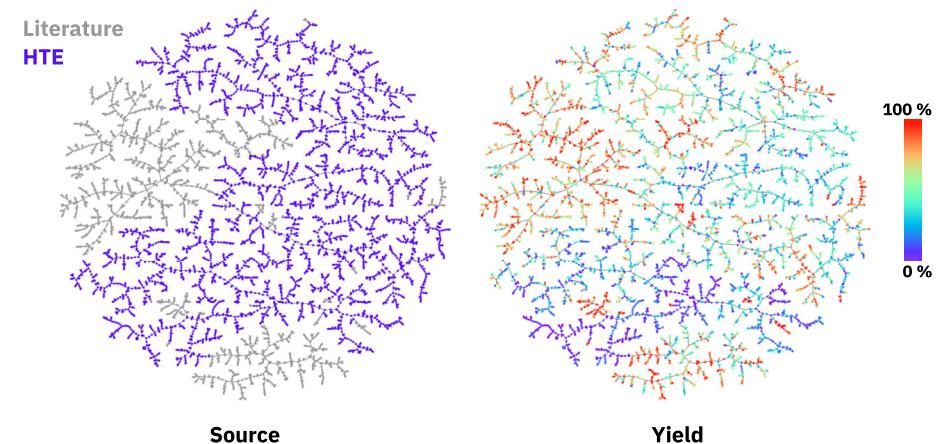


FIGURE 7 Literature and HTE Buchwald–Hartwig reactions with the corresponding yields. Every point in the TMAP^{57,205} is a reaction

independent of the size of the involved molecules. The resulting features can be formatted as a vector of predefined dimensions, commonly called reaction fingerprints. Most approaches for reaction classification differ mainly in how the reaction fingerprint is constructed and in the machine learning architecture processing the fingerprint.

A common choice to produce reaction fingerprints is to characterize chemical reactions by difference vectors of chemical descriptors between the product and reactants, as described by Schneider et al.⁵⁴ In the same study, Schneider et al.⁵⁴ developed a classifier that, given the reaction fingerprint as input, could distinguish 50 name reaction classes assigned by the NameRXN tool.⁵³ NameRXN categorizes reactions into more than 1800 classes if the reactions match expert templates. In contrast to the three-level class hierarchy in NameRXN, Ghiandoni et al.²⁰³ introduced the four-level Sheffield hierarchical reaction classification system (SHREC). However, NameRXN, as well as the SHREC tool, are closed-source.

In the Schneider fingerprint,⁵⁴ reagents are included by computing additional features (MW, NumAtoms, NumRings, LogP, NumRadicalElectrons, TPSA, NumHeteroAtoms, NumHAcceptors, and NumHDonors), which are weighted and concatenated with the products-reactants difference vectors. In contrast, Probst et al.²⁰⁴ describe a differential reaction fingerprint (DRFP), for which no distinction is made between reactants and reagents and the symmetric difference between precursors and products is taken.

Recently, Schwaller et al.⁵⁷ developed Transformer-based reaction classification models. In this case, the classification model is trained directly from the reaction SMILES representation. Apart from remarkable classification performance, the output of the reaction encoder is a vector that corresponds to a purely data-driven reaction fingerprint, which the authors called rxnfp. This fingerprint can be combined with the tree-map (TMAP) dimensionality reduction algorithm²⁰⁵ to build interactive reaction atlases that enable chemists to efficiently inspect the local and global features of their datasets.

To date, all data-driven reaction classification models^{54,57,203} have been trained on class labels generated by rule-based tools like NameRXN⁵³ or extracted from the atom-mapping predicted by RXNMapper.³⁶

The utility of mapping reactions starting from reaction fingerprints is exemplified in Figure 7. There, TMAP plots of HTE and literature reactions illustrate how these reactions cover different parts of chemical space and how the yields reported in the literature are biased toward higher values—suggesting that it is challenging to combine literature and HTE reactions for the training of machine intelligence models.

4 | CONCLUSION

Machine intelligence emerged as an important asset to explore chemical reaction spaces. For a long time, chemists have recognized the benefits of classifying large reaction databases into related chemical reactions and searching for transformations similar to reactions under study. The development of expert systems first and data-driven models afterward demonstrated the predictive power of machine intelligence beyond reaction classification, addressing tasks like forward reaction prediction, retrosynthesis, reaction optimization, catalysts design, inference of experimental procedures, to name a few. The field is evolving rapidly, and possibly some of the approaches described in this review will be

superseded in the next months or years. However, few challenges still remain of greatest interest: the possibility to learn from negative data instead of using only positive chemical reaction records; the characterization of underexplored regions in connection with low data regimes; improved predictions in the extrapolation regime and finally, the use of the chemical reaction knowledge to design entirely novel chemical transformation. While the scientific efforts will focus on addressing some of these challenges, the entire chemistry community will be responsible for driving adoption with different initiatives: from a modernization of the education courses to the outreach of the chemical community to stimulate chemical reaction data sharing on public and open platforms. Only via a concerted global effort, chemistry will be transformed into a modern data-driven science.

ACKNOWLEDGMENT

This publication was created as part of NCCR Catalysis (grant number 180544), a National Centre of Competence in Research funded by the Swiss National Science Foundation.

CONFLICT OF INTEREST

The authors have no conflict of interest.

AUTHOR CONTRIBUTIONS

Philippe Schwaller: Conceptualization (lead); project administration (lead); writing – original draft (lead); writing – review and editing (lead). **Alain C. Vaucher:** Conceptualization (equal); writing – original draft (equal); writing – review and editing (equal). **Ruben Laplaza:** Writing – original draft (equal); writing – review and editing (equal). **Charlotte Bunne:** Writing – original draft (equal); writing – review and editing (equal). **Andreas Krause:** Writing – original draft (equal); writing – review and editing (equal). **Clemence Corminboeuf:** Writing – original draft (equal); writing – review and editing (equal). **Teodoro Laino:** Conceptualization (equal); writing – original draft (equal); writing – review and editing (equal).

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Philippe Schwaller  <https://orcid.org/0000-0003-3046-6576>

Alain C. Vaucher  <https://orcid.org/0000-0001-7554-0288>

Ruben Laplaza  <https://orcid.org/0000-0001-6315-4398>

Charlotte Bunne  <https://orcid.org/0000-0003-1431-103X>

Andreas Krause  <https://orcid.org/0000-0001-7260-9673>

Clemence Corminboeuf  <https://orcid.org/0000-0001-7993-2879>

Teodoro Laino  <https://orcid.org/0000-0001-8717-0456>

RELATED WIREs ARTICLES

[Automatic reaction mapping and reaction center detection](#)

[Selectivity in organocatalysis-From qualitative to quantitative predictive models](#)

[Computer-aided synthesis design: 40 years on](#)

REFERENCES

1. Turing AM. I.—Computing machinery and intelligence. *Mind*. 1950;LIX:433–60.
2. Corey EJ. General methods for the construction of complex molecules. *Pure Appl Chem*. 1967;14:19–38.
3. Corey EJ, Long AK, Rubenstein SD. Computer-assisted analysis in organic synthesis. *Science*. 1985;228:408–18.
4. Corey EJ, Todd Wipke W, Cramer RD III, Howe WJ. Computer-assisted synthetic analysis. Facile man-machine communication of chemical structure by interactive computer graphics. *J Am Chem Soc*. 1972;94:421–30.
5. Todd Wipke W, Dyott TM. Simulation and evaluation of chemical synthesis. Computer representation and manipulation of stereochemistry. *J Am Chem Soc*. 1974;96:4825–34.
6. Wipke WT, Braun H, Smith G, Choplin F, Sieber W. SECS—simulation and evaluation of chemical synthesis: strategy and planning. In *Computer-Assisted Organic Synthesis*, Chapter 5, pp 97–127, ACS Symposium Series, Vol. 61. ACS Publications; 1977. <https://doi.org/10.1021/bk-1977-0061.ch005>.

7. Gelernter HL, Sanders AF, Larsen DL, Agarwal KK, Boivie RH, Spritzer GA, et al. Empirical explorations of synchem. *Science*. 1977; 197:1041–9.
8. Herbert Gelernter J, Rose R, Chen C. Building and refining a knowledge base for synthetic organic chemistry via the methodology of inductive and deductive machine learning. *J Chem Inf Comput Sci*. 1990;30:492–504.
9. Hendrickson JB. Systematic characterization of structures and reactions for use in organic synthesis. *J Am Chem Soc*. 1971;93:6847–54.
10. Dugundji J, Ugi I. An algebraic model of constitutional chemistry as a basis for chemical computer programs. *Computers in chemistry*. Berlin, Germany: Springer; 1973. p. 19–64.
11. Gasteiger J, Jochum C. Eros a computer program for generating sequences of reactions. *Organic compounds*. Berlin, Germany: Springer; 1978. p. 93–126.
12. Gasteiger, J, Hutchings, MG, Christoph, B, Gann, L, Hiller, C, Löw, P, et al. A new treatment of chemical reactivity: development of eros, an expert system for reaction prediction and synthesis design. *Organic Synthesis, Reactions and Mechanisms*; 1987. p. 19–73.
13. Jorgensen WL, Laird ER, Gushurst AJ, Fleischer JM, Gothe SA, Nelson HE, et al. Cameo: a program for the logical prediction of the products of organic reactions. *Pure Appl Chem*. 1990;62:1921–32.
14. Gasteiger J, Ihlenfeldt WD, Röse P. A collection of computer methods for synthesis design and reaction prediction. *Recl Trav Chim Pays-Bas*. 1992;111:270–90.
15. Satoh H, Funatsu K. Sophia, a knowledge base-guided reaction prediction system-utilization of a knowledge base derived from a reaction database. *J Chem Inf Comput Sci*. 1995;35:34–44.
16. Szymkuć S, Gajewska EP, Klucznik T, Molga K, Dittwald P, Startek M, et al. Computer-assisted synthetic planning: the end of the beginning. *Angew Chem Int Ed*. 2016;55:5904–37.
17. Klucznik T, Mikulak-Klucznik B, McCormack MP, Lima H, Szymkuć S, Bhowmick M, et al. Efficient syntheses of diverse, medicinally relevant targets planned by computer and executed in the laboratory. *Chem*. 2018;4:522–32.
18. Molga K, Szymkuć S, Grzybowski BA. Chemist ex machina: advanced synthesis planning by computers. *Acc Chem Res*. 2021;54: 1094–106.
19. D.M. Lowe. Extraction of chemical structures and reactions from the literature. [PhD thesis], University of Cambridge; 2012.
20. Daniel Lowe. Chemical reactions from US patents (1976–Sep2016). 6 2017.
21. Coley CW, Barzilay R, Jaakkola TS, Green WH, Jensen KF. Prediction of organic reaction outcomes using machine learning. *ACS Cent Sci*. 2017;3:434–43.
22. Jin W, Coley C, Barzilay R, Jaakkola T. Predicting organic reaction outcomes with weisfeiler-lehman network. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, et al., editors. *Advances in neural information processing systems*. Redhook, USA: Curran Associates, Inc; 2017. p. 2607–16.
23. Segler MHS, Preuss M, Waller MP. Planning chemical syntheses with deep neural networks and symbolic ai. *Nature*. 2018;555:604–10.
24. Schwaller P, Gaudin T, Lanyi D, Bekas C, Laino T. “Found in translation”: predicting outcomes of complex organic chemistry reactions using neural sequence-to-sequence models. *Chem Sci*. 2018;9:6091–8.
25. Bradshaw, J, Kusner, MJ, Paige, B, Segler, MHS, Hernández-Lobato, JM. A generative model for electron paths. In 7th International Conference On Learning Representations, ICLR 2019; 2019.
26. Coley CW, Jin W, Rogers L, Jamison TF, Jaakkola TS, Green WH, et al. A graph-convolutional neural network model for the prediction of chemical reactivity. *Chem Sci*. 2019;10:370–7.
27. Schwaller P, Laino T, Gaudin T, Bolgar P, Hunter CA, Bekas C, et al. Molecular transformer: A model for uncertainty-calibrated chemical reaction prediction. *ACS Cent Sci*. 2019;5:1572–83.
28. Schwaller P, Petraglia R, Zullo V, Nair VH, Haeuselmann RA, Pisoni R, et al. Predicting retrosynthetic pathways using transformer-based models and a hyper-graph exploration strategy. *Chem Sci*. 2020;11:3316–25.
29. Thakkar A, Kogej T, Reymond J-L, Engkvist O, Bjerrum EJ. Datasets and their influence on the development of computer assisted synthesis planning tools in the pharmaceutical domain. *Chem Sci*. 2020;11:154–68.
30. Sacha M, Blaz M, Byrski P, Dabrowski-Tumanski P, Chrominski M, Loska R, et al. Molecule edit graph attention network: modeling chemical reactions as sequences of graph edits. *J Chem Inf Model*. 2021;61:3273–84.
31. Pesciullesi G, Schwaller P, Laino T, Reymond J-L. Transfer learning enables the molecular transformer to predict regio-and stereoselective reactions on carbohydrates. *Nat Commun*. 2020;11:1–8.
32. Mikulak-Klucznik B, Golebiowska P, Bayly AA, Popik O, Klucznik T, Szymkuc S, et al. Computational planning of the synthesis of complex natural products. *Nature*. 2020;588:83–8.
33. Shields BJ, Stevens J, Li J, Parasram M, Damani F, Martinez Alvarado JI, et al. Bayesian reaction optimization as a tool for chemical synthesis. *Nature*. 2021;590:89–96.
34. Grisoni F, Huisman BJH, Button AL, Moret M, Atz K, Merk D, et al. Combining generative artificial intelligence and on-chip synthesis for de novo drug design. *Sci Adv*. 2021;7:eabg3338.
35. Chen WL, Chen DZ, Taylor KT. Automatic reaction mapping and reaction center detection. *WIREs Comput Mol Sci*. 2013;3:560–93.
36. Schwaller P, Hoover B, Reymond J-L, Strobel H, Laino T. Extraction of organic chemistry grammar from unsupervised learning of chemical reactions. *Sci Adv*. 2021;7:eaabe4166.
37. Lin A, Dyubankova N, Madzhidov TI, Nugmanov RI, Verhoeven J, Gimadiev TR, et al. Atom-to-atom mapping: a benchmarking study of popular mapping algorithms and consensus strategies. *Molecular Informatics*. 2020;2100138.
38. Vollmer JJ. Wiswesser line notation: an introduction. *J Chem Educ*. 1983;60(3):192.

39. Weininger D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J Chem Inf Comput Sci.* 1988;28:31–6.
40. Weininger D, Weininger A, Weininger JL. SMILES. 2. Algorithm for generation of unique SMILES notation. *J Chem Inf Comput Sci.* 1989;29:97–101.
41. Ash S, Cline MA, Webster Homer R, Hurst T, Smith GB. Sybyl line notation (sln): A versatile language for chemical structure representation. *J Chem Inf Comput Sci.* 1997;37(1):71–9.
42. Dalby A, Nourse JG, Douglas Hounshell W, Gushurst AKI, Grier DL, Leland BA, et al. Description of several chemical structure file formats used by computer programs developed at molecular design limited. *J Chem Inf Comput Sci.* 1992;32(3):244–55.
43. Daylight Theory: SMILES. <https://www.daylight.com/dayhtml/doc/theory/theory.smiles.html>. Accessed 15 Nov 2021.
44. Segler MHS, Kogej T, Tyrchan C, Waller MP. Generating focused molecule libraries for drug discovery with recurrent neural networks. *ACS Cent Sci.* 2018;4:120–31.
45. Fujita S. Description of organic reactions based on imaginary transition structures. 1. Introduction of new concepts. *J Chem Inf Comput Sci.* 1986;26:205–12.
46. Hoonakker F, Lachiche N, Varnek A, Wagner A. Condensed graph of reaction: considering a chemical reaction as one single pseudo molecule. *Int J Artif Intell Tools.* 2011;20:253–70.
47. Daylight Theory: SMARTS. <https://www.daylight.com/dayhtml/doc/theory/theory.smarts.html>. Accessed 15 Nov 2021
48. Daylight Theory: SMIRKS. <https://www.daylight.com/dayhtml/doc/theory/theory.smirks.html>. Accessed 15 Nov 2021
49. Judson PN, Ihlenfeldt W-D, Patel H, Delannée V, Tarasova N, Nicklaus MC. Adapting chmtrn (chemistry translator) for a new use. *J Chem Inf Model.* 2020;60:3336–41.
50. Delannée V, Nicklaus MC. ReactionCode: format for reaction searching, analysis, classification, transform, and encoding/decoding. *J Chem.* 2020;12:1–13.
51. Krenn M, Häse F, Nigam AK, Friederich P, Aspuru-Guzik A. Self-referencing embedded strings (SELFIES): A 100% robust molecular string representation. *Mach Learn: Sci Technol.* 2020;1:045024.
52. Dai H, Li C, Coley C, Dai B, Song L. Retrosynthesis prediction with conditional graph logic network. *Adv Neur In.* 2019;32:8872–82.
53. Nextmove Software NameRXN. <http://www.nextmovesoftware.com/namerxn.html>. Accessed 29 Jul 2021
54. Schneider N, Lowe DM, Sayle RA, Landrum GA. Development of a novel fingerprint for chemical reactions and its application to large-scale reaction classification and similarity. *J Chem Inf Model.* 2015;55:39–53.
55. Schneider N, Stiefl N, Landrum GA. What's what: the (nearly) definitive guide to reaction role assignment. *J Chem Inf Model.* 2016;56: 2336–46.
56. Liu B, Ramsundar B, Kawthekar P, Shi J, Gomes J, Nguyen QL, et al. Retrosynthetic reaction prediction using neural sequence-to-sequence models. *ACS Cent Sci.* 2017;3:1103–13.
57. Schwaller P, Probst D, Vaucher AC, Nair VH, Kreutter D, Laino T, et al. Mapping the space of chemical reactions using attention-based neural networks. *Nat Mach Intell.* 2021;3:144–52.
58. Nextmove Software Pistachio. <http://www.nextmovesoftware.com/pistachio.html>. Accessed 29 Jul 2021.
59. Reaxys database. <https://www.reaxys.com>. Accessed 29 Jul 2021.
60. CAS SciFinder database. <https://scifinder.cas.org>. Accessed 29 Jul 2021.
61. Thieme Science of Synthesis database. <https://www.thieme.de/en/thieme-chemistry/about-science-of-synthesis-54781.htm>. Accessed 29 Jul 2021.
62. Jiang S, Zhang Z, Zhao H, Li J, Yang Y, Lu B-L, et al. When SMILES smiles, practicality judgment and yield prediction of chemical reaction via deep chemical language processing. *IEEE Access.* 2021;9:85071–83.
63. Fitzner M, Wuenschik G, Koller RJ, Adam J-M, Schindler T, Reymond J-L. What can reaction databases teach us about Buchwald–Hartwig cross-couplings. *Chem Sci.* 2020;11:13085–93.
64. Schwaller P, Vaucher AC, Laino T, Reymond J-L. Prediction of chemical reaction yields using deep learning. *Mach Learn: Sci Technol.* 2021;2:015016.
65. Toniato A, Schwaller P, Cardinale A, Geluykens J, Laino T. Unassisted noise reduction of chemical reaction datasets. *Nat Mach Intell.* 2021;3:485–94.
66. Ahneman DT, Estrada JG, Lin S, Dreher SD, Doyle AG. Predicting reaction performance in C–N cross-coupling using machine learning. *Science.* 2018;360:186–90.
67. Nielsen MK, Ahneman DT, Riera O, Doyle AG. Deoxyfluorination with sulfonyl fluorides: navigating reaction space with machine learning. *J Am Chem Soc.* 2018;140:5004–8.
68. Perera D, Tucker JW, Brahmbhatt S, Helal CJ, Chong A, Farrell W, et al. A platform for automated nanomole-scale reaction screening and micromole-scale synthesis in flow. *Science.* 2018;359:429–34.
69. Kearnes SM, Maser MR, Wlekliński M, Kast A, Doyle AG, Dreher SD, et al. The open reaction database. *J Am Chem Soc.* 2021;143:18820–6.
70. Ihlenfeldt W-D, Gasteiger J. Computer-assisted planning of organic syntheses: the second generation of programs. *Angew Chem Int Ed.* 1996;34:2613–33.
71. Todd MH. Computer-aided organic synthesis. *Chem Soc Rev.* 2005;34:247–66.
72. Warr WA. A short review of chemical reaction database systems, computer-aided synthesis design, reaction prediction and synthetic feasibility. *Mol Inform.* 2014;33:469–76.

73. Engkvist O, Norrby P-O, Selmi N, Lam Y-H, Peng Z, Sherer EC, et al. Computational prediction of chemical reactions: current status and outlook. *Drug Discov Today*. 2018;23:1203–18.
74. Kayala MA, Azencott C-A, Chen JH, Baldi P. Learning to predict chemical reactions. *J Chem Inf Model*. 2011;51:2209–22.
75. Kayala MA, Baldi P. ReactionPredictor: prediction of complex chemical reactions at the mechanistic level using machine learning. *J Chem Inf Model*. 2012;52:2526–40.
76. Fooshee D, Mood A, Gutman E, Tavakoli M, Urban G, Liu F, et al. Deep learning for chemical reaction prediction. *Mol Syst Des Eng*. 2018;3:442–52.
77. Wei JN, Duvenaud D, Aspuru-Guzik A. Neural networks for the prediction of organic chemistry reactions. *ACS Cent Sci*. 2016;2: 725–32.
78. Segler MHS, Waller MP. Neural-symbolic machine learning for retrosynthesis and reaction prediction. *Chem A Eur J*. 2017;23:5966–71.
79. Rogers D, Hahn M. Extended-connectivity fingerprints. *J Chem Inf Model*. 2010;50:742–54.
80. Do, K, Tran, T, Venkatesh, S. Graph transformation policy network for chemical reaction prediction. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining; 2019. p. 750–760.
81. Nikitin F, Isayev O, Strijov V. DRACON: disconnected graph neural network for atom mapping in chemical reactions. *Phys Chem Chem Phys*. 2020;22:26478–86.
82. Nam, J, Kim, J. Linking the neural machine translation and the prediction of organic chemistry reactions. arXiv: 1612.09529; 2016. Accessed 29 Aug 2019.
83. Tetko IV, Karpov P, Van Deursen R, Godin G. State-of-the-art augmented NLP transformer models for direct and single-step retrosynthesis. *Nat Commun*. 2020;11:1–11.
84. Qian, WW, Russell, NT, Simons, CLW, Luo, YW, Burke, MD, Peng, J. Integrating deep neural networks and symbolic inference for organic reactivity prediction. *Theor Comp Chem* 2020. <https://doi.org/10.26434/chemrxiv.11659563.v1>
85. Mann V, Venkatasubramanian V. Predicting chemical reaction outcomes: a grammar ontology-based transformer framework. *AIChE J*. 2021;67:e17190.
86. Bi, H, Wang, H, Shi, C, Coley, C, Tang, J, Guo, H. Non-autoregressive electron redistribution modeling for reaction prediction. arXiv preprint arXiv:2106.07801; 2021.
87. Kovács DP, McCorkindale W, Lee AA. Quantitative interpretation explains machine learning models for chemical reaction prediction and uncovers bias. *Nat Commun*. 2021;12:1–9.
88. Kreutter D, Schwaller P, Reymond J-L. Predicting enzymatic reactions with a molecular transformer. *Chem Sci*. 2021;12:8648–59.
89. Wang L, Zhang C, Bai R, Li J, Duan H. Heck reaction prediction using a transformer model based on a transfer learning strategy. *Chem Commun*. 2020;56:9368–71.
90. Yejian W, Zhang C, Wang L, Duan H. A graph-convolutional neural network for addressing small-scale reaction prediction. *Chem Commun*. 2021;57:4114–7.
91. Seidl, P, Renz, P, Dyubankova, N, Neves, P, Verhoeven, J, Wegner, JK, Hochreiter, S, Klambauer, G. Modern hopfield networks for few-and zero-shot reaction prediction. arXiv preprint arXiv:2104.03279; 2021.
92. Probst D, Manica M, Teukam YGN, Castrogiovanni A, Paratore F, Laino T. Biocatalysed synthesis planning using data-driven learning. *Nat Commun*. 2021;13(1):1–11.
93. Coley CW, Rogers L, Green WH, Jensen KF. Computer-assisted retrosynthesis based on molecular similarity. *ACS Cent Sci*. 2017;3: 1237–45.
94. Sutskever I, Vinyals O, Le QV. Sequence to sequence learning with neural networks. Advances in neural information processing systems. Redhook, USA: Curran Associates, Inc; 2014. p. 3104–12.
95. Yang Q, Sresht V, Bolgar P, Hou X, Klug-McLeod JL, Butler CR, et al. Molecular transformer unifies reaction prediction and retrosynthesis across pharma chemical space. *Chem Commun*. 2019;55:12152–5.
96. Karpov, P, Godin, G, Tetko, IV. A transformer model for retrosynthesis. In International Conference on Artificial Neural Networks; Springer; 2019. p. 817–830.
97. Duan H, Wang L, Zhang C, Guo L, Li J. Retrosynthesis with attention-based nmt model and chemical analysis of “wrong” predictions. *RSC Adv*. 2020;10:1371–8.
98. Zheng S, Rao J, Zhang Z, Jun X, Yang Y. Predicting retrosynthetic reactions using self-corrected transformer neural networks. *J Chem Inf Model*. 2019;60:47–55.
99. Kim E, Lee D, Kwon Y, Park MS, Choi Y-S. Valid, plausible, and diverse retrosynthesis using tied two-way transformers with latent variables. *J Chem Inf Model*. 2021;61:123–33.
100. Ucak UV, Kang T, Ko J, Lee J. Substructure-based neural machine translation for retrosynthetic prediction. *J Chem*. 2021;13:1–15.
101. Chen S, Jung Y. Deep retrosynthetic reaction prediction using local reactivity and global attention. *JACS Au*. 2021;1:1612–20.
102. Shi, C, Minkai, X, Guo, H, Zhang, M, Tang, J. A graph to graphs framework for retrosynthesis prediction. In International Conference on Machine Learning, PMLR; 2020. p. 8818–8827.
103. Somnath, VR, Bunne, C, Coley, CW, Krause, A, Barzilay, R. Learning graph models for Retrosynthesis predictions. In 35rd Conference on Neural Information Processing Systems (NeurIPS); 2021.
104. Seo, S-W, Song, YY, Yang, JY, Bae, S, Lee, H, Shin, J, Hwang, SJ, Yang, E. GTA: graph truncated attention for retrosynthesis. Proceedings of the AAAI Conference on Artificial Intelligence; 2021. Vol. 35, p. 531–539.

105. Wang X, Li Y, Qiu J, Chen G, Liu H, Liao B, et al. Retroprime: A diverse, plausible and transformer-based method for single-step retrosynthesis predictions. *Chem Eng J.* 2021;420:129845.
106. Thakkar A, Selmi N, Reymond J-L, Engkvist O, Bjerrum EJ. “Ring breaker”: neural network driven synthesis prediction of the ring system chemical space. *J Med Chem.* 2020;63:8791–808.
107. Fortunato ME, Coley CW, Barnes BC, Jensen KF. Data augmentation and pretraining for template-based retrosynthetic prediction in computer-aided synthesis planning. *J Chem Inf Model.* 2020;60:3398–407.
108. Beker W, Gajewska EP, Badowski T, Grzybowski BA. Prediction of major regio-, site-, and diastereoisomers in diels–alder reactions by using machine-learning: the importance of physically meaningful descriptors. *Angew Chem Int Ed.* 2019;58:4515–9.
109. Coley CW, Thomas DA, Lummiss JAM, Jaworski JN, Breen CP, Schultz V, et al. A robotic platform for flow synthesis of organic compounds informed by AI planning. *Science.* 2019;365(6453).
110. Genheden S, Thakkar A, Chadimová V, Reymond J-L, Engkvist O, Bjerrum E. AiZynthFinder: a fast, robust and flexible open-source software for retrosynthetic planning. *J Chem.* 2020;12:1–9.
111. Watson IA, Wang J, Nicolaou CA. A retrosynthetic analysis algorithm implementation. *J Chem.* 2019;11:1–12.
112. Lin K, Youjun X, Pei J, Lai L. Automatic retrosynthetic route planning using template-free models. *Chem Sci.* 2020;11:3355–64.
113. IBM RXN for Chemistry platform. <https://rxn.res.ibm.com>. Accessed 15 Nov 2021.
114. Schreck JS, Coley CW, Bishop KJM. Learning retrosynthetic planning through simulated experience. *ACS Cent Sci.* 2019;5:970–81.
115. Chen, B, Li, C, Dai, H, and Song, L. Retro*: learning retrosynthetic planning with neural guided A* search. International Conference on Machine Learning. PMLR, 2020. p. 1608–1616.
116. Kim, J, Ahn, S, Lee, H, Shin, J. Self-improved retrosynthetic planning. arXiv preprint arXiv:2106.04880; 2021.
117. Guo Z, Stephen W, Ohno M, Yoshida R. Bayesian algorithm for retrosynthesis. *J Chem Inf Model.* 2020;60:4474–86.
118. Yan, C, Ding, Q, Zhao, P, Zheng, S, Yang, J, Yu, Y, Huang, J. Retroxpert: decompose retrosynthesis prediction like a chemist. arXiv preprint arXiv:2011.02893, 2020.
119. Wang X, Qian Y, Gao H, Coley CW, Mo Y, Barzilay R, et al. Towards efficient discovery of green synthetic pathways with Monte Carlo tree search and reinforcement learning. *Chem Sci.* 2020;11:10959–72.
120. Shibukawa R, Ishida S, Yoshizoe K, Wasa K, Takasu K, Okuno Y, et al. CompRet: a comprehensive recommendation framework for chemical synthesis planning with algorithmic enumeration. *J Chem.* 2020;12:1–14.
121. Heifets, A, Jurisica, I. Construction of new medicines via game proof search. Proceedings of the AAAI Conference on Artificial Intelligence. Vol 26; 2012.
122. Gao H, Coley CW, Struble TJ, Li L, Qian Y, Green WH, et al. Combining retrosynthesis and mixed-integer optimization for minimizing the chemical inventory needed to realize a who essential medicines list. *React Chem Eng.* 2020;5:367–76.
123. Gao H, Pauphilet J, Struble TJ, Coley CW, Jensen KF. Direct optimization across computer-generated reaction networks balances materials use and feasibility of synthesis plans for molecule libraries. *J Chem Inf Model.* 2020;61:493–504.
124. Weber JM, Lió P, Lapkin AA. Identification of strategic molecules for future circular supply chains using large reaction networks. *React Chem Eng.* 2019;4(11):1969–81.
125. Mo Y, Guan Y, Verma P, Guo J, Fortunato ME, Lu Z, et al. Evaluating and clustering retrosynthesis pathways with learned strategy. *Chem Sci.* 2021;12:1469–78.
126. Gao H, Struble TJ, Coley CW, Wang Y, Green WH, Jensen KF. Using machine learning to predict suitable conditions for organic reactions. *ACS Cent Sci.* 2018;4:1465–76.
127. Walker E, Kammeraad J, Goetz J, Robo MT, Tewari A, Zimmerman PM. Learning to predict reaction conditions: relationships between solvent, molecular structure, and catalyst. *J Chem Inf Model.* 2019;59:3645–54.
128. Maser MR, Cui AY, Ryou S, DeLano TJ, Yue Y, Reisman SE. Multilabel classification models for the prediction of cross-coupling reaction conditions. *J Chem Inf Model.* 2021;61:156–66.
129. Vaucher AC, Schwaller P, Geluykens J, Nair VH, Iuliano A, Laino T. Inferring experimental procedures from text-based representations of chemical reactions. *Nat Commun.* 2021;12:2573.
130. Vaucher AC, Zipoli F, Geluykens J, Nair VH, Schwaller P, Laino T. Automated extraction of chemical synthesis actions from experimental procedures. *Nat Commun.* 2020;11:3601.
131. Chuang KV, Keiser MJ. Comment on “Predicting reaction performance in C–N cross-coupling using machine learning”. *Science.* 2018;362(6416).
132. Estrada JG, Ahneman DT, Sheridan RP, Dreher SD, Doyle AG. Response to Comment on “Predicting reaction performance in C–N cross-coupling using machine learning”. *Science.* 2018;362(6416).
133. Sandfort F, Strieth-Kalthoff F, Kühnemund M, Beecks C, Glorius F. A structure-based platform for predicting chemical reactivity. *Chem.* 2020;6:1379–90.
134. Schwaller P, Vaucher AC, Laino T, Reymond J-L. Data augmentation strategies to improve reaction yield predictions and estimate uncertainty. *Theor Comp Chem.* 2020. <https://doi.org/10.26434/chemrxiv.13286741.v1>
135. Haywood, A, Redshaw, J, Hanson-Heine, M, Taylor, A, Brown, A, Mason, A, Gaertner, T & Hirst, J. Kernel methods for predicting yields of chemical reactions. *J Chem Inf Model.* 2021. <https://doi.org/10.1021/acs.jcim.1c00699>
136. Eyke NS, Green WH, Jensen KF. Iterative experimental design based on active machine learning reduces the experimental burden associated with reaction screening. *Reaction Chemistry & Engineering.* 2020;5:1963–72.
137. Santanilla AB, Regalado EL, Pereira T, Shevlin M, Bateman K, Campeau L-C, et al. Nanomole-scale high-throughput chemistry for the synthesis of complex molecules. *Science.* 2014;347:49–53.

138. Granda JM, Donina L, Dragone V, Long D-L, Cronin L. Controlling an organic synthesis robot with machine learning to search for new reactivity. *Nature*. 2018;559:377–81.
139. Smith A, Keane A, Dumesic JA, Huber GW, Zavala VM. A machine learning framework for the analysis and prediction of catalytic activity from experimental data. *Appl Catal B*. 2020;263:118257.
140. Wodrich MD, Sawatlon B, Busch M, Corminboeuf C. The genesis of molecular volcano plots. *Acc Chem Res*. 2021;54:1107–17. PMID: 33570407.
141. Cordova M, Wodrich MD, Meyer B, Sawatlon B, Corminboeuf C. Data-driven advancement of homogeneous nickel catalyst activity for aryl ether cleavage. *ACS Catal*. 2020;10:7021–31.
142. Meyer B, Sawatlon B, Heinen S, von Lilienfeld OA, Corminboeuf C. Machine learning meets volcano plots: computational discovery of cross-coupling catalysts. *Chem Sci*. 2018;9:7069–77.
143. Zhong M, Tran K, Min Y, Wang C, Wang Z, Dinh C-T, et al. Accelerated discovery of CO₂ electrocatalysts using active machine learning. *Nature*. 2020;581:178–83.
144. Yang Z, Gao W, Jiang Q. A machine learning scheme for the catalytic activity of alloys with intrinsic descriptors. *J Mater Chem A*. 2020;8:17507–15.
145. Ma S, Liu Z-P. Machine learning for atomic simulation and activity prediction in heterogeneous catalysis: current status and future. *ACS Catal*. 2020;10:13213–26.
146. Andersen M, Reuter K. Adsorption enthalpies for catalysis modeling through machine-learned descriptors. *Acc Chem Res*. 2021;54: 2741–9.
147. Ree N, Goller AH, Jensen JH. RegioSQM20: improved prediction of the regioselectivity of electrophilic aromatic substitutions. *J Chem*. 2021;13:1–9.
148. Struble TJ, Coley CW, Jensen KF. Multitask prediction of site selectivity in aromatic C–H functionalization reactions. *React Chem Eng*. 2020;5:896–902.
149. Guan Y, Coley CW, Wu H, Ranasinghe D, Heid E, Struble TJ, et al. Regio-selectivity prediction with a machine-learned reaction representation and on-the-fly quantum mechanical descriptors. *Chem Sci*. 2021;12:2198–208.
150. Li X, Zhang S-Q, Li-Cheng X, Hong X. Predicting regioselectivity in radical C–H functionalization of heterocycles through machine learning. *Angew Chem*. 2020;132:13355–61.
151. Banerjee S, Sreenithya A, Sunoj RB. Machine learning for predicting product distributions in catalytic regioselective reactions. *Phys Chem Chem Phys*. 2018;20:18311–8.
152. Tavakoli, M, Mood, A, Van Vranken, D, Baldi, P. Quantum mechanics and machine learning synergies: graph attention neural networks to predict chemical reactivity. *arXiv preprint arXiv:2103.14536*; 2021.
153. Moon S, Chatterjee S, Seeberger PH, Gilmore K. Predicting glycosylation stereoselectivity using machine learning. *Chem Sci*. 2021;12: 2931–9.
154. Zahrt AF, Henle JJ, Rose BT, Wang Y, Darrow WT, Denmark SE. Prediction of higher-selectivity catalysts by computer-driven workflow and machine learning. *Science*. 2019;363:eaau5631.
155. Ian Rinehart N, Zahrt AF, Henle JJ, Denmark SE. Dreams, false starts, dead ends, and redemption: a chronicle of the evolution of a chemoinformatic workflow for the optimization of enantioselective catalysts. *Acc Chem Res*. 2021;54:2041–54.
156. Huang B, von Lilienfeld OA. Quantum machine learning using atom-in-molecule-based fragments selected on the fly. *Nat Chem*. 2020; 12:945–51.
157. Gallarati S, Fabregat R, Laplaza R, Bhattacharjee S, Wodrich MD, Corminboeuf C. Reaction-based machine learning representations for predicting the enantioselectivity of organocatalysts. *Chem Sci*. 2021;12:6879–89.
158. Gao CW, Allen JW, Green WH, West RH. Reaction mechanism generator: automatic construction of chemical kinetic mechanisms. *Comput Phys Commun*. 2016;203:212–25.
159. Heinen S, von Rudorff GF, von Lilienfeld OA. Toward the design of chemical reactions: machine learning barriers of competing mechanisms in reactant space. *J Chem Phys*. 2021;155:064105.
160. von Rudorff GF, Heinen SN, Bragato M, von Lilienfeld OA. Thousands of reactants and transition states for competing E2 and SN2 reactions. *Mach Learn: Sci Technol*. 2020;1:045026.
161. Christensen AS, Bratholm LA, Faber FA, von Lilienfeld OA. FCHL revisited: faster and more accurate quantum machine learning. *J Chem Phys*. 2020;152:044107.
162. Friederich P, dos Passos Gomes G, De Bin R, Aspuru-Guzik A, Balcells D. Machine learning dihydrogen activation in the chemical space surrounding Vaska's complex. *Chem Sci*. 2020;11:4584–601.
163. Singh AR, Rohr BA, Gauthier JA, Nørskov JK. Predicting chemical reaction barriers with a machine learning model. *Catal Lett*. 2019; 149:2347–54.
164. Grambow CA, Pattanaik L, Green WH. Reactants, products, and transition states of elementary chemical reactions based on quantum chemistry. *Sci Data*. 2020;7(1):1–8.
165. Yang K, Swanson K, Jin W, Coley C, Eiden P, Gao H, et al. Analyzing learned molecular representations for property prediction. *J Chem Inf Model*. 2019;59:3370–88.
166. Grambow CA, Pattanaik L, Green WH. Deep learning of activation energies. *J Phys Chem Lett*. 2020;11:2992–7.
167. Jorner K, Brinck T, Norrby P-O, Buttar D. Machine learning meets mechanistic modelling for accurate prediction of experimental activation energies. *Chem Sci*. 2021;12:1163–75.

168. Pattanaik L, Ingraham JB, Grambow CA, Green WH. Generating transition states of isomerization reactions with deep learning. *Phys Chem Chem Phys*. 2020;22:23618–26.
169. Lemm D, von Rudorff GF, von Lilienfeld OA. Machine learning based energy-free structure predictions of molecules, transition states, and solids. *Nat Commun*. 2021;12:4468.
170. Bittracher A, Banisch R, Schütte C. Data-driven computation of molecular reaction coordinates. *J Chem Phys*. 2018;149:154103. <https://doi.org/10.1063/1.5035183>
171. Brandt S, Sittel F, Ernst M, Stock G. Machine learning of biomolecular reaction coordinates. *J Phys Chem Lett*. 2018;9:2144–50. <https://doi.org/10.1021/acs.jpclett.8b00759>
172. Rogal J. Reaction coordinates in complex systems — a perspective. *Eur Phys J B*. 2021;94(11):1–9. <https://doi.org/10.1140/epjb/s10051-021-00233-5>
173. Mockus, J. On bayesian methods for seeking the extremum. In: Guri I. Marchuk, editor. Optimization Techniques, IFIP Technical Conference, Novosibirsk, USSR, July 1–7; Springer; 1974. Lecture Notes in Computer Science. Vol. 27, p. 400–404.
174. Rasmussen CE, Williams CKI. Gaussian processes for machine learning. Adaptive computation and machine learning. Cambridge, USA: MIT Press; 2006.
175. Jones DR, Schonlau M, Welch WJ. Efficient global optimization of expensive black-box functions. *J Glob Optim*. 1998;13:455–92.
176. Srinivas N, Krause A, Kakade S, Seeger M. Information-theoretic regret bounds for gaussian process optimization in the bandit setting. *IEEE Trans Inf Theory*. 2012;58:3250–65.
177. Romero PA, Krause A, Arnold FH. Navigating the protein fitness landscape with gaussian processes. *PNAS*. 2013;110:E193–201.
178. Hase F, Roch LM, Kreisbeck C, Aspuru-Guzik A. Phoenics: a bayesian optimizer for chemistry. *ACS Cent Sci*. 2018;4:1134–45.
179. Burger B, Maffettone PM, Gusev VV, Aitchison CM, Bai Y, Wang X, et al. A mobile robotic chemist. *Nature*. 2020;583:237–41.
180. Felton KC, Rittig JG, Lapkin AA. Summit: benchmarking machine learning methods for reaction optimisation. *Chem Methods*. 2021;1: 116–22.
181. Elton DC, Boukouvalas Z, Fuge MD, Chung PW. Deep learning for molecular design—a review of the state of the art. *Mol Syst Des Eng*. 2019;4:828–49.
182. Rezende, DJ, Mohamed, S, Wierstra, D. Stochastic backpropagation and approximate inference in deep generative models. In Proceedings of the 31th International Conference on Machine Learning, ICML 2014, Beijing, China, 21–26 June 2014, Volume 32 of *JMLR* Workshop and Conference Proceedings, JMLR.org; 2014. p. 1278–1286.
183. Kingma DP, Welling M. Auto-encoding Variational Bayes. In: Bengio, Y, LeCun, Y, editors. 2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14–16, 2014, Conference Track Proceedings; 2014.
184. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. In: Ghahramani Z, Welling M, Cortes C, Lawrence N, Weinberger KQ, editors. Advances in neural information processing systems. Volume 27. Redhook, USA: Curran Associates, Inc.; 2014.
185. De Cao, N, Kipf, T. Molgan: an implicit generative model for small molecular graphs. *CoRR*, abs/1805.11973, 2018.
186. Yabo Dan Y, Zhao XL, Shaobo Li MH, Hu J. Generative adversarial networks (GAN) based efficient sampling of chemical composition space for inverse design of inorganic materials. *npj Comput Mater*. 2020;6:1–7.
187. Gómez-Bombarelli R, Wei JN, Duvenaud D, Hernández-Lobato JM, Sánchez-Lengeling B, Sheberla D, et al. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Cent Sci*. 2018;4:268–76.
188. Boitreaud J, Mallet V, Oliver C, Waldspuhl J. Optimol: optimization of binding affinities in chemical space for drug discovery. *J Chem Inf Model*. 2020;60:5658–66.
189. Gao W, Coley CW. The synthesizability of molecules proposed by generative models. *J Chem Inf Model*. 2020;60:5714–23.
190. Button A, Merk D, Hiss JA, Schneider G. Automated de novo molecular design by hybrid machine intelligence and rule-driven chemical synthesis. *Nat Mach Intell*. 2019;1:307–15.
191. Bradshaw, J, Paige, B, Kusner, MJ, Segler, MHS, Hernández-Lobato, JM. A model to search for synthesizable molecules. 33rd Conference on Neural Information Processing Systems (NeurIPS 2019); Curran Associates, Inc.; 2020. p. 7905–7917.
192. Bradshaw J, Paige B, Kusner MJ, Hernández-Lobato JM. Barking up the right tree: an approach to search over molecule synthesis dags. In: Larochelle H, Ranzato M, Hadsell R, Balcan MF, Lin H, editors. Advances in neural information processing systems. Volume 33. Redhook, USA: Curran Associates, Inc; 2020. p. 6852–66.
193. Korovina, K, Xu, S, Kandasamy, K, Neiswanger, W, Poczos, B, Schneider, J, Xing, E. ChemBO: Bayesian optimization of small organic molecules with synthesizable recommendations. International Conference on Artificial Intelligence and Statistics, PMLR; 2020. p. 3393–3403.
194. Sutton RS, Barto AG. Reinforcement learning: an introduction. Adaptive computation and machine learning. Cambridge, USA: MIT Press; 1998.
195. Puterman ML. Markov decision processes: discrete stochastic dynamic programming. Hoboken, USA: Wiley Series in Probability and Statistics. Wiley; 1994.
196. Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, et al. Mastering the game of go without human knowledge. *Nature*. 2017;550:354–9.
197. Lee J, Hwangbo J, Wellhausen L, Koltun V, Hutter M. Learning quadrupedal locomotion over challenging terrain. *Sci Rob*. 2020;5(47).
198. Simm, GNC, Pinsler, R, Hernández-Lobato, JM. Reinforcement learning for molecular design guided by quantum mechanics. Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13–18 July 2020, Virtual event, Volume 119 of Proceedings of Machine Learning Research, PMLR, 2020. p. 8959–8969.

199. Gottipati, SK, Sattarov, B, Niu, S, Pathak, Y, Wei, H, Liu, S, Liu, S, Blackburn, S, Thomas, K, Coley, C, Tang, J, Chandar, S, Bengio, Y. Learning to navigate the synthetically accessible chemical space using reinforcement learning. In: Daumé III, H, Singh, A, editors, Proceedings of the 37th International Conference on Machine Learning, Volume 119 of Proceedings of Machine Learning Research, PMLR, 13–18 July 2020; 2020. p. 3668–3679.
200. Bickerton GR, Paolini GV, Besnard J, Muresan S, Hopkins A. Quantifying the chemical beauty of drugs. *Nat Chem*. 2012;4(2):90–8.
201. You, J, Liu, B, Ying, R, Pande, V, Leskovec, J. Graph convolutional policy network for goal-directed molecular graph generation. *NeurIPS*; 2018.
202. Chen L. Reaction classification and knowledge acquisition. *Handbook of chemoinformatics*. Weinheim, Germany: Wiley-VCH Verlag GmbH; 2008. p. 348–90.
203. Ghiandoni GM, Bodkin MJ, Chen B, Hristozov D, Wallace JEA, Webster J, et al. Development and application of a data-driven reaction classification model: comparison of an electronic lab notebook and medicinal chemistry literature. *J Chem Inf Model*. 2019;59:4167–87.
204. Probst, D, Schwaller, P, Reymond, J-L. Reaction classification and yield prediction using the differential reaction fingerprint DRFP. *Chemrxiv:2021-mc870*; 2021. <https://doi.org/10.33774/chemrxiv-2021-mc870>.
205. Probst D, Reymond J-L. Visualization of very large high-dimensional data sets as minimum spanning trees. *J Chem*. 2020;12:1–13.

How to cite this article: Schwaller P, Vaucher AC, Laplaza R, Bunne C, Krause A, Corminboeuf C, et al. Machine intelligence for chemical reaction space. *WIREs Comput Mol Sci*. 2022;12:e1604. <https://doi.org/10.1002/wcms.1604>