Machine learning for retrosynthesis and synthesisable molecule generation in drug discovery

Ewa Wieczorek

Green Templeton College

University of Oxford

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Michaelmas 2024

Acknowledgements

I would like to thank all my supervisors, both academic and industrial, for all the support they have given me. My journey in drug discovery has started with working in Paul's lab, and it was great to have his input in my PhD research. Thank you to Fernanda for welcoming me to her group and letting me work on such interesting topics!

Working closely with Exscientia has allowed me to see a different side of cheminformatics research and put my work in perspective of the real-life drug discovery applications. I am very grateful to Liam for providing me with the opportunity to spend part of my PhD there and his mentorship - his ideas and input into my research have been invaluable. Thank you also to Nina, Kirubin and Mason for welcoming me to the team and showing me the ropes!

I have enjoyed the three years I spent in the Duarte group immensely - thank you to all the members for creating a great atmosphere. In particular my thanks go to the members of ML subgroup (past and present), including Chloe, Matina, Ally and Josh among others, for interesting discussions and input into my research. Also to all the members of the tiny office - Sara, Hanwen and Tom - thank you for all the chats and coffee breaks. You have made my last year at Oxford the most enjoyable!

Finally, thank you to all my friends and family who kept me sane and motivated - I would not have made it through the last 4 years without you!

Abstract

Drug discovery is a notoriously difficult and slow process, with high research and development costs and a decreasing success rate. Computer-Aided Drug Design methods show promise in improving the efficiency of early stage drug discovery, increasing the number of compounds that can be evaluated per design cycle and allowing for pre-filtering of molecules with fast computational methods before they are synthesised. However, many of the compounds designed *in silico* are not synthesisable in practice or the synthesis routes towards them are not obvious. This leads to computational resources being wasted on designing molecules that can never be tested experimentally. This thesis explores new methods for two approaches assessing and improving synthesisability in drug discovery: retrosynthesis prediction and synthesisability-constrained molecule generation.

First, the problem of retrosynthesis prediction for molecules containing heterocyclic scaffolds is considered. Four domain adaptation approaches are benchmarked to develop a single-step retrosynthesis prediction model with improved performance for ring disconnections. Accuracy for heterocycle formations and all reaction classes, as well as computational cost, are considered. A further fine-tuning workflow for continual retraining of the model with newly published data is introduced. The application of the most versatile model, trained with a mixed fine-tuning strategy, is then demonstrated in multi-step retrosynthesis in a retrospective analysis for two drug-like compounds.

Next, the development of *retro-active*, a method for synthesisable molecule generation and optimisation, is described. *Retro-active* generates molecules based on a known synthesis route and a provided starting material pool. The use of active learning for starting material selection allows for the optimisation of the resulting product molecules for user-defined scoring

functions. A benchmark of starting material acquisition and product enumeration methods is included, as well as a comparison to alternative non-machine learning-based starting material selection approaches. The applicability of *retro-active* for both ligand-based and structure-based drug discovery is demonstrated.

The use case of *retro-active* is then extended to multi-parameter optimisation, to simulate a real-life drug discovery scenario. The compounds are optimised for their structural, physicochemical, and ADMET properties, with a scoring function that combines physics-based and machine learning-based scores. The robustness of the method is demonstrated with both convergent and linear synthesis route topologies and ligands for different target proteins.

The thesis concludes with final remarks regarding retrosynthesis prediction and synthesisable molecule generation with *retro-active*, including future research directions and challenges in the field.

Contents

Chapter 1

Introduction

Pharmacological drug discovery is famously a time-consuming and expensive process, with the median research and development investment to bring a new drug to market recently estimated at $985.3 million.[1] Moreover, the productivity of drug discovery seems to be decreasing with time and the failure rate of drug discovery projects is rising.[2, 3] In recent years, the use of computational methods in early drug discovery has become increasingly popular, with the goal of conserving resources and speeding up the design process.[4] This widespread use of computational methods has been enabled by the greater availability of cloud and graphics processing unit (GPU) computing resources,[5] the developments in machine learning and deep learning methodologies,[6, 7] as well as the growth of virtual chemical spaces, such as make-on-demand libraries,[8] and the growth in the number of 3D protein structures available due to recent advances in crystallography,[9] cryo-electron microscopy[10, 11] and predictions of machine learning models .[12–14]

However, despite those developments and the increasing utility of Computer-Aided Drug Design (CADD), many of the molecules designed *in silico* are not possible to synthesise in the laboratory, leading to computational resources and time being wasted.[15] Fuelled by the belief that a greater focus needs to be placed on the syn-

thesisability of the computationally designed molecules to take full advantage of the computational methods for drug discovery, the aim of this thesis is to develop and benchmark computational methods to test and guarantee molecule synthesisability.

This chapter provides an introduction to synthesisability prediction and synthesis planning in the context of drug discovery. We begin with an overview of the early drug discovery process and the computational methods involved in it. A discussion of synthetic accessibility scores follows, together with an introduction to various reaction prediction related tasks. Finally, the background to Computer-Aided Synthesis Planning (CASP) tools is presented, with an overview of approaches used in single-step and multi-step retrosynthesis. We conclude with an outline of this thesis.

1.1 Drug Discovery

Drug discovery is the multidisciplinary process through which new medicines are discovered. It can be split into the pre-clinical phase, consisting of (1) target identification and validation, (2) hit generation, (3) hit to lead, (4) lead optimisation and (5) pre-clinical research stages, and the clinical trials phase, consisting of (6) phase 1, 2 and 3 clinical trials as well as (7) regulatory approval (Figure 1.1).[16] While in recent years therapies based on more complex molecules, such as antibodies,[17] peptides,[18] and PROTACs[19] have started to emerge, this thesis will discuss only the aspects relevant to classical small molecule drug discovery.

The drug discovery process begins with target identification, the aim of which is to find the biological component for the drug to target - this is most often a protein: enzyme, receptor, ion channel or a transcription factor, but it could also be an oligonucleotide such as DNA.[20, 21] The biological target should be one of the main drivers of the disease of interest and its mode of action and *in vivo* knockout effects are usually further studied during the target validation stage before committing to

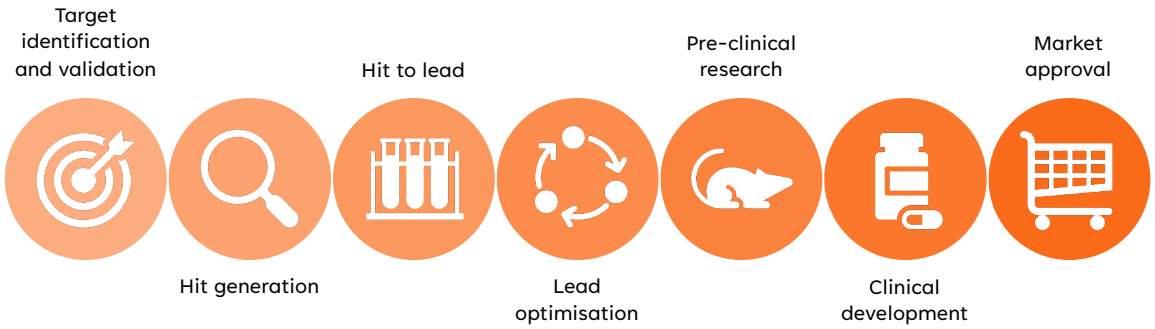


Figure 1.1: Stages in the drug discovery pipeline (left to right). Computational chemistry and cheminformatics methods are most commonly applied in the hit generation, hit to lead and lead optimisation stages.

the drug discovery campaign. Other considerations might also be taken into account at this point, such as availability of structural data or whether there is competition working on similar targets.

Once the biological target is established, the process of finding small molecule binders begins with the hit generation stage. The simplest approach is to search literature, patents, and public databases for known binders to the protein of interest or a similar target, and use them as a starting point. However, in many cases no such data is available. Under those circumstances it is most common to use high throughput screening (HTS) to identify potential hits. In HTS, thousands of compounds can be simultaneously tested for binding affinity towards the target using biological assays.[22] While very fast, this approach has a relatively high chance of producing false positive results and the hit finding rate is quite low. Computational methods can also be applied to perform HTS in what is referred to as virtual screening (VS), for example by performing protein-ligand docking.[23, 24] While faster than traditional HTS, VS requires further experimental validation and the rate of finding true hit molecules is even lower. An alternative approach to performing HTS using compound libraries is fragment-based drug discovery.[25] It involves screening a set of much smaller molecules, usually with molecular weight <300 Da, against the target

of interest. While the founds hits are typically less potent than in traditional HTS, the hit rate is significantly higher and the fragments can be grown, linked or merged to improve affinity.[26, 27]

Once a set of hit compounds with reasonable activity towards the target has been found, they are progressed to the hit to lead (or lead generation) stage.[16] The main goal now is to modify and expand the hit molecules to improve their binding affinity and selectivity towards the target, while also keeping in mind the pharmacokinetic properties. Structural modifications are applied to the hit molecules, most commonly focusing on one area of the molecule at a time, to assess the structure-activity relationship (SAR). A great focus is also placed on maintaining the so-called "drug-likeness" of the molecules, often defined by obeying the Lipinski rule of 5, describing the appropriate size, lipophilicity, and number of hydrogen bond donors (HBDs) and hydrogen bond acceptors (HBAs) for the molecule.[28] As the drug discovery campaign progresses and the compound potency improves, the most promising molecules are advanced to the lead optimisation stage. Here, further improvements to the molecule are made, often focusing on optimising multiple properties at the same time. On top of activity and selectivity, the ADMET (absorption, distribution, metabolism, excretion, and toxicity) properties of the lead candidate are considered, such as solubility, cell permeability or hERG and CYP450 inhibition.

Traditionally during hit to lead and lead optimisation stages medicinal chemists would rely on their expertise, intuition, and judgement to design compounds and decide which structural modifications would be most promising. This would usually be done in design-make-test-analyse (DMTA) cycles, where the chemists would iteratively design new compounds with the goal of optimising the property of interest, synthesise them, test them in relevant biological assays, and analyse the results to inform their subsequent design decisions (Figure 1.2).[29] With the developments in computational chemistry methods, the efficiency and speed of this process can be

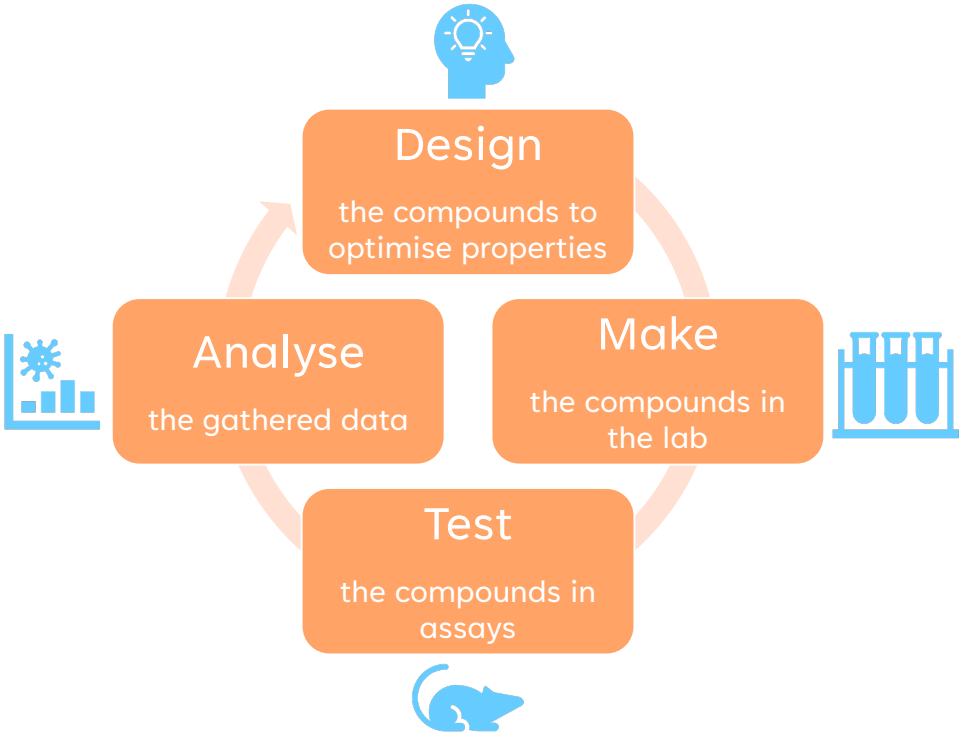


Figure 1.2: Design-make-test-analyse (DMTA) cycle. In the lead generation and optimisation stages of drug discovery, medicinal chemists go through iterative cycles of hypothesis-driven design of new compounds, their synthesis, and testing to inform the next design decisions.

improved by considering more molecules at the design stage and using *in silico* methods to pre-filter them, as well as standardising and rationalising the design decisions. While computational methods are now commonly employed at the "design" stage of the DMTA cycle (as discussed later in Section 1.2), their use to streamline the "make" stage is less prevalent.

Once a promising lead candidate is identified it is progressed first to the pre-clinical research stage, where further *in vivo* studies are performed to ensure its safety, and then to clinical trials.[30] In phase 1 clinical trials only the human safety of the drug and the potential dosage are studied on a small population of <50 subjects. The therapeutic effect is only started to be considered during phase 2 trials, where a small population of patients is assessed for the biological effect of the drug. Phase 3 trials

involve a larger population of patients and study the effectiveness of the treatment. Once the drug is confirmed to be safe and working, it can progress through regulatory approval and be brought to market.

1.2 Computer-Aided Drug Design

Computational chemistry, cheminformatics, and machine learning methods can be employed between the hit generation and lead optimisation stages of drug discovery to address the issues of high cost, low efficiency, and long timelines in drug discovery, in what is referred to as Computer-Aided Drug Design (CADD).[4, 31–34] The first uses of CADD date back to the 1970s when quantitative structure-activity relationship (QSAR) studies were first introduced.[35] Since then CADD has undergone several hype and disillusionment phases[35] but nowadays its utility in drug discovery is firmly established with most pharmaceutical companies employing it in their drug discovery campaigns. CADD can be generally split into structure-based drug design (SBDD), which utilises structural information about the target to design molecules, and ligand-based drug design (LBDD), which only requires information about known binders (ligands) towards the target. A brief overview of the most commonly used techniques in CADD is provided below.

1.2.1 Virtual screening

Virtual screening is usually performed at the hit generation stage of drug discovery and involves screening a large library of *in silico* defined molecules with a selected computational method or model.[36–38] It can be applied both in structure-based and ligand-based drug discovery, depending on the scoring function used.

When used for structure-based drug discovery, virtual screening usually involves docking the molecule library into the active (or allosteric) binding site on the target

protein.[23, 24, 39, 40] Traditional docking involves a conformational search of the ligands within the defined binding site, followed by the scoring and ranking of selected poses, usually by estimating their binding energy.[41–43] The scoring functions used can be either physics-based,[44–46] knowledge-based,[47, 48] empirical,[41, 49] or machine learning-based.[50, 51] While the binding mode prediction of most modern docking programs is very accurate, with a difference of <2Å between the true and predicted poses considered to be good,[52] the predicted binding affinity values are not as reliable. Recently, many docking methods dependent on deep learning models have started to emerge, with the hope of reducing the computational cost of docking.[53–56] In those methods, a machine learning model is trained on protein-ligand complex structures and then used to predict the binding mode, replacing the costly conformational search process. While the ligand pose prediction of these deep docking methods is faster than traditional docking, more work needs to be done to ensure the physical and chemical validity of the predicted structures before they can be reliably employed in virtual screening, as revealed by recent benchmarks such as PoseBusters.[57] Alternative approaches to improve docking efficiency and increase the size of chemical space screened include combining traditional docking with active learning to prioritise molecules for screening[58] or docking fragments or synthons from a combinatorial library into the active site and enumerating the top scoring ones.[59, 60]

In the case of ligand-based virtual screening, the compound library is compared to a known ligand or set of ligands for the target protein.[61] While having the advantage of not requiring the target protein structure to be known and a lower computational cost, ligand-based virtual screening tends to be used less frequently as similarities between ligands are more difficult to translate to binding affinity or activity towards the target. The simplest and quickest scoring functions in ligand-based virtual screening are based on 2D molecule similarity, for example by using fingerprint similarity or

employing specific substructure searches.[62] When a larger set of ligands is known, QSAR models can also be trained and used to perform the screen. 3D-based methods tend to be more informative but also more computationally expensive. Those can be purely shape-based, where conformers are generated for the compounds in the screening library, superimposed onto the known actives and the shape overlay is calculated.[62, 63] On the other hand, in pharmacophore-based screens, the 3D overlay/similarity of other electronic and chemical features of the molecule is considered in addition to the steric features, such as location of positive/negative charges, hydrophobic/hydrophilic areas, hydrogen bond donors and acceptors, or aromatic rings.[63]

The size of compound libraries used to perform virtual screens is increasing, together with improved computational methods and greater access to computing power.[4, 64, 65] Smaller, traditional virtual screens might involve testing the compounds available in-house, molecules present in literature-based databases or in-stock compounds from vendor catalogues, with around $10^6 - 10^7$ molecules being routinely screened per run. However, recent years have seen the rise in popularity of virtual on-demand databases (or combinatorial libraries), such as Enamine REAL.[66] These libraries consist of a set of building blocks and a set of common and reliable 2- or 3-component reactions that can be applied to them, with the full compound library obtained by enumerating the building blocks with compatible reactions. At around $10^{10} - 10^{15}$ compounds, it might be too computationally expensive to screen the whole on-demand database with expensive 3D-based scoring functions, so cheaper pre-filtering steps can often be employed to reduce the size of the chemical space. While the reported synthesis success rate for compounds in those virtual combinatorial libraries is quite high, at over 80%, developing better and quicker synthesisability prediction methods could aid in such pre-filtering steps to ensure all the screened compounds can be made.

1.2.2 *De novo* design

De novo design is a computational technique in drug discovery that aims to generate novel compounds with the desired pharmacological or physicochemical properties.[67–69] In contrast to virtual screening, which explores only a small fraction of the theoretical chemical space but scores it completely (with maximum $10^8 - 10^9$ molecules screened[23] out of the estimated $10^{23} - 10^{60}$ possible[70]), *de novo* design attempts to explore the full chemical space but only score it sparsely by optimising the generated molecules towards the desired properties. Although when used nowadays *de novo* design usually refers to deep learning based methods, the term dates back to the 1990s and the use of rule-based transformation algorithms.

Some of the earliest rule-based algorithms relied on atom-by-atom generation of molecules within a known binding pocket.[71, 72] While allowing for high structural variety and finer control over the molecular properties, the size of chemical space that could be accessed by those methods was impossible to fully explore. Moreover, the synthesisability of the generated molecules was low, with no consideration being given to what reactions would form the bonds between the added atoms. Fragment-based algorithms (such as fragment growing[73–75] or linking[73, 74, 76]) and reaction-based generation[77] addressed the issues of atom-based generation by limiting the combinatorial space and introducing explicit synthesisability consideration, leading to their greater popularity.

Another class of early *de novo* design algorithms, including genetic algorithms and matched molecular pair analysis (MMPA), relied on rule-based transformations applied to a known binder to create novel molecules instead of generating them completely from scratch. Genetic algorithms involve an iterative cycle of mutations or crossovers being applied to the molecule population, scoring of the newly generated compounds, and selection of the most optimal ones.[78] While they have been successfully applied in drug discovery,[79] the need to manually define the genetic oper-

ations can be a limitation and significantly affect the outcome of the process. On the other hand, matched molecular pair analysis involves mining a database of previously published compounds for molecules differing by a single transformation.[80–82] The observed trends in changes to molecule properties arising from this transformation can then be used to inform subsequent design decisions. While MMPA has the advantage of using already published data (including data for different targets) and not needing to run any time-consuming assays or calculations, the effect of the molecular transformation can differ between projects, lowering MMPA's reliability.[83]

The last 5 years have seen a huge growth in the use of deep learning for *de novo* design (deep generative modelling), starting with the SMILES-based molecular autoencoder developed by Gómez-Bombarelli *et al.*[84] Since then, both text-based (SMILES[85] and SELFIES[86]) and graph-based (2D topological graphs and 3D geometric graphs) molecular representations have been used for *de novo* molecule generation with a variety of deep learning algorithms: variational autoencoders,[84, 87, 88] recurrent neural networks,[89, 90] generative adversarial networks,[91, 92] normalizing flows,[93], autoregressive models,[94, 95] and diffusion models.[96, 97]

Deep generative modelling has traditionally been split into two task: distributional learning and goal-based design, with the latter further split into conditional generation and molecule optimisation.[68, 98] In distributional learning, the underlying chemical distribution of a set of molecules (usually known binders) is learned and new molecules are sampled from this distribution. The practical applications of distributional learning are limited to the creation of virtual screening libraries or the generation of molecules structurally similar to known actives that can then be used as a starting point for optimisation with other methods. On the other hand, goal-based generation, which aims to generate molecules with a set of desired properties, is much more common in drug discovery. The models can be conditioned towards specific pharmacological or physicochemical properties (such as binding affinity or

ADMET), to change or retain molecular substructures, or to optimise the molecule for interactions with a known target structure. Goal-based generation is routinely applied in lead optimisation, scaffold hopping and linker design.[99]

Together with the growth in the number of available deep generative modelling methods, benchmarks have started to emerge that compare their performance, beginning with the GuacaMol benchmark.[98] While the ability of the methods to optimise the generated molecules towards desired properties and the chemical validity of the generated compounds are usually assessed, synthesisability is not considered as frequently. Meanwhile, these methods can have the tendency to exploit the reward function and produce complex and improbable molecules, with a high proportion of synthetically inaccessible molecules being generated as highlighted by Gao *et al.*[15] While the use of synthesisability scores or synthesis planning tools (described later in Sections 1.3 and 1.5) has been widely adopted in *de novo* generation workflows (either at post-processing or included in the reward function) following the work of Gao *et al.*, the synthesisability of *de novo* generated molecules is still viewed as an ongoing challenge.[68]

1.2.3 Molecular property prediction

Another very common use of machine learning in drug discovery is the application of molecular property models for prediction of activity and physicochemical and pharmacokinetic properties in what is often referred to as quantitative structure-activity relationship (QSAR) or quantitative structure-property relationship (QSPR) modelling.[100–102] Those models can be utilised either in virtual screening workflows, as part of the reward function in *de novo* generation or to score human-designed molecules in lead optimisation.

The earliest applications of QSAR started with linear models that correlated a single molecular property or descriptor to the compound activity towards the target

protein.[103] Later, the use of machine learning models has been adopted in order to model the relationship between multiple molecular features and the compound activity. Random Forests (RFs) have been widely used for this tasks, due to their reliable prediction performance across various datasets and the easy interpretability of the feature contributions to the final prediction.[104–106] In recent years the use of deep learning algorithms has also become prevalent - however, while they might improve the predictive performance marginally, they suffer from interpretability issues.[107–109] While QSAR models are now widely adopted and successfully used in many drug discovery campaigns, one of their limitations is that they still struggle with prediction of "activity cliffs" - molecules which are structurally very similar but differ greatly in their binding affinity.[110, 111]

Apart from predicting binding affinity or activity, machine learning models have also been applied to prediction of many physicochemical and ADMET properties, which are often given more consideration in later stages of drug discovery such as lead optimisation. For physicochemical properties, many models have been developed to predict lipophilicity of the compounds, defined by either the octanol–water partition coefficient (log P) or pH-dependent distribution coefficient (log D), and tend to perform well for molecules within the training data distribution.[112–114] On the other hand, solubility prediction is considered to be more difficult and while models have been developed, their performance is significantly worse.[115–118] The last commonly predicted physicochemical property is cell permeability, with models developed for Madin-Darby canine kidney (MDCK) cells,[119] the parallel artificial membrane permeability assay (PAMPA)[120] and the human colon adenocarcinoma (Caco-2) cell line.[121] Out of the ADMET properties, absorption is one of the most reliably predicted, with many human intestinal absorption models.[122, 123] For distribution, models of both plasma protein binding,[124, 125] and blood-brain barrier permeation[126] have been published, with the latter being a more difficult task.

Metabolism prediction can be split into predicting the sites[127, 128] and the isoforms responsible[129] for the metabolism. While the developed models are relatively successful, the main focus remains on CYP450-mediated metabolism. Excretion is one of the more difficult and rarer ADMET properties to model due to the complex mechanisms behind it, however models to predict the volume of distribution[130] and human plasma clearance[131] have been developed. On the other hand, toxicity prediction has been extensively studied and benchmarked, for example in the Tox21 challenge.[132]

While the growth in the number of QSAR/QSPR models published in recent years might seem optimistic for their use in accelerating drug discovery, several concerns regarding their real-life application and the rigour of statistical testing remain.[133] The practical applicability of the deep learning models is often limited, when the incremental improvement in performance comes at a cost of a much longer computing time.[101] Moreover, the relevance of the datasets used to benchmark those methods, such as MoleculeNet,[134] to real-world drug discovery is questionable.[135, 136] Finally, random splitting of the datasets can lead to overestimation of the models' performance, sometimes even to above what is realistic given the experimental errors.[137]

1.3 Synthetic accessibility prediction

Determining synthetic accessibility of computationally-designed molecules is an important task in drug discovery, to be better able to select or optimise molecules proposed for experimental testing. While many tools have been created with the aim of predicting synthetic pathways towards molecules (as discussed later in Section 1.5), their prediction time is very slow due to the complex nature of the problem. Synthetic accessibility (or synthetic complexity) scores are a useful surrogate for synthesis plan-

ning tools, with a much shorter prediction time, albeit at the cost of reduced accuracy and not providing the synthesis route.

The earliest efforts in predicting synthetic complexity relied on expert-selected molecular descriptors.[138–140] In 2009 Ertl *et al.* introduced the synthetic accessibility score (SAscore), which combined a "fragment score" (based on substructure frequency in molecules from PubChem[141]) and a "complexity penalty" (that accounts for rings, macrocycles, stereochemistry and size of the molecules).[142] While still widely used to this date and implemented in RDKit,[143] SAscore tends to underestimate the accessibility of large and complex molecules containing known substructures.

Early work also attempted to explicitly capture the synthetic chemists' opinions and intuition and convert it into a synthetic accessibility score or use it to select weights for the components of synthetic complexity scores.[144, 145] However, with the task of manually assessing molecule synthesisability being relatively time-consuming, the collected dataset sizes were quite modest, with between 100-4000 molecules being scored by only 5 chemists. Together with a low agreement between the chemists when ranking the molecules based on synthetic accessibility, this highlighted the difficulties in taking a human-based approach.

With the continuing developments in machine learning and deep learning, more ML-based synthetic accessibility scores started emerging in the last 15 years. Those can be generally split into two categories: substructure-based approaches[146–148] and reaction-based approaches.[149–155]

In substructure-based approaches, the similarity of structural features to those present in previously synthesised molecules is captured in order to assess synthetic accessibility. SYBA (SYnthetic Bayesian Accessibility) was trained to classify molecules into easy-to-synthesise and hard-to-synthesise, with the hard-to-synthesise molecules in the training set being artificially generated.[147] GASA (Graph Attention-based

assessment of Synthetic Accessibility) followed a similar approach, but used a graph representation for the molecules and was much more successful at distinguishing the effect of small structural differences on synthesisability.[146]

Conversely, in reaction-based approaches, the knowledge from known synthesis pathways is used to train the models. SCScore (Synthetic Complexity score) was trained on Reaxys[156] data with the assumption that a product of a reaction is more synthetically complex than the reactants.[153] Neeser *et al.* follow the same approach to pre-train their Focused Synthesizability score (FSscore), but then fine-tune it with human feedback for improved performance on chemical areas of interest.[155] On the other hand, both Retrosynthetic Accessibility score (RAscore)[149] and RetroGNN[151] depend on data generated by retrosynthesis prediction tools to assess synthesisability. Extending this idea, Calvi *et al.* provide the option to include information about intermediates at inference to boost the performance of their Leap score.[150] Focusing on medicinal chemistry-relevant reactions, Kim *et al.* trained DFRscore (Drug-Focused Retrosynthetic score) to predict the number of reaction steps based on a curated set of reaction templates.[154]

The recent study of Skoraczynski *et al.*[157] benchmarked four of the most popular synthetic accessibility scores: SAscore,[142] SYBA,[147] SCScore[153] and RAscore.[149] They found that SAscore and RAscore reflected the outcome of synthesis planning tools quite well and could be reliably used as their surrogates. However, with the only rule-based score among the benchmarked, SAscore, performing the best, they highlighted the need to include the human chemists' intuition in synthesisability scoring and showed that there is still room for improvement for the ML-based scores.

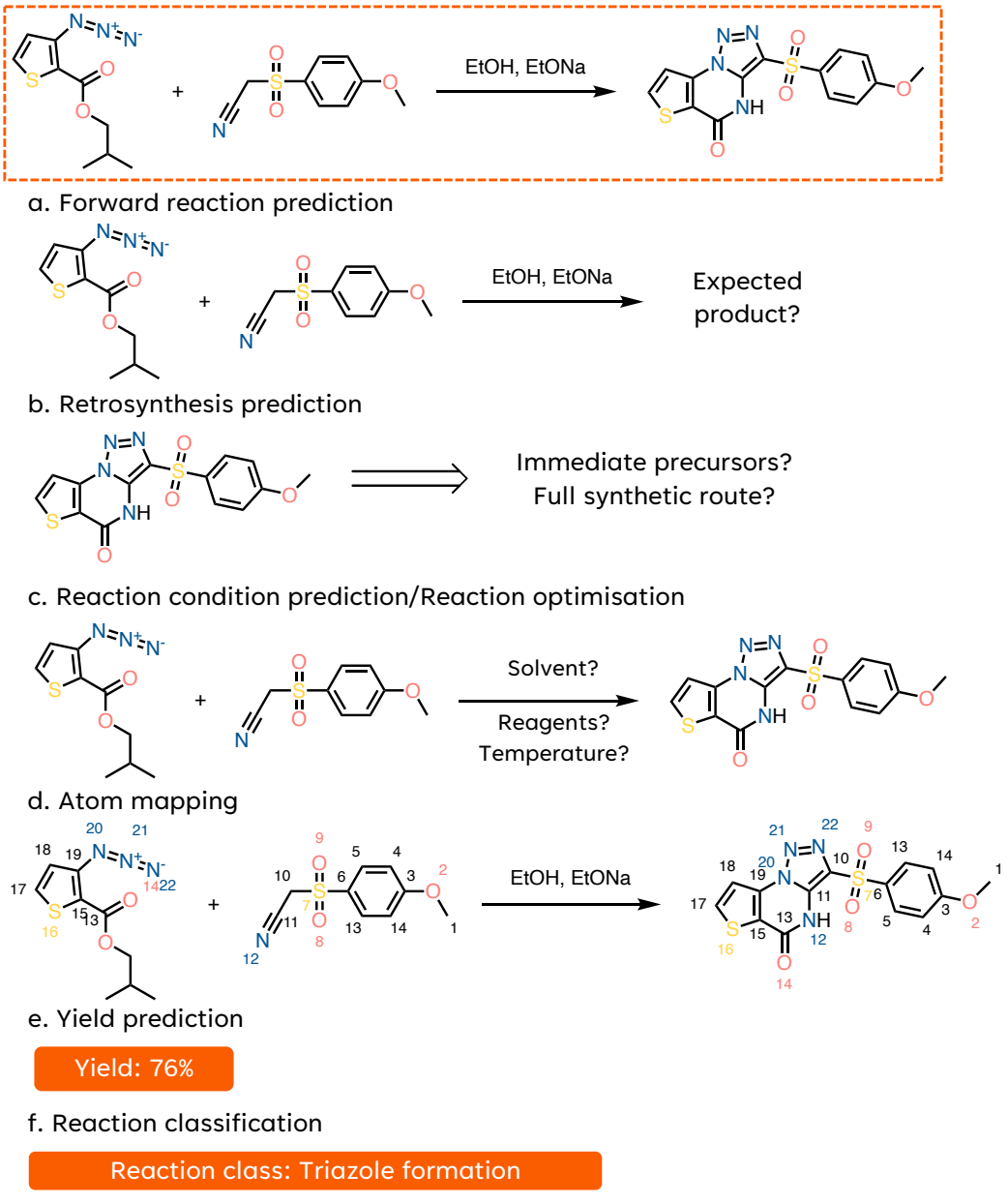


Figure 1.3: Overview of common reaction prediction related tasks in machine learning. (a.) In forward reaction prediction, the product of a reaction is predicted based on the input reactants. (b.) In retrosynthesis prediction, either just the reactants are predicted based on the input product (single-step retrosynthesis) or the full synthetic route (multi-step retrosynthesis). (c.) Reaction optimisation aims to select the best conditions (solvent, catalyst, temperature) to achievie the highest product yield. (d.) In atom mapping, the correspondence between atoms in the product and reactants is established. Models for various reaction (e.) regression and (f.) classification tasks have been developed to assess reaction performance and aid in prediction analysis.

1.4 Reaction prediction tasks

Although not as popular as molecular property prediction in drug discovery, machine learning has also been applied to various reaction-related prediction tasks, such as reaction product prediction, retrosynthesis prediction, reaction condition optimisation, atom-mapping, or various regression and classification tasks (Figure 1.3).[158–160] Even though not all of those tasks relate directly to predicting molecule synthesisability, they can be used in drug discovery in DMTA cycles to select molecules made through reactions with the most likelihood of success or the shortest/easiest synthesis route.

Forward reaction prediction

The task of forward reaction prediction (Figure 1.3a) is defined as the prediction of a reaction product given the reactants (and potentially also the reagents) as input. In most cases this is a qualitative prediction of the major reaction product, with yield and selectivity prediction defined as separate tasks. While forward reaction prediction models cannot be directly used to assess molecule synthesisability, they are often employed to check the predictions of retrosynthesis models.

The early work in forward reaction prediction treated it as a multi-class classification problem, with the goal being to predict the reaction type or reaction template most applicable to the reactants, which then implicitly defines the reaction product.[161–163] Both Wei *et al.*[161] and Segler *et al.*[162] trained their models to predict the most likely template based on a concatenated or summed reactant fingerprint representation. Coley *et al.* employed a two-step approach, where first a library of general templates was applied to the reactants and then a machine learning model was used to rank the predicted products.[163] While showing very good performance on their test sets, the applicability of those models is limited to the chemistry encoded in the extracted templates.

In recent years, the use of template-based models in forward reaction predic-

tion has almost disappeared, as they are outperformed by template-free methods: both translation-based and graph-edit based. In translation-based approaches, a sequence-based representation of molecules is used (usually SMILES) and the product prediction problem can be cast as a reaction to product translation, and therefore approaches for neural machine translation can be adopted. The first models were based on recurrent neural networks,[164, 165] but a major breakthrough came with the application of the Transformer architecture by Schwaller *et al.* in their Molecular Transformer.[166] Later developments in translation-based approaches utilised the Transformer architecture, but employed pretraining,[167] updated the SMILES representation,[168] or made use of graph-aware encoders[169, 170] with the aim of improving performance.

Coley *et al.* developed a graph-edit based approach involving a two-stage pipeline, where two separate graph convolutional networks are used first to identify reactive sites by predicting scores for each bond change and then to re-rank the products of the most likely bond changes.[171] An alternative approach was used by Sacha *et al.*, who represented the reaction as a sequence of graph edits (bond or atom deletion/addition/edit) that can be applied to the reactants.[172] In a pseudo-mechanistic approach, Bradshaw *et al.* instead choose to model the sequence of electron flow in ELECTRO.[173]

While the performance of template-free models surpasses 90% top-1 accuracy on general benchmarks, recent work has focused on improving their performance on specialised reaction classes. Pesciulessi *et al.* used transfer learning to improve the Molecular Transformer for carbohydrate reactions.[174] Zhang *et al.*[175] and Wang *et al.*[176] followed a similar approach for Baeyer-Villiger and Heck reactions respectively. Kreutter *et al.*[177] and Probst *et al.*[178] extended the Molecular Transformer for use with biocatalysed reactions, representing the enzymes either by their names or Enzyme Commission numbers.

Retrosynthesis prediction

Retrosynthesis prediction (Figure 1.3b) can refer to either the prediction of the reactants given a product of the reaction (single-step retrosynthesis) or the prediction of a full synthesis route towards a molecule (multi-step synthesis planning).[179–181] Single-step retrosynthesis might initially look like the inverse to forward synthesis prediction, and indeed similar machine learning approaches have been applied to both problems. However, while forward synthesis prediction can be thought of as a one-to-one function, with only a single reaction product assumed to be possible, retrosynthesis prediction is a one-to-many problem, with often many chemically valid ways of synthesising the same molecule. This makes the retrosynthesis prediction task more ambiguous and the evaluation and comparison of the developed methods more difficult.

As the retrosynthesis prediction task is directly involved in molecule synthesisability prediction and is one of the main subjects of this thesis, it will be discussed separately in more detail in Section 1.5.

Reaction condition prediction

Connected to retrosynthesis prediction is the task of reaction condition (or reagent) prediction (Figure 1.3c). As most retrosynthesis prediction models are not trained with reagent data included, a separate model is often used to recommend the most optimal conditions, i.e., those expected to result in the highest reaction yield. These models can be global[182, 183] (trained on a variety of reaction classes) or local[184, 185] (specific to one reaction type). While the global models of Maser *et al.*[182] and Gao *et al.*[183] have shown an improved performance over the popularity baseline (most common reagent selected), the issues of noise and bias in literature data impacting model performance have recently been highlighted by Beker *et al.*[186]

Although the models mentioned above have been trained to predict the chemical identity of reagents, they are not capable of providing other, more physical, reaction

conditions, such as concentrations, order of addition or purification methods. To solve this problem Vaucher *et al.* developed a model to predict experimental procedures, however the identity of all chemicals (including reagents) was required as input.[187]

Recent years have also seen an increase in popularity for reaction optimisation methods, which aim to improve the reaction outcome through iterative cycles of prediction and experiments. Among those, the use of Bayesian Optimisation is the most prominent,[188, 189] with methods based on active learning with surrogate ML models or deep reinforcement learning also developed.[190, 191]

Atom mapping

The task of atom mapping (Figure 1.3d) involves the creation of a mapping between the atoms present in the reactants and the product of a reaction. Atom mapping is usually performed to assist in other reaction prediction tasks, for example to separate reactants from reagents or extract reaction templates for forward reaction prediction, retrosynthesis, or reaction classification.

Traditional approaches for atom mapping were either optimisation-based, which focused on minimising the number of broken and formed bonds in a reaction, or involved maximum common substructure (MCS) algorithms.[192] However, the efficiency and accuracy of those methods were low and atom mapping was viewed as a difficult problem. An alternative approach was based on matching expert-encoded rules (templates) to the reaction, with NameRxn being the most prominent example.[193] In recent years, deep learning approaches have started to emerge and have shown promising results. Schwaller *et al.* developed RXNMapper, which mapped the reaction based on attention weights of a Transformer model.[194] A similar unsupervised learning strategy was applied in the GraphomerMapper, but using a graph-based Transformer and training on a larger reaction dataset.[195] Even more recently, Chen *et al.* developed LocalMapper, which was trained on human-labelled atom mapping data, with an active learning loop used to improve labelling efficiency.[196]

Regression tasks

While there are many reaction-related regression tasks, such as reaction rate constant and selectivity prediction, yield prediction (Figure 1.3e) is one of the most directly relevant to synthesisability and reaction performance assessment. Most yield prediction models to date have been trained on high-throughput experimentation (HTE) datasets, such as the Buchwald-Hartwig amination,[197–200] Suzuki-Miyaura cross-coupling,[197, 198, 201] Negishi reaction,[198] or alcohol deoxyfluorination[202] datasets. Both simple machine learning models, such as Random Forests and Support Vector Machines,[198, 200, 202] as well as deep learning models, such as the YieldBERT of Schwaller *et al.*,[197] have been used for this tasks, generally showing good performance on HTE data. Less success has been seen for more diverse datasets, such as patent data from the USPTO dataset or literature data from Reaxys.[197, 203] The reasons for this are most likely two-fold: the noisiness of the experimental data, with reactions yields being commonly acknowledged as difficult to reproduce by organic chemists; and the increased diversity of the data and confusing underlying trends, such as the change in yield distribution based on reaction scale uncovered by Schwaller *et al.*[197]

Out of the other regression tasks, reaction rate constant prediction is also often undertaken, although most work focuses on models for a specific reaction mechanism, such as S_N1 or S_N2 reactions.[204–206] Most models for reaction selectivity focus on predicting the reaction site (in a classification task),[207–209] but regression-based models that quantify stereoselectivity or enantioselectivity have also been introduced.[210, 211]

Reaction classification

Historically, the most prominent software for reaction classification (Figure 1.3f) is NameRxn, which depends on expert-encoded rules and uses a 3-tier reaction type classification system.[193] In recent years, machine learning models have started to

emerge for reaction classification based on different reaction fingerprint representations, but most often trained to predict the reaction labels assigned by NameRxn. Schneider *et al.* developed a reaction fingerprint based on the difference in atom-pair fingerprints of products and reactants, with the reagent data added as their physicochemical properties.[212] This fingerprint was then used to train a classification model for 50 reaction classes from NameRxn (out of now over 1800). Probst *et al.* also used a differential reaction fingerprint for reaction classification, but made no distinction between reactants and reagents.[213] Schwaller *et al.* developed a Transformer-based reaction classification model, and the output of its encoder can also be used as a reaction fingerprint (rxnfp).[214]

The use of reaction classification in reaction prediction workflows is limited. Early work in forward reaction prediction or retrosynthesis used reaction type as model input to aid the prediction. On the other hand, reaction classification is more commonly used in retrospective analysis, for example to look at trends in reactions performed in medicinal chemistry or to analyse predictions of other models.[215]

1.5 Computer-Aided Synthesis Planning

Retrosynthesis is the thought process of iteratively disconnecting a molecule into smaller fragments, breaking bonds for which there are known reactions that can form them, until all remaining fragments are commercially available.[216] Before the introduction of Computer-Aided Synthesis Planning (CASP) tools and searchable databases of reactions, synthetic chemists would need to go through this process every time they wanted to make a new molecule. In computational chemistry, retrosynthesis prediction (or CASP) is usually split into two separate tasks: single-step retrosynthesis, where models are trained to predict the most likely set of precursors for a molecule, and multi-step retrosynthesis, where those models are used with tree

search algorithms to predict full synthesis routes (Figure 1.4a).

The earliest efforts in computational retrosynthesis prediction (Figure 1.4b) date back to 1970s and Corey's work on LHASA (Logic and Heuristics Applied to Synthetic Analysis).[217] For a long time retrosynthesis prediction tools relied on the iterative application of expert-defined reaction rules (templates), with Chematica (now Synthia) being one of the most prominent examples.[218, 219] However, the task of manually curating the templates is laborious and time-consuming, rendering this approach inefficient. A breakthrough in retrosynthesis prediction came with Segler *et al.*'s work in 2018,[220] that signified a move from expert-based methods to the broad application of deep learning. Their approach utilised a neural network to select the most promising templates in single-step retrosynthesis and combined it with Monte-Carlo Tree Search (MCTS) to streamline multi-step retrosynthesis. Coley *et al.* and Genheden *et al.* adopted a similar approach, with the use of MCTS and template-based models, to create their open-source tools (ASKCOS[221] and AiZynthFinder[222]) which are widely used to this day. While the template-based approach has been very successful, it has the downside of depending on the template extraction process, which can be either time-consuming (when manual) or prone to error (when automated). To circumvent this, template-free approaches have been developed, with the first sequence-to-sequence model published by Liu *et al.*[223] Since then a plethora of both sequence-based and graph-based approaches have been developed, with IBM RXN being the first open-access (but not open-source) software employing a template-free approach.[224, 225]

A brief overview of methods used for single-step retrosynthesis and multi-step retrosynthesis is provided below.

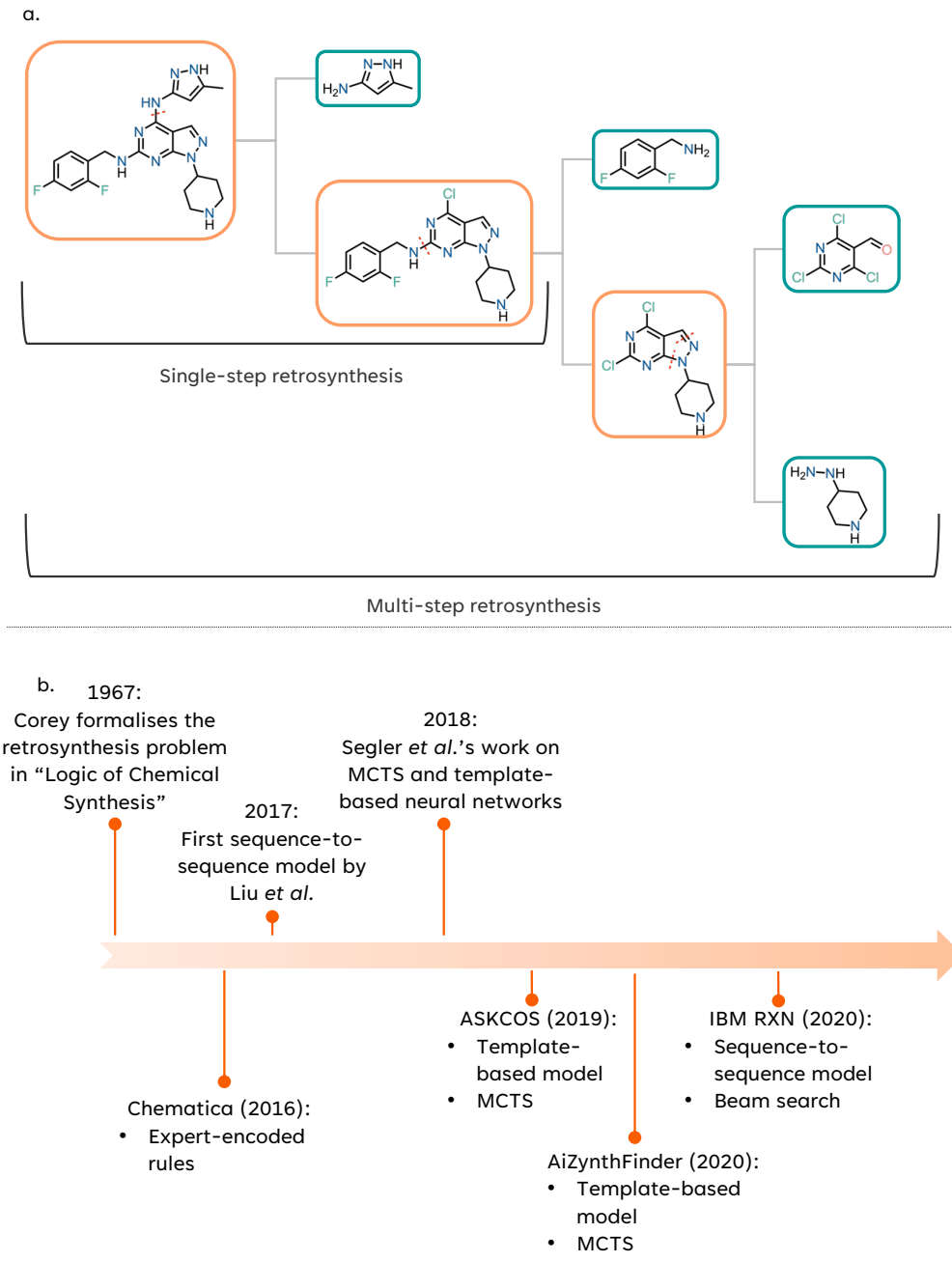


Figure 1.4: Overview of retrosynthesis prediction. (a.) Retrosynthesis prediction can be split into two tasks: single-step retrosynthesis, where one disconnection is predicted at a time, and multi-step retrosynthesis, where a full route is predicted. (b.) A timeline of the most important developments in retrosynthesis prediction.

1.5.1 Single-step retrosynthesis

1.5.1.1 Models

The methods used for single-step retrosynthesis can be generally split into template-based and template-free models, with the latter further split into sequence-based and graph-based approaches.

Template-based approaches rely on reaction templates, which are substructure patterns that describe the atom and bond changes between the product molecule and reactants. Deep-learning template-based approaches have started to gain popularity with the move from manual template curation to automatic extraction of templates from reaction databases with the help of atom mapping. NeuralSym was trained to predict the most likely template based on the extended connectivity fingerprints of the reactants.[162] Fortunato *et al.* improved the performance of this approach and expanded the chemistry scope using pretraining and data augmentation.[226] Nowadays, the state-of-the-art template-based model is LocalRetro, which divides the templates into atom-change, bond-change and multiple-change templates in order to be better able to focus on the local environment of the reaction.[227]

Sequence-based methods treat the task of retrosynthesis prediction as a product-to-reactant translation problem, which is enabled by the use of a text-based SMILES representation. Liu *et al.* pioneered this approach with their model based on long short-term memory (LSTM) cells,[223] however this architecture was soon replaced by the Transformer, starting with the work of Karpov *et al.*[228] Later work aimed to improve the performance by applying SMILES syntax correction,[229] performing data augmentation,[230] and using pre-training strategies.[167] Irwin *et al.* have shown that while pre-training significantly improved the top-1 accuracy of the Transformer model, it also decreased the top-10 accuracy, indicating the reduced diversity of the predictions.[167] Other approaches, such as that of Zhong *et al.*,[168] focused

on improving the reaction representation as a means of model performance improvement. They created root-aligned SMILES with a tightly aligned one-to-one mapping between product and reactant SMILES that reduced the edit distance and therefore relieved the model from learning the complex syntax. By combining this representation with the data augmentation strategy proposed by Tetko *et al.*[230] they achieved state-of-the-art performance for sequence-based methods. On the other hand, Ucak *et al.* treat the retrosynthesis problem as a translation between the atomic environments, instead of SMILES, in their RetroTRAE and achieve a further incremental increase in top-1 accuracy.[231] Most recently, Han *et al.* have introduced an iterative sequence editing model that works on SMILES and achieves top-1 accuracy of 60.8%, outperforming all other models on the USPTO-50k benchmark.[232]

In contrast to sequence-based approaches, graph-based approaches use a graph representation of the product and apply a sequence of graph edits to convert it into plausible reactants. The most representative model is MEGAN, which uses a graph encoder-decoder architecture to predict an edit action (deletion, addition, or modification of atom/bond) given the product or intermediate structure.[172] Graph2Edits employed the same principle of applying graph edits, but simplified the neural network architecture and adjusted the edit actions to increase the top-1 accuracy to 55.1%.[233] Several semi-template based approaches have also been developed, which follow a two-step approach of first breaking the product into synthons and then predicting the reactants based on those synthons.[234, 235]

1.5.1.2 Datasets

Most open-source single-step retrosynthesis methods are developed on the USPTO (United States Patent and Trademark Office) dataset of reactions extracted by Lowe *et al.*,[236] or its subsets developed for benchmarking: USPTO-50k,[237] USPTO-MIT,[238] or USPTO-FULL.[239] While USPTO-50k is most commonly employed

for retrosynthesis benchmarking, the small size and limited diversity of the reactions can lead to overestimation of performance for real-life applications. USPTO-MIT is much bigger at about 400k reactions, but less diverse than USPTO-50k and does not include any stereochemistry. USPTO-FULL is both the biggest and most diverse subset, at around 1M reactions.

While the wide use of the USPTO dataset can be attributed to it being the only publicly available dataset, many commercial reaction datasets are much greater in size. Their reaction sources can include either patent data, as is the case for Pistachio (over 13M reactions),[240] or both literature and patent data, as in the Chemical Abstracts Service[241] (CAS, over 150M reactions) and Reaxys[156] (over 70M reactions). The diversity of the literature-based data is also higher than that of patent datasets, and they occupy distinct chemical spaces, as highlighted by Thakkar *et al.*[242]

The noisiness of the reaction datasets remains a large issue in retrosynthesis prediction. To this end, Toniato *et al.* proposed a dataset de-noising framework that was based on the phenomenon of catastrophic forgetting in neural networks, with the reactions that are "forgotten" during training being assumed to be noise rather than true data.[243] In an alternative approach, atom mapping can also be used to clean datasets, with only the reactions that can be mapped retained as correct.[242]

1.5.1.3 Metrics

The most commonly used metric for assessing single-step retrosynthesis prediction models is top-k accuracy, which corresponds to the proportion of test reactions for which at least one correct reactant set (matching the ground truth) appears in the top-k predictions. Variations on this metric have been employed to make it less stringent, for example by only considering the main, largest reactant as in MaxFrag accuaracy.[230] While easy to implement, top-k accuracy has been criticised in recent

years for not reflecting real-life applications of retrosynthesis when only classifying ground truth matches as correct predictions - while if a prediction matches the ground truth it is expected to be correct, there can also be other possible correct disconnections not matching the ground truth which will not be accounted for by the top-k accuracy.[224, 244] The alternative proposed is to use round-trip accuracy, which employs a forward reaction prediction model to predict the proportion of the proposed disconnections that are chemically viable.[224] While inclusive of the correct but not ground truth disconnections, the dependence of this approach on another machine learning model (often trained on the same data) can introduce errors into the evaluation. Other less commonly used, evaluation metrics include top-k invalid rate (for SMILES in sequence-based models), human expert evaluation, and reaction class diversity.[224]

1.5.2 Multi-step retrosynthesis

In multi-step retrosynthesis a synthesis route tree is constructed by the recursive application of single-step models to the target molecule until a suitable stop condition is reached: either all building blocks are available in the provided building block stock (usually a commercial molecule database or in-house compound catalogue) or the maximum number of iterations or synthesis tree depth has been reached. With the extremely large reaction space and the exponential growth of the tree at each layer, exhaustive searches are not viable, and so a number of graph search algorithms have been proposed for this task, such as MCTS,[220] Retro* [245] or beam search.[224]

All of those algorithms follow the same general framework with three phases: selection, expansion and update (Figure 1.5). First, the most promising molecule node is selected, based on a value function (using heuristics or a surrogate model). The node is then expanded, usually with a single-step model, to produce a set of precursor molecule nodes. Some filtering may follow to discard impossible or unlikely

reactions. The values of all molecule nodes along the expanded pathway are then updated in preparation for the next selection step. In MCTS there is a rollout step before the update, where the predicted precursors are evaluated with further iterative expansion, often using a more lightweight model.[220] Retro* (based on A* search) replaces the need for this rollout step by including the prediction of route cost in their value function.[245]

Unlike for single-step retrosynthesis, which has many quantitative metrics such as top-k and round-trip accuracy, the historical evaluation of multi-step retrosynthesis has mostly been performed by humans and qualitative in nature.[220] Recently, Genheden *et al.* introduced the PaRoutes benchmark, which included full synthesis route data extracted from the USPTO dataset.[246] They also proposed a number of quantitative metrics to be used with this benchmark, such as average search time, number of solved targets, number of route clusters, or tree edit distance to the reference route. Maziarz *et al.* provide a software framework (Syntheseus) for retrosynthesis benchmarking and propose a number of best practises in how multi-step retrosynthesis should be performed and benchmarked.[244]

1.6 Thesis aims and outline

This thesis aims to develop and benchmark machine learning methods that improve the synthesisability of compounds created and selected in Computer-Aided Drug Design, either through improving synthesis planning tools or generating molecules with a focus on their synthesisability.

Chapter 1 provided an introduction to uses of machine learning for medicinal and organic chemistry in drug discovery with a particular focus on synthesisability and synthesis prediction. An overview of the drug discovery process and Computer-Aided Drug Design was first provided, before a detailed background on synthetic accessi-

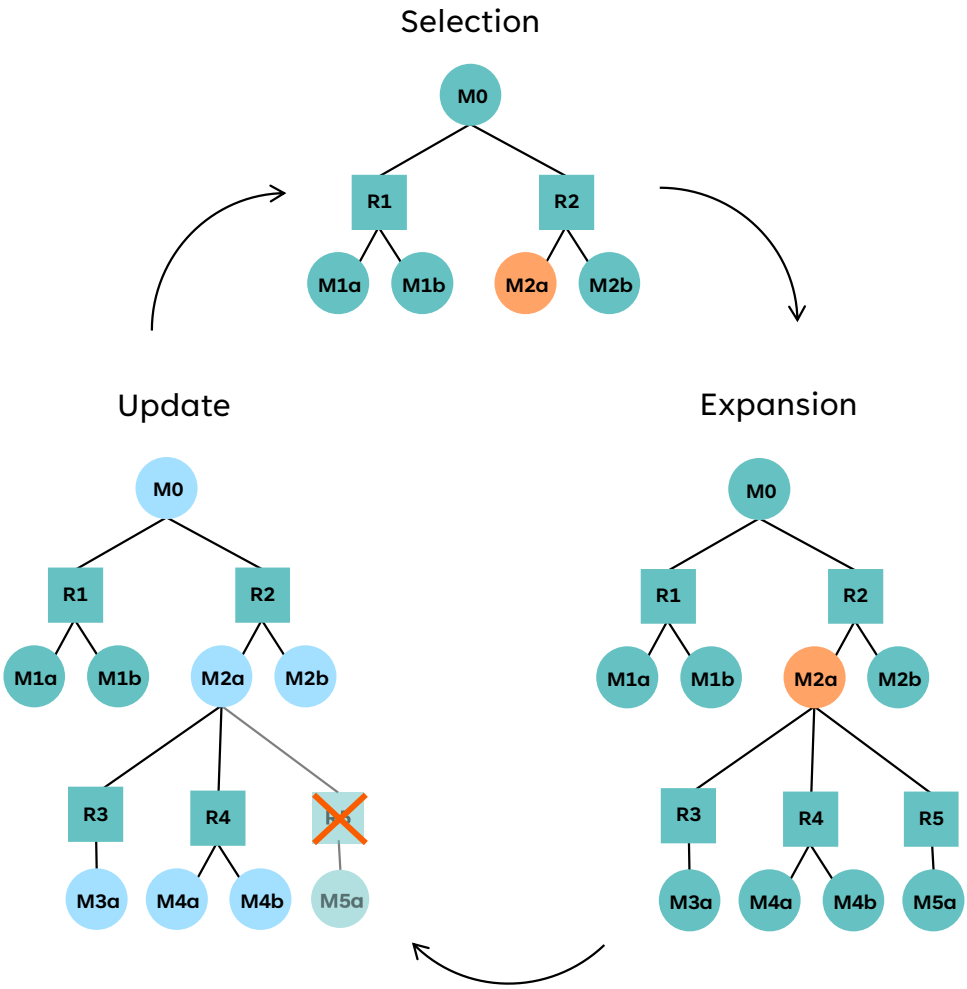


Figure 1.5: A generalised example iteration of a multi-step retrosynthesis algorithm. The synthesis route is represented as a tree with reaction and molecule nodes. First, the node to expand is selected based on a value function. Then, a single-step model is used to predict its precursors. The predictions are then evaluated and the value function is updated.

bility scores, reaction prediction tasks, and Computer-Aided Synthesis Planning was presented.

In Chapter 2, the theoretical background is introduced for the cheminformatics and machine learning methods used in this thesis. First, different molecule and reaction representations are discussed. Then, an overview of machine learning methods, including dataset splitting strategies, the models used, and training approaches in low data availability scenarios, is presented.

Chapter 3 outlines the development of a heterocycle-focused single-step retrosynthesis prediction model. Four domain adaptation approaches are benchmarked in order to improve the performance for ring-breaking disconnections. Moreover, a protocol is introduced for further training of the model as new reaction data becomes available. Finally, the applicability of the model in drug discovery is demonstrated in two retrospective case studies for newly published small molecule probes.

Chapter 2

Theory

This chapter provides the necessary computational background to the methods presented in this thesis. We will begin with an overview of the chemical data representations that can be used for machine learning and cheminformatics methods. Then, the machine learning models used later in this thesis will be introduced. Finally, a description of model training approaches in low data availability scenarios is provided.

2.1 Chemical data representations

This thesis focuses on two types of chemical entities: molecules and reactions. An overview of most commonly used representations for both of them is provided below. The representations that will be used later in this thesis are SMILES, SMARTS, and fingerprints for molecules and their substructures, and reaction SMILES, reaction SMARTS, and reaction fingerprints for reactions.

2.1.1 Molecule representations

The most intuitive molecule representation is a graph representation (Figure 2.1a), with atoms represented as nodes and bonds as edges. While the traditional graph

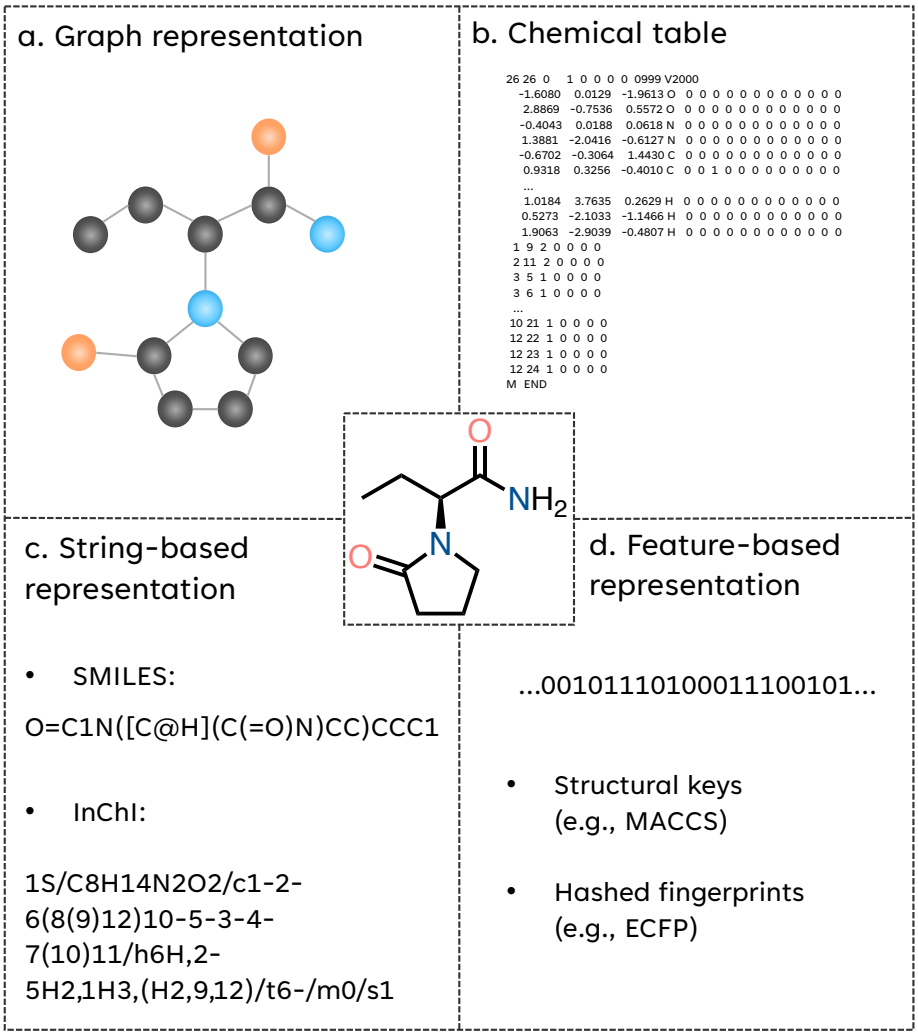


Figure 2.1: Most common molecule representations: (a.) graphs, (b.) chemical tables, (c.) string-based representations and (d.) feature-based representations.

representation is a 2D representation, 3D information, such as stereochemistry or atomic coordinates, can be encoded as node attributes. While graph representations have been directly used as input for some deep learning architectures (such as graph neural networks), they also serve as a starting point for the construction of other molecule representations discussed below.

One of the issues with graphs is the high memory requirement for the storage of this representation, with the size of matrices or tables that contain them growing as a function of the square of the number of atoms. Chemical tables, an example of which is Molfile, (Figure 2.1b) are closely related to the graph representation but are more memory efficient due to the way they encode the atom and bond data. The atom data is stored in the atom block, with atomic coordinates and atom types included, while the bond data is stored in the bond block, where bonds are defined by the indices of atoms that belong to them.

String-based representations (Figure 2.1c) are even more storage efficient, with the molecules described as a 1D sequence of characters that can convey 2D, and even some elements of 3D structure information (such as chirality). The first, and to this day the most commonly used, string-based representation is the Simplified Molecular Input Line Entry System (SMILES) developed by Weininger *et al.* in 1988.[85] In this system atoms are represented by their element symbol (lower case denoting an aromatic atom), with additional characters used for different types of bonds and structural features: '=' and '#' for double and triple bonds, '@' and '@@' to denote chirality, '\' and '/' to describe stereochemistry at a double bond, parentheses to show a branched chain and numbers to link atoms connected in a ring (Figure 2.2). The SMILES string is constructed by starting at a random node in the molecular graph and traversing through it, adding each atom to the string. Due to the nature of their construction, there can be multiple valid SMILES strings describing the same molecule This redundancy would be a problem in many machine learning applications,

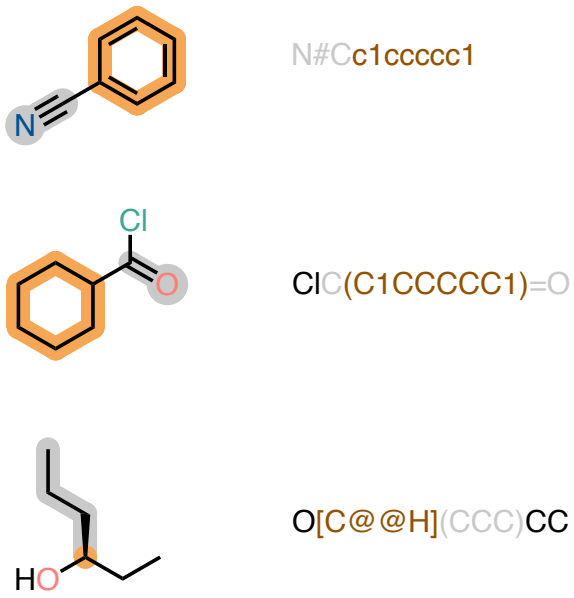


Figure 2.2: Example SMILES representations of molecules. Representative substructures and substrings are highlighted in the same colour (grey or orange).

so methods to canonicalise the representation are usually employed.[247]

An additional downside of SMILES strings is that not every generated string corresponds to a valid chemical structure. This has been a problem when using SMILES with generative models or genetic algorithms, and has lead to the development of SELF-referencIng Embedded Strings (SELFIES).[86] However, the use of SELFIES has not been widely adopted, potentially due to their more complex syntax and greater string length.

An alternative string-based representation are the InChIs (International Chemistry Identifier).[248] They consist of three text blocks containing the chemical formula, atom connectivity and hydrogen atom information. A hashed version of InChIs, InChI keys is also available. However, the less intuitive representation in InChIs is likely the cause of their much lower popularity when compared to SMILES.

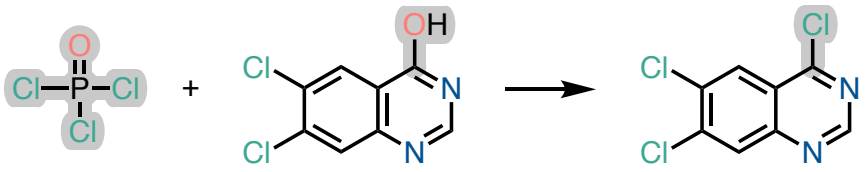
SMILES Arbitrary Target Specification (SMARTS) are an extension of the SMILES representation that has been specifically designed for substructure searching.[249] On top of the characters used in SMILES, it contains symbols meant to represent more

general constructs, such as any atom ('*') or any bond ('∼'). Moreover, logical operators can be used to combine patterns into more complex queries.

In contrast to the previously described representations which retain atomic information and can be converted back to the 2D molecular graph, feature-based representations (Figure 2.1d) extract higher level structural or physicochemical properties. They can be split into structural keys and hashed fingerprints. Structural keys, such as MACCS,[250] encode the absence or presence of predefined structural motifs into a bit vector. On the other hand, in hashed fingerprints the features encoded depend on the molecule itself. The most popular algorithm of this type are the Extended Connectivity Fingerprints (ECFPs), often calculated with the Morgan algorithm and then known as Morgan fingerprints.[251] They belong to the class of circular fingerprints and are generated by iteratively assessing the circular environment of each atom with increasing radius. The identifiers of each atom are updated with its neighbours until the maximum radius is reached. They are then converted to a single integer value through hashing, and the list of those integers forms a basis for the fingerprint. At the last step the identifiers are deduplicated and converted to the bit string representation.

2.1.2 Reaction representations

The SMILES system used for molecule representation can also be extended to reaction data (Figure 2.3). In reaction SMILES the reactants, reagents, and products are separated by '>' (in the format reactants>reagents>products), with the reagents being optional or sometimes combined with reactants (reactants and reagents>>products). The molecules on the same side of the reaction are then further separated by a period. Reaction SMILES can also include the atom mapping, with the mapping represented as numbers after the ':' symbol. Analogously to the SMARTS representation for molecular patterns, reaction SMARTS can be used to represent reaction



Reaction SMILES:

O=P(Cl)(Cl)Cl.OC1=NC=NC2=C1C=C(Cl)C(Cl)=C2>>ClC1=NC=NC2=CC(Cl)=C(Cl)C=C12

Reaction SMILES with atom mapping:

O=P(Cl)(Cl)[Cl:12].O[c:1]1[n:2][cH:3][n:4]c2[cH:5][c:6]([Cl:11])[c:7]([Cl:10])[cH:8][c:9]21
>>c12[n:4][cH:3][n:2][c:1]([Cl:12])[c:9]1[cH:8][c:7]([Cl:10])[c:6]([Cl:11])[cH:5]2

Reaction SMARTS corresponding to the reaction template:

O-[c:1].Cl-P(-Cl)(=O)-[Cl:2]>>[c:1]-[Cl:2]

Figure 2.3: String-based reaction representations: reaction SMILES without and with atom mapping, and reaction SMARTS corresponding to the template. Reaction template is highlighted in grey on the reaction scheme.

patterns and transformations. Most reaction templates extracted for retrosynthesis and reaction-based enumeration come in this format.

In the same way as in molecular feature-based representations, reactions can also be represented through fingerprints. The most commonly used fingerprints are either based on the difference in atom-pair fingerprints of products and reactants,[212, 213] or based on the learned representation of the encoder of a BERT (Bidirectional Encoder Representations from Transformers) model.[214]

2.2 Machine learning

Machine learning is a subfield of artificial intelligence (AI) that focuses on the study and development of statistical algorithms which aim to learn patterns from data. Such models can generalise to unseen data, performing tasks without pre-programmed instructions.[252] Machine learning can be split into supervised, unsupervised, and reinforcement learning, with the focus of this thesis being supervised learning. In

supervised learning models are trained on labelled data, with the input and output values for each datapoint being known. Common supervised learning algorithms include: Linear Regression, Logistic Regression, Random Forests, Support Vector Machines and Neural Networks.

The aim of supervised learning is to learn the parameters of a predictive model that minimise a loss function describing the difference between the target and predicted value. In practice, the dataset is first divided into training, test, and (optionally) validation sets. The model is fit on the training set to optimise the data-dependent parameters of the model. The predictions of the fitted model on the validation set might be used to estimate its performance to optimise non-data-dependent parameters of the model ("hyperparameters"). Finally, the model is used to make predictions on a test set to assess its performance on unseen data. The use of a test set is important to check for over-fitting, a phenomenon where the model matches/memorises the training set too closely and is not able to generalise to new data. While early work used random splitting strategies to split the data into training, validation, and test sets, the use of more complex splitting strategies such as scaffold-based, time-based, or Tanimoto similarity-based splits has been advocated to better assess the model's generalisability.

Deep learning is a subset of machine learning that uses artificial neural networks to learn from raw data instead of hand-crafted features, with popular deep learning applications including computer vision and natural language processing.[252, 253] Neural networks usually consist of multiple layers, with the input layer receiving the input data, the output layer generating the output, and intermediate layers (known as hidden layers) processing and transforming the data. The layers are formed from individual neurons, and perform non-linear transformations on their input, typically of the form:

$$h = \sigma(Wx + b) \tag{2.1}$$

where W is a learnable weight matrix, x is a vector of inputs, b is a learnable vector of biases and σ is an optional non-linear activation function. The output h is then used as input for the next layer or provided as the output of the model. During model training, the weights are iteratively updated to minimise a predefined loss function. Most of the methods used for loss optimisation are gradient-based, such as stochastic gradient descent. These algorithms use a step size defined by the learning rate, one of the neural network's hyperparameters. The gradient-based optimisation of multi-layer neural network is enabled by backpropagation, a chain rule-based algorithm which adjusts the network's parameters each training epoch.[254]

The two main machine learning models used in this thesis are Random Forests and the Transformer, and as such they are described in detail below.

2.2.1 Model architectures

2.2.1.1 Random Forest

Random Forest (RF) is an ensemble of decision trees, in which every tree is trained on a random subset of the training set and with a random subset of features to prevent over-fitting.[255] For regression tasks, the output of RF is calculated as the mean of the predictions given by every tree.

During the training process of a regression decision tree the dataset is recursively split based on key features, starting with the root question node. When a new decision node is added, all possible data splits are considered and the feature boundary that minimises the mean squared error of the new nodes is selected. This process is repeated until the stopping condition is reached, for example maximum tree depth. Each leaf node is then assigned a target value, most commonly based on the average of the samples present in that node. The tree can then be used to make predictions on new data by traversing down in starting from the root and following the path determined by the new sample's features (Figure 2.4).

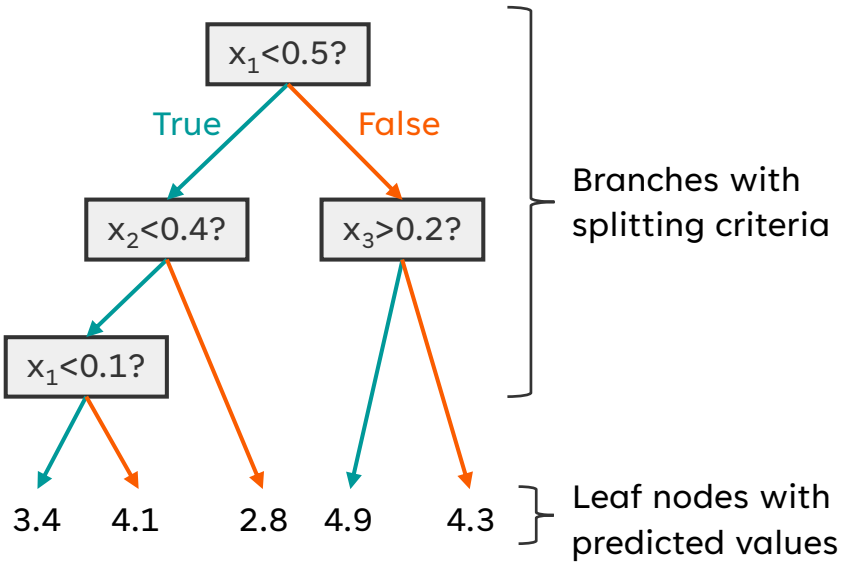


Figure 2.4: Example of a decision tree for a regression task. x_1, x_2 and x_3 denote different features.

While individual decision trees are very computationally efficient and interpretable, they are prone to over-fitting. The ensembling process used in RFs improves the performance and renders them robust to over-fitting, while retaining the interpretability and short training time, making this one of the more popular models for drug discovery tasks.

2.2.1.2 Transformer

The Transformer architecture was introduced by Vaswani *et al.* in 2017 and was a major breakthrough in the field of machine translation.[256] The original Transformer was developed for translation between languages with strings of text used as input. Transformers have an encoder-decoder architecture, where the encoder converts the input sequence of symbols into a context vector and the decoder uses this representation together with previously generated symbols to predict the output sequence in an auto-regressive process. The encoder and decoder consist of several stacked encoder or decoder blocks (6 in the original Transformer paper, but it can vary), with each

encoder block made up of a self-attention and a feed-forward layer, and each decoder block consisting of a self-attention, a encoder-decoder attention, and a feed-forward layer (Figure 2.5).

Multi-head attention

One of the key features of the Transformer model is the attention mechanism, which allows it to attend to different (and positionally distant) parts of the input data simultaneously. While attention was used in previous work,[257] Vaswani *et al.* introduced the concept of multi-head attention.

Attention takes as input the query (Q), key (K), and value (V) matrices. The Q, K and V matrices are created by multiplying the matrix of packed embeddings by a set of trainable weight matrices. The matrix product of Q and the transposed K is calculated and then scaled down by $\sqrt{d_k}$ (where d_k is the dimensionality of K), before applying a softmax function to obtain the attention weights of the values. The final attention score is the product of those attention weights and V and is computed as:

$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}})V \tag{2.2}$$

In multi-head attention this process is followed multiple times in parallel (8 in the original Transformer paper), but with different weight matrices used to produce Q, K and V for each head. The multi-head attention is then obtained by concatenating the heads and transforming them with trainable output weights:

$$MultiHead(Q, K, V) = Concat(head_1, ..., head_h)W^O \tag{2.3}$$

where

$$head_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \tag{2.4}$$

The use of multi-head attention improves the performance of the attention layer by allowing it to attend to different patterns simultaneously.

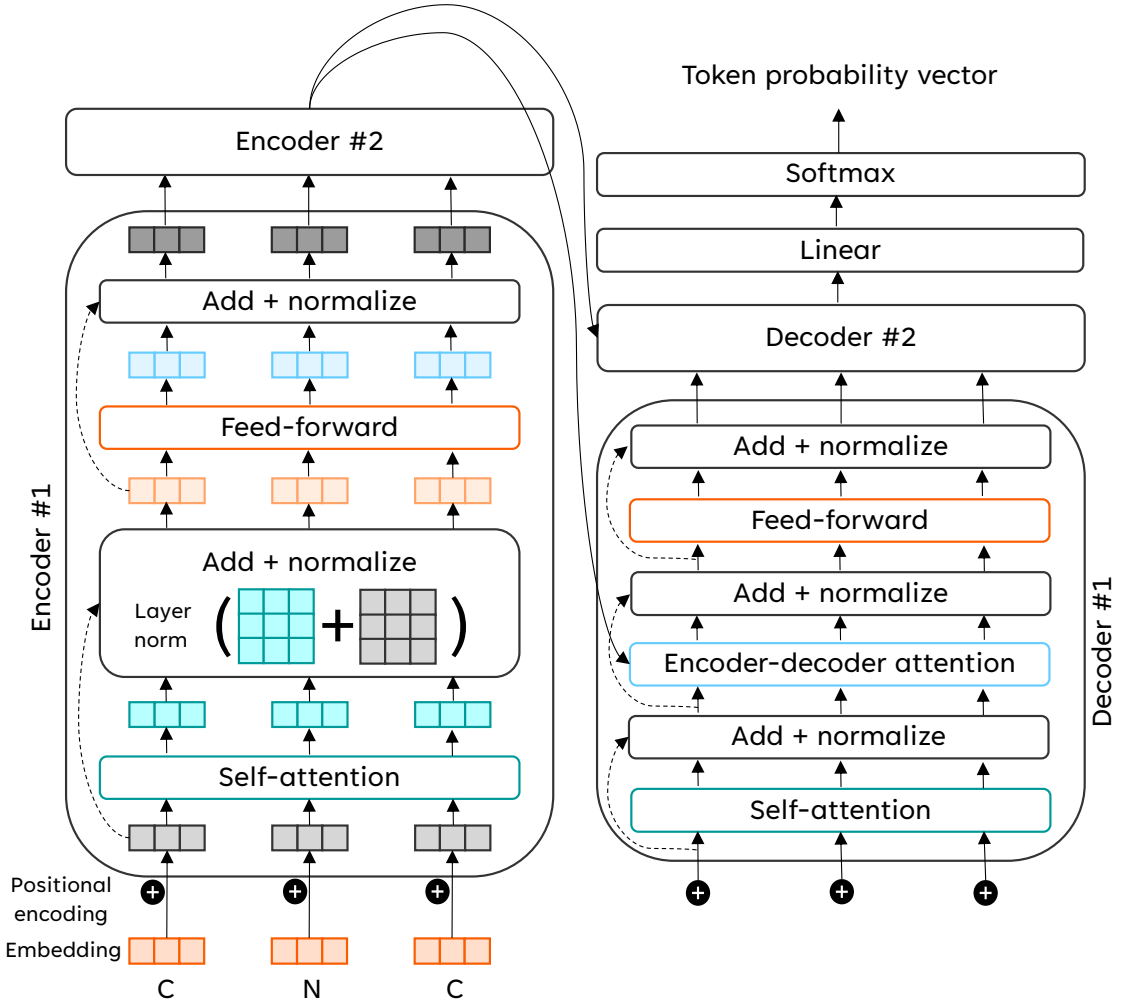


Figure 2.5: Architecture of a Transformer model. The input sequence is first embedded and positional information is encoded. The embeddings then enter a series of encoder blocks, each consisting of a multi-head attention layer (turquoise) and a point-wise feed-forward neural network (orange). After each of those layers, the residual connections are added and the layer is normalised. The output of the encoder block is the learned representation, which is then passed to the decoder. The input for the decoder is the previously predicted output, which also undergoes embedding and positional encoding. The decoder block consists of a self-attention layer, an encoder-decoder attention layer, and a point-wise feed forward neural network, with addition of residual connections and layer normalisation between each of those layers. The final output, which is a token probability vector, is obtained after passing through a linear and a softmax layer.

Tokenisation and input embedding

The input sequence is first split into words, sub-words or characters in a process known as tokenisation. For chemistry applications using SMILES, Schwaller *et al.* suggested a tokenisation into individual characters of the SMILES string (mostly corresponding to individual atoms).[166] The set of tokens used for a model is known as its vocabulary.

Each input token is then turned into a vector using an embedding algorithm. The embedding only happens in the first encoder block and each of the resulting vectors flows through its own path in the encoder.

Positional encoding

The Transformer model as described so far has no way to account for the order of the input tokens. To solve this issue in the original Transformer a positional encoding vector was added to each embedding before it is passed to the first encoder block. The positional encoding used there are sine and cosine functions of the token positions, but more recent approaches use positional encodings learned during training.[258]

Output generation

To convert the vector of floats output by the last decoder block into a predicted token it goes through a final linear and a softmax layer. The linear layer projects the output vector into a much larger logits vector that is the size of the model's vocabulary, with each logit corresponding to the score of a unique token. Those scores are then turned into probabilities by passing through the softmax layer and the token with the highest probability is selected as the prediction at this step time.

2.2.2 Training approaches in low data regimes

Chemical data is costly and time consuming to acquire. As such, many tasks in drug discovery, including those presented in this thesis, suffer from data scarcity. Two distinct machine learning approaches, transfer learning and active learning (Figure

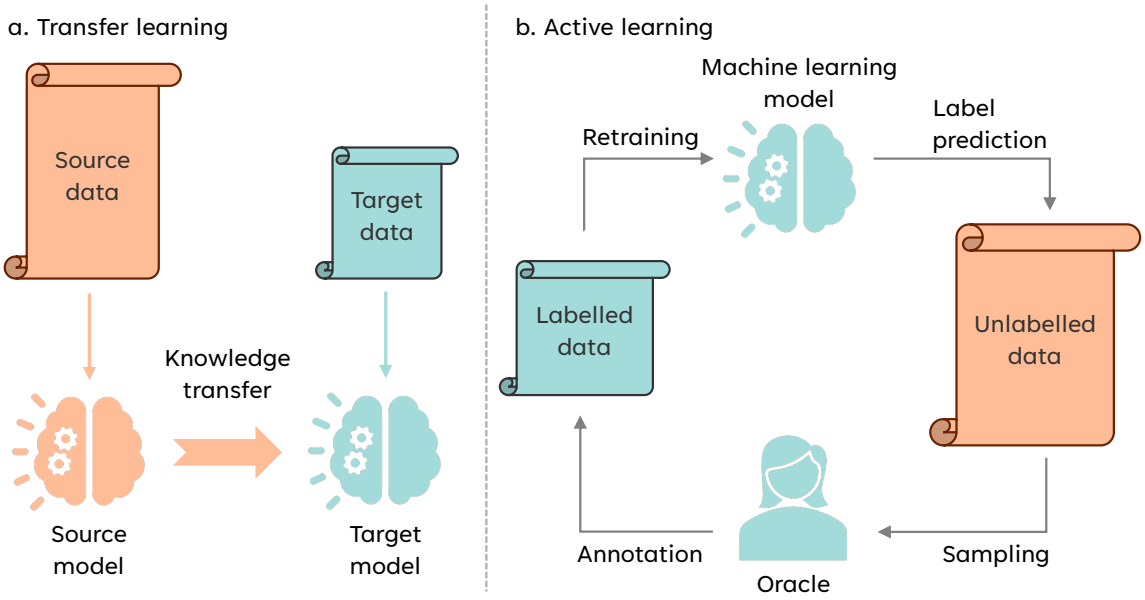


Figure 2.6: ML approaches for low data scenarios: (a.) transfer learning and (b.) active learning.

2.6), can be used in such scenarios and are employed in this thesis. A brief overview of each approach is included below.

2.2.2.1 Transfer learning

In transfer learning (Figure 2.6a) the knowledge learned from one task or dataset (source) is used to boost performance for another separate, but related, application (target).[259] As such, this approach is most applicable for tasks for which little data (labelled or unlabelled) is available, but data for either a similar task or the same task but from a different domain can be acquired.

Transfer learning can be split into three categories based on the differences in source and target tasks and domains. In the inductive transfer learning setting, the source and target tasks are different. The same is true for the second category, unsupervised transfer learning, but here the focus is solely on solving unsupervised tasks in the target domain. Finally, in transductive transfer learning the source and target tasks are the same but the domains are different. The work presented in this

thesis fits into the latter category.

One approach to transfer learning is feature extraction, where a pre-trained model is used as a fixed feature extractor. The final layers of this model are removed and replaced with new layers specific to the target task. The original layers are frozen and only the weights of the new layers are trained on the target dataset. In practice, during fine-tuning some of the pre-trained model's layers may be unfrozen and also trained with the new dataset.

A related concept to transfer learning is multi-task learning, where several different, but conceptually similar, tasks are learned at the same time, sharing an internal representation.[260] In contrast to transfer learning, all tasks are treated as equal here and knowledge transfer happens in all directions.

2.2.2.2 Active learning

Active learning (AL) can be applied in situations where there is an abundance of unlabelled data, but the labelling process is costly. AL involves an iterative cycle of model training, followed by prediction of the unlabelled data, the results of which form the basis for the selection of next samples to be labelled (Figure 2.6b).[261] Most approaches (including ours) use pool-based sampling for evaluation of the unlabelled data with the machine learning model, where labels are predicted for the whole unlabelled dataset.[262] This can be time- and memory-intensive, so alternatives such as stream-based selective sampling have been proposed.[263] In traditional AL the only goal is to improve the performance of the model and the strategies used to select samples for labelling reflect this, often being based on the uncertainty of the model's prediction or the expected effect on model error.

In the context of drug discovery, AL is most frequently applied in molecule screens, either for more costly *in silico* methods or experimental screens, where a surrogate machine learning model is trained to predict the output of those methods. Here, the

goal of AL is not only to improve the surrogate model's performance but also to select highly scoring molecules for labelling. As such, sample acquisition methods based more on exploitation, such as highest predicted score, might be more applicable.[264]

Chapter 3

Improving prediction of ring-breaking disconnections using transfer learning

Heterocycles are important scaffolds in medicinal chemistry that can be used to modulate the binding mode as well as pharmacokinetic properties of drugs. The importance of heterocycles has been exemplified by the publication of numerous datasets containing heterocyclic rings and their properties. However, those datasets lack synthetic routes towards the published heterocycles. Consequently, heterocycles with uncommon and novel substitution patterns, as well as theoretically designed novel scaffolds are not easily synthetically accessible by basing the proposed routes on previous literature. While retrosynthesis prediction models could usually be used to assist synthetic chemists, their performance is poor for heterocycle formation reactions due to low data availability.

This chapter explores different ways of overcoming the low data availability problem and improving the performance of single-step retrosynthesis prediction models for ring-breaking disconnections. Four domain adaptation approaches are benchmarked

against a general retrosynthesis model and models trained only on ring formation data. Out of the domain adaptation approaches, the *mixed fine-tuned* model is shown to be the most versatile, achieving top-1 accuracy of 36.5% and top-1 ring-breaking round-trip accuracy of 62.1%. Furthermore, a workflow for further fine-tuning the model is introduced to facilitate effective model retraining on new reaction data as it becomes available. Finally, the applicability of the *mixed fine-tuned* model in drug discovery is demonstrated by recreating synthetic routes towards two drug-like targets published last year.

The work presented in this chapter was previously published as a preprint: Wieczorek, E. *et al.* Transfer learning for Heterocycle Synthesis Prediction. *ChemRxiv* (2024). The training and evaluation of the template-based model and the predictions for multi-step case studies presented in this chapter were performed by Josh Sin.

3.1 Introduction

Heterocycles are key motifs in drug design, with 85% of the top 200 best-selling small molecule drugs of 2022 featuring heterocyclic rings (Figure 3.1).[266] The heterocyclic scaffolds can determine and restrict the shape of the molecule, helping to maintain key protein-ligand interactions. Moreover, through bioisosteric replacement of the rings with other heterocycles, pharmacokinetic and toxicological properties of lead compounds can often be improved.[267–269] Although numerous virtual libraries document theoretically synthesisable heterocyclic scaffolds,[270] synthetic pathways towards novel heterocycles remain underexplored, with the focus in medicinal chemistry being on ring derivatisation rather than ring formation.[215, 271]

Employing Computer-Aided Synthesis Planning (CASP) tools (see Section 1.5) to predict synthetic pathways towards those scaffolds could stimulate the exploration of novel heterocyclic molecules, potentially fueling new therapeutic breakthroughs.

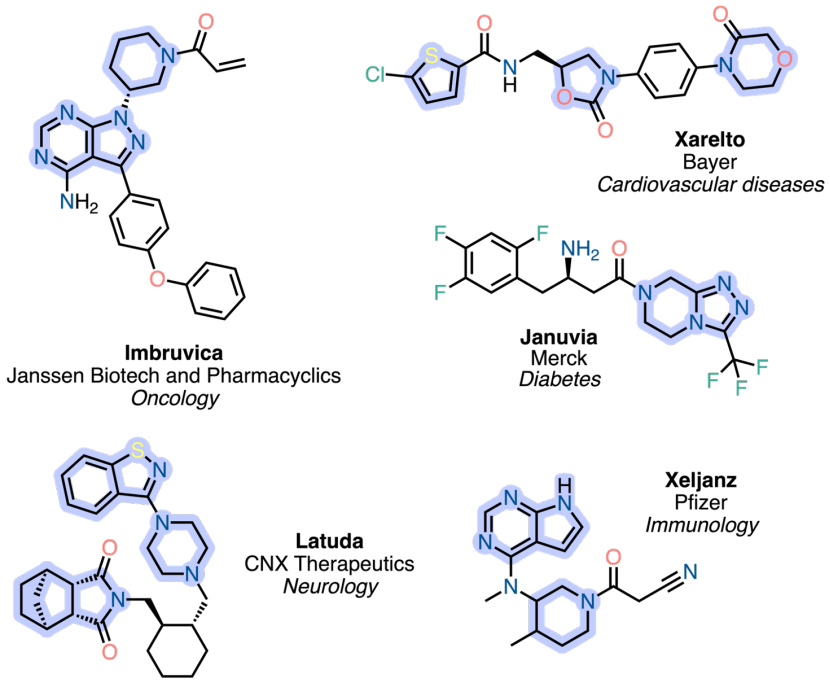


Figure 3.1: Example top 50 best-selling small molecule drugs in 2022 across different disease areas. Heterocycle substructures are highlighted in blue.

However, despite the high efficacy of CASP tools on general reaction datasets, predicting retrosynthetic disconnections for specific, less prevalent areas of chemistry remains a significant challenge due to dataset bias.[174, 272] Unfortunately, heterocycle formation reactions are an example of such underrepresented reaction class, accounting for only 5% of reported chemical reactions in the USPTO dataset,[236] the reaction dataset most commonly used to train retrosynthesis models.[272] This indicates the necessity to improve the performance of retrosynthesis models for ring-breaking disconnections in order for them to be useful in synthesis planning for heterocycle-containing molecules.

Recently, transfer learning, a machine learning approach where knowledge learned from one task is used to boost the performance on a related task (see Section 2.2.2.1), has been successfully introduced for reaction prediction tasks in low data scenarios. Two transfer learning approaches, fine-tuning and multi-task learning, have been applied for the forward reaction prediction of carbohydrate reactions[174] and Heck

reactions[176], as well as forward reaction and retrosynthesis prediction of enzymatic reactions[177, 178]. While both of those transfer learning approaches depend on the use of a large general reaction dataset and a smaller specialised dataset for training the model, they differ in how those datasets are used for model training: sequentially in fine-tuning (first longer pre-training on the general dataset and then short fine-tuning on the specialised dataset) and simultaneously in multi-task learning. Improvement was seen for the specialised task with both approaches, however each approach comes with different limitations. For example, in the reported examples, fine-tuning showed a quick training time and increased accuracy for reactions of interest but showed low performance for common reactions. Conversely, multi-task learning maintained good performance across reaction types but required longer training time, making it less suitable for frequent retraining as new data emerges.

While the use of transfer learning in reaction prediction is relatively novel, it has become widespread in other, more popular, machine learning applications, such as machine translation. In those areas, many different transfer learning (or domain adaptation) approaches have been developed that address the aforementioned limitations. For example, in mixed fine-tuning the sequential and simultaneous nature of, respectively, fine-tuning and multi-task learning are combined: the pre-trained model is fine-tuned on both the general and specialised dataset.[273] Meanwhile, in ensemble decoding it is the models' predictions themselves that are combined at the inference time - usually from a pre-trained and a fine-tuned model.[274] In both cases, this allows for a shorter training time without a significant loss in performance on the general dataset. As the task of adapting reaction prediction models towards specific reaction classes can be compared to adapting translation models for niche applications, it should be possible to use these domain adaptation approaches developed for language translation in retrosynthesis prediction.

This work aims to benchmark and use the various domain adaptation approaches

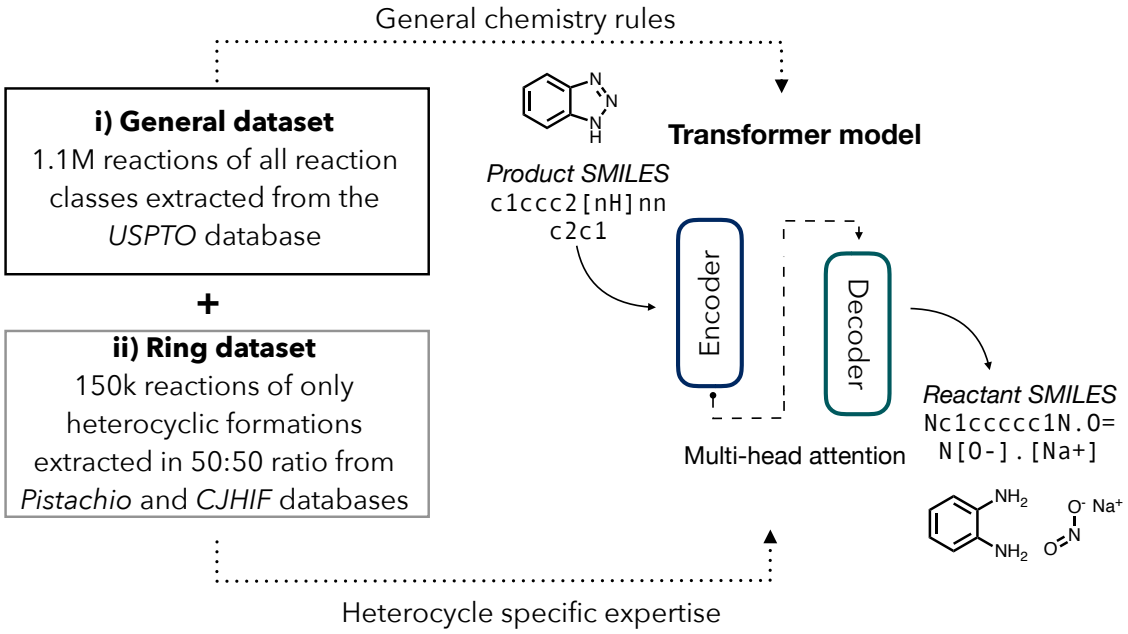


Figure 3.2: Overview of the domain adaptation approach used for training heterocycle-focused retrosynthesis prediction models in this study. All models are trained on both (i) a large dataset of all reaction classes and (ii) a much smaller dataset of only ring formations to, respectively, gain an understanding of general chemistry rules and ring disconnections.

discussed above to enhance the performance of CASP tools for heterocycle retrosynthesis by using them to train sequence-to-sequence (seq2seq) single-step retrosynthesis prediction models (Figure 3.2). We compare those methods to the template-based approach reported by Thakkar *et al.*, 'Ring Breaker', trained specifically for ring-forming reaction prediction.[272] To train these models we use a large dataset of all reaction types based on USPTO ('*General*') and a smaller dataset of just heterocycle formations ('*Ring*'). Our results show that the *mixed fine-tuned* model is the best for multi-step retrosynthesis, with a 10% increase in accuracy over the baseline for heterocycle formations and similar performance for other reactions. We test the model on recently developed heterocycle formations and demonstrate how it can be further fine-tuned to improve its accuracy on this new data. Finally, we demonstrate the applicability of the *mixed fine-tuned* model by predicting retrosynthetic routes for two recently published heterocycle-containing drug-like targets.

3.2 Materials and Methods

3.2.1 Data curation

Three separate dataset were used in this work: the *General*, *Ring* and *Recent* datasets. Both *General* and *Ring* datasets were used to train and benchmark the seq2seq models using various transfer learning strategies. The *Recent* dataset was used for further fine-tuning experiments. The details of each dataset's preparation are included below.

3.2.1.1 *General* and *Ring* Datasets

In this study, we utilised the USPTO dataset preprocessed by Pesciullesi *et al.*[174], containing ∼1.2M reactions from a wide variety of reaction classes. This dataset is henceforth referred to as the *General* dataset.

Additionally, we curated a dataset of ∼165k ring formation reactions, referred

here to as the *Ring* dataset, comprising about 80k reactions extracted from academic journals (CJHIF dataset[275]) and 80k reactions from additional patent data (Pistachio dataset, accessed 28th June 2022, version 2022Q1).[240] Both datasets first underwent a series of standardisations and validity filters. The reactions were canonicalised: each component was separately canonicalised with RDKit and the reactants were sorted alphabetically, ions in salts (and other fragments belonging together) were separated by "∼". Duplicate entries were removed together with reactions already present in USPTO. Reactions with more than one product were filtered out. Reagents in CJHIF were converted into the SMILES format with Chemical Identifier Resolver (https://cactus.nci.nih.gov). Reactions in CJHIF were atom-mapped with RXNmapper and entries with mapping confidence <50% were removed.[214] For Pistachio, the ring formation reactions were extracted based on the assigned reaction superclass "Heterocycle formations" (82,486 reactions). For CJHIF, the reactions were extracted based on the difference in the number of rings in the product and in the reactants as calculated by RDKit (83,623 reactions). Reactions from the two sources were then combined and duplicates dropped, resulting in the final *Ring* dataset containing 165,216 reactions.

A TMAP[276] visualisation of the chemical space of the datasets is shown in the Figure 3.3, indicating that ring-breaking reactions occupy distinct areas of the chemical space. The TMAP was created using rxnfp[214] embeddings.

3.2.1.2 *Recent* Dataset

For further fine-tuning experiments we manually extracted a set of 1,475 heterocycle formations from 47 scientific publications from 2022 reporting new methodologies for heterocycle synthesis, referred to as the *Recent* dataset. The articles used as the source of data are included in Appendix B. The preprocessed reactions can be found at `https://github.com/duartegroup/Het-retro/tree/main/data/`

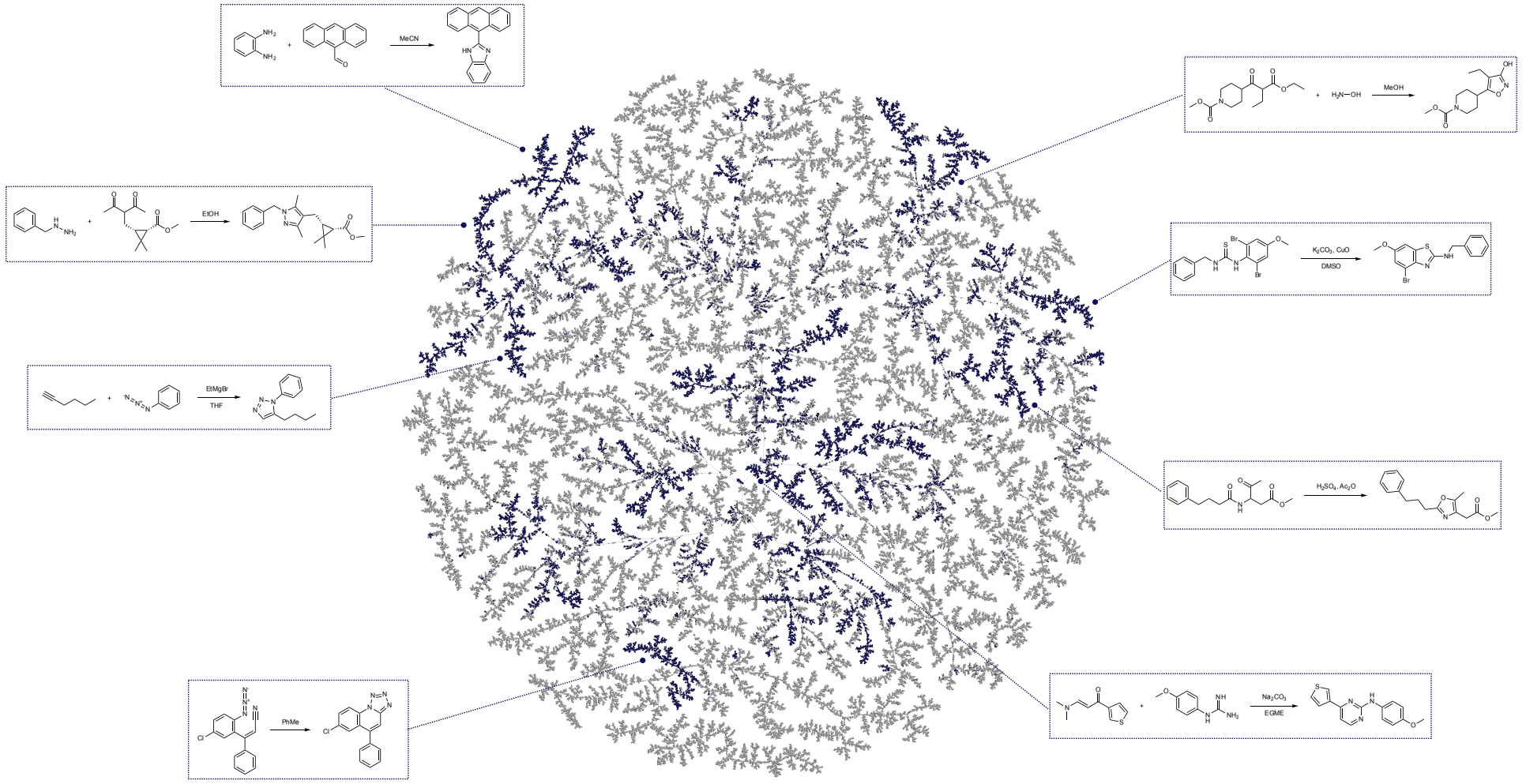


Figure 3.3: A visual representation of the chemical space of *Ring* (navy) and *General* (grey) datasets. Representative reactions from some of the larger *Ring* dataset clusters are shown.

`recent_dataset`.

3.2.1.3 Dataset splitting

The *General* dataset splits were retained from the original source.[174] The dataset was split randomly with a 90:5:5 train:validation:test set ratio, ensuring that all reactions with the same product were included in the same split.

The *Ring* dataset was split into train, validation, and test sets with 90:5:5 ratio based on the Tanimoto similarity of reaction products[277] using DeepChem.[278] Additionally, we performed a random split of the *Ring* dataset and trained the mixed fine-tuned model on the randomly split dataset to assess the effect of dataset splitting (Appendix C).

Due to the smaller size, the *Recent* dataset was split randomly into a train, validation and test sets with a ratio of 80:10:10.

The exact sizes for each dataset split are included in Table 3.1.

Table 3.1: Training, validation and test set sizes for each dataset used in this study.

Dataset	Train	Validation	Test
General	1,090,034	60,430	60,548
Ring	148,694	8,260	8,262
Random Ring	148,749	8,276	8,191
Recent	1,180	147	148

3.2.2 Model training and inference

Two different single-step retrosynthesis model architectures were used in this study: sequence-to-sequence and template-based models. All domain adaptation strategies described in Section 3.2.2.2 were attempted with sequence-to-sequence models, specifically Transformers. A template-based model was also trained for comparison, follow-

ing the approach of Ring Breaker.[272] The details of the models' hyperparameters and the domain adaptation strategies are included below.

3.2.2.1 Model architectures and hyperparameters

Sequence-to-sequence models

The Transformer models were trained using the OpenNMT-py package.[279]

The datasets were first preprocessed:

```
onmt_preprocess -train_src ${DATASET}/tgt-train.txt
${DATASET_TRANSFER}/product-train.txt -train_tgt ${DATASET}/src-train.txt
${DATASET_TRANSFER}/reactant-train.txt -train_ids general ring
-valid_src ${DATASET_TRANSFER}/product-valid.txt -valid_tgt
${DATASET_TRANSFER}/reactant-valid.txt -save_data ${DATADIR}/mft_retro
-src_seq_length 3000 -tgt_seq_length 3000 -src_vocab_size 3000
-tgt_vocab_size 3000 -share_vocab
```

An example command used to train the *mixed fine-tuned* model is included below:

```
onmt_train -data $DATADIR/mft_retro  -save_model
$MODELDIR/mft_model -save_checkpoint_steps 1000
-data_ids general ring --data_weights $WEIGHT1 $WEIGHT2
-seed $SEED -gpu_ranks 0   -train_steps 256000 -param_init 0
-param_init_glorot  -train_from $MODELDIR/retro_model_pretrained.pt
-max_generator_batches 32   -batch_size 6144 -batch_type tokens
-normalization tokens -max_grad_norm 0 -accum_count 4
-optim adam -adam_beta1 0.9 -adam_beta2 0.998 -decay_method noam
-warmup_steps 8000 -learning_rate 2 -label_smoothing 0.0
-layers 4 -rnn_size 384 -word_vec_size 384
-encoder_type transformer -decoder_type transformer
```

```
-dropout 0.1 -position_encoding -share_embeddings
-global_attention general -global_attention_function softmax
-self_attn_type scaled-dot -heads 8 -transformer_ff 2048
--tensorboard -tensorboard_log_dir $DATADIR/logs
```

All hyperparameters (apart from number of training steps and dataset weights) were kept constant for all models and used as in the work of Pesciullesi *et al.*[174] The -train_from argument was used for *fine-tuned* and *mixed fine-tuned* models to provide a path for the pre-trained model. The -data_weights argument was used for *multi-task* and *mixed fine-tuned* models to set the dataset weights.

The following command was used to make predictions:

```
onmt_translate -model $MODELDIR/retro_mft_model.pt
-src $DATADIR/ring_dataset/product-test.txt
-output $PREDDIR/retro_mft_ring_predictions_top5.txt
-n_best 5 -beam_size 5 -max_length 300 -batch_size 64 -gpu 0
```

The code used to train the models is available at `https://github.com/duartegroup/Het-retro` and the trained models are available at `https://figshare.com/articles/journal_contribution/Transfer_Learning_for_Heterocycle_Retrosynthesis/25723818`.

Template-based model

We trained a single-step template-based retrosynthesis prediction model on only ring-forming reactions based on the approach used by Thakkar *et al.* in 'Ring Breaker'.[272] Our dataset comprised reactions from the *Ring* dataset and ring formations extracted from the *General* dataset. Atom-mapping of reaction data was conducted using RXN-Mapper,[194] and reaction templates were subsequently extracted using RDKit[143] and RDChiral.[280] We used TensorFlow[281] to construct the multilabel classification neural network for prediction. The template-based model architecture utilises a multi-label classification neural network with one dense layer and was adapted from

'Ring Breaker'. We scanned various neural network hyperparameters using the model cross-entropy loss and top-1/3/5 reactant accuracy to identify the best-performing hyperparameter set. The scanned hyperparameters are summarised in Table 3.2. After the scan, we selected a dense layer size of 512 and a dropout rate of 0.8. All other hyperparameters were retained as in 'Ring Breaker'.

Table 3.2: Template-based neural network hyperparameters scanned.

Hyperparameter	Values scanned
Dense layer size	128, 256, 512, 1024
Dropout rate	0.5, 0.7, 0.8

3.2.2.2 Domain adaptation approaches

Four different domain adaptation approaches were benchmarked in this study: fine-tuning, multi-task learning, mixed fine-tuning and ensemble decoding (Figure 3.4). A *baseline* model was pre-trained on the *General* dataset for 250k steps (32 epochs) for further use in fine-tuning, mixed fine-tuning and ensemble decoding. Additionally, we trained a *ring-only* model as a second baseline, training it for 250k steps (244 epochs) on the *Ring* dataset only.

For both fine-tuning and multi-task learning, we followed the approach previously taken by Pesciullesi *et al.*[174] and used the number of training steps and dataset weight ratio that was found to be optimal in that study (Figure 3.4a). Therefore, the *fine-tuned* model was trained for 6000 steps (6 epochs) starting from the *baseline* model. The *multi-task* model was trained from scratch for 250k steps on both the *General* and *Ring* datasets with a 9:1 dataset weight ratio.

As mixed fine-tuning was not previously used for reaction prediction tasks, we benchmarked the effect of the number of fine-tuning steps and the dataset weight ratio (Appendix A). The final *mixed fine-tuned* model was trained on the *General*

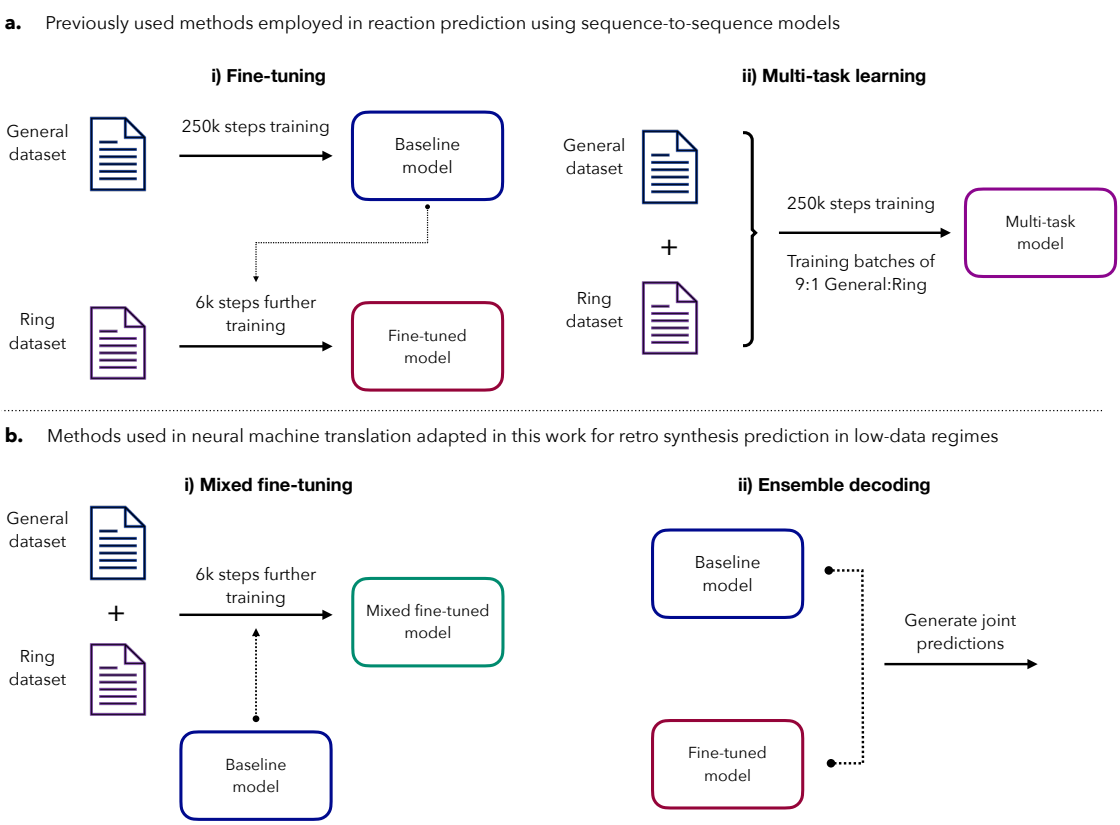


Figure 3.4: Overview of domain adaptation approaches used in this work for heterocycle retrosynthesis prediction. (a) Methods previously used for forward reaction prediction and retrosynthesis. Fine-tuning consists of pre-training a *baseline* model on a large dataset of all reaction classes, which is then fine-tuned on a smaller dataset of only reactions of interest. In multi-task learning, the model is trained on both datasets at the same time. (b) Methods previously only used in NLP tasks. In mixed fine-tuning, the *baseline* model is fine-tuned on both datasets. In ensemble decoding the prediction is made jointly with the *baseline* and *fine-tuned* model.

and *Ring* datasets with a weight ratio of 1:1 for 6000 steps starting from the *baseline* model (Figure 3.4bi).

Ensemble decoding was performed using the in-built OpenNMT-py functionality by providing two models at the inference time: the *baseline* and *fine-tuned* models (Figure 3.4bii). In this implementation, the final prediction is made by averaging the prediction distributions of both models.

3.2.3 Evaluation metrics

All the models described above were tested on the *Ring* and *General* datasets to assess their performance for ring-breaking disconnections and other reaction types. Two classes of metrics were used to evaluate the models (Figure 3.5):

- Recall-based metrics (top-N accuracy and reactant accuracy) that historically have been most commonly used to assess retrosynthesis prediction models and measure how well the models can recover the ground truth reactants from the test set.

- Precision-based metrics (round-trip accuracy, coverage, and ring-breaking round-trip accuracy), a newer set of metrics that aim to measure the chemical validity of the models' predictions.[224]

A description of each metric is included below.

Top-N accuracy: The proportion of test reactions where at least one of the top-N predictions contains all the ground truth precursors (reactants and reagents). As both reactants and reagents were used for model training, they were both included in this assessment, making this a harsher metric for our model than models trained only on reactant data.

Reactant accuracy: The proportion of test reactions where at least one of the top-N predictions contains all the ground truth reactants. This metric aligns more

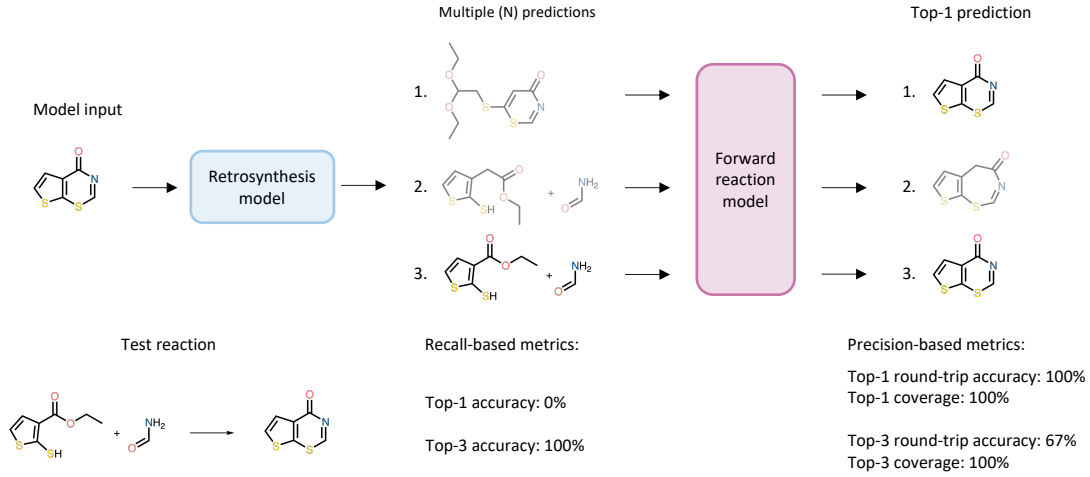


Figure 3.5: The workflow for calculating model metrics demonstrated on an example test reaction. To calculate the recall-based metrics the top-N predictions (here N=3) of the retrosynthesis model are compared to the ground truth test reaction. To calculate the precision-based metrics a forward reaction prediction model is used to predict the top-1 product of the output of the retrosynthesis model and this prediction is compared to the initial model input. The top-N accurate (middle) and round-trip accurate (right) predictions are highlighted.

closely with traditional top-N accuracy for models trained only on reactant data. The reactants were separated from reagents using the role assignments provided in the CJHIF and Pistachio datasets.

Round-trip accuracy: The proportion of all top-N predictions which are round-trip accurate. A round-trip accurate prediction is defined as a set of predicted precursors that, when used as input for a forward reaction prediction model, reproduce the original product molecule as a top-1 prediction.

Coverage: The proportion of test reactions for which at least one of the top-N predictions is round-trip accurate.

Ring-breaking round-trip accuracy: The proportion of all top-N predictions which are round-trip accurate and correspond to a ring breaking disconnection. A ring breaking disconnection is defined as one where the sum of the number of rings in the reactants is lower than the number of rings in the product, as calculated with RDKit.[143] The predicted reactants were separated from reagents through atom

mapping with RXNMapper[194] for the purpose of calculating this metric.

For all precision-based metrics, the forward reaction prediction model used was a multi-task model trained on the *Ring* and *General* datasets following the approach of Pesciullesi *et al.*[174] Table 3.3 contains the top-1/3/5 accuracies of this multi-task model together with the same metrics for a baseline forward reaction prediction model.

Table 3.3: Accuracy of the forward reaction prediction models on the *General* and *Ring* test sets.

	Test set					
Model	*General*			*Ring*		
	top-1	top-3	top-5	top-1	top-3	top-5
Baseline	78.0%	84.5%	86.7%	61.3%	76.9%	79.4%
Multi-task	78.4%	86.1%	87.3%	74.3%	88.3%	89.8%

3.2.4 Multi-step retrosynthesis

For the case studies, to adapt the trained single-step retrosynthesis prediction models to multi-step route planning tools, we used a neural-based A* search algorithm based on Retro*.[245] Multi-step route planning tools were constructed for both the baseline and mixed fine-tuned single-step models. The default search parameters were used as in Retro*. The stock molecule database chosen was eMolecules (version accessed with Retro* code implementation from Chen et al.[245], 11th January, 2019). Only the lowest cost route as output by Retro* was considered.

3.3 Results

3.3.1 Benchmarking domain adaptation approaches

We commenced our study by comparing the performance of different domain adaptation approaches, focusing on methods previously used for chemical reaction prediction (i.e., multi-task learning and fine-tuning) and methods employed in the NLP domain (mixed fine-tuning and ensemble decoding) (Figure 3.4). These approaches were benchmarked against the *baseline* and *ring-only* models as the two extremes of the training approaches: one trained only on the *General* data and one only on the *Ring* dataset. Additionally, the methods were compared to a *template-based* model trained following the strategy adapted by Ring Breaker,[272] as a benchmark against the only other heterocycle-specific retrosynthesis prediction model. The performance of all models was first tested on the *Ring* test set to measure the improvement in prediction of the ring-breaking disconnections and then on the *General* test set to assess their suitability for use in multi-step retrosynthesis, where other reaction classes also need to be predicted well.

3.3.1.1 Performance for ring-breaking disconnections

Figure 3.6 showcases the trends in the models' performance on the *Ring* test set with more detailed metrics for all models included in Table 3.4 (recall-based metrics) and Table 3.5 (precision-based metrics). Our results show that on the *Ring* test set, the *fine-tuned* model outperforms all other approaches, achieving a top-1 reactant accuracy of 40.5% (Figure 3.6A). Moreover, 69.5% of all its top-1 predictions are chemically valid and correspond to ring-breaking reactions (Figure 3.6B). The three other domain adaptation approaches also show improvement over the *baseline* model with top-1 reactant accuracies of around 36% and top-1 ring-breaking round-trip accuracies of around 62%. However, they perform similarly to the *ring-only* model,

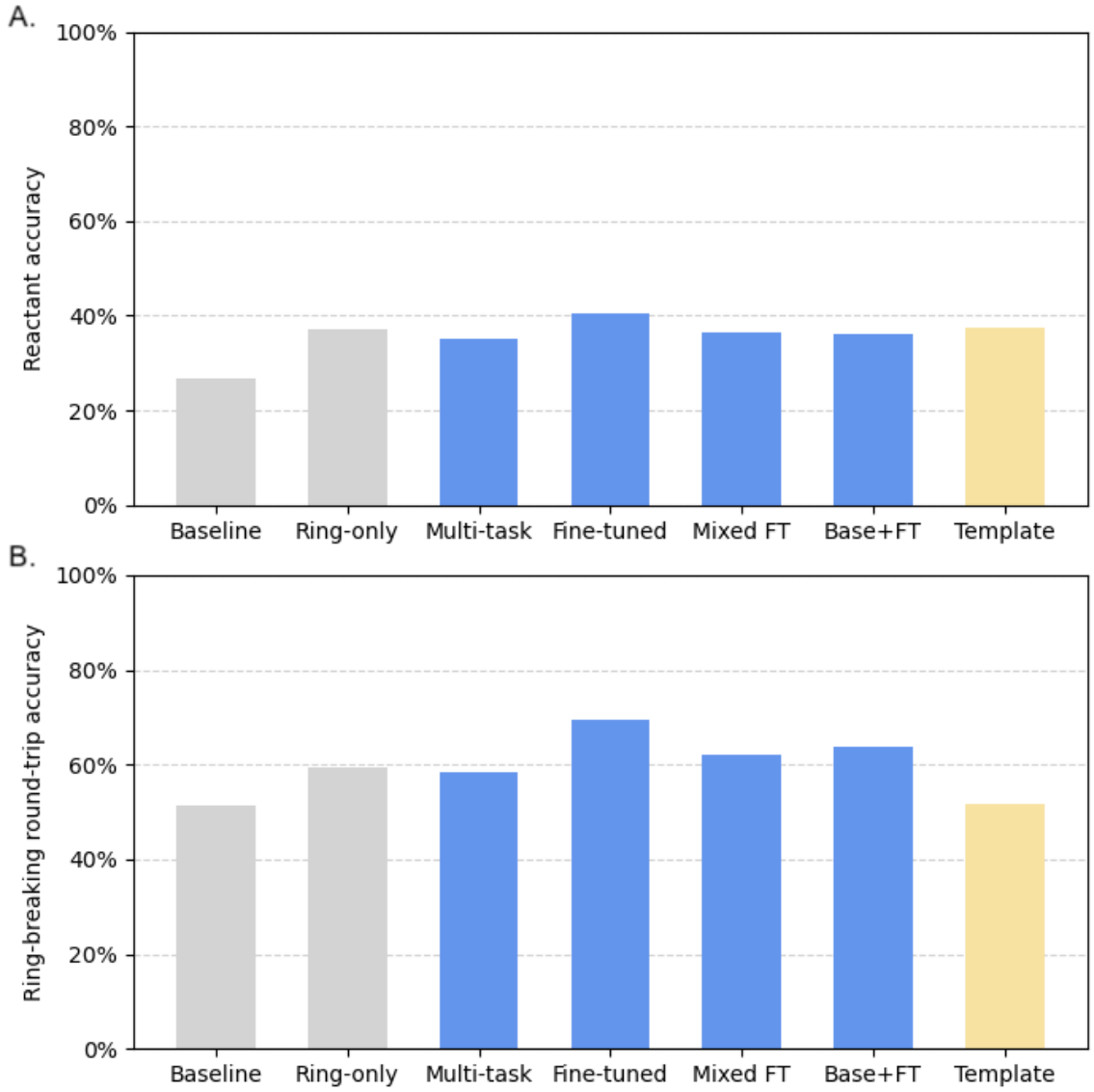


Figure 3.6: Comparison of model performance for ring-breaking disconnections. (A) Top-1 reactant accuracy and (B) top-1 ring-breaking round-trip accuracy are shown for the *Ring* test set. All domain adaptation approaches (blue); *multi-task, fine-tuned, mixed fine-tuned* (Mixed FT) models and ensemble decoding (Base+FT); are compared to the two baselines (grey); *baseline* and *ring-only* model; and the *template-based* (Template) model (yellow). All reported accuracies are from single model training runs due to resource constraints.

Table 3.4: Recall-based metrics assessing all models' performance on the *Ring* test set for top-1/3/5 predictions.

Model	Reactant accuracy			Top-N accuracy		
	top-1	top-3	top-5	top-1	top-3	top-5
Baseline	26.9%	35.9%	38.9%	6.5%	11.0%	12.6%
Ring-only	37.2%	47.5%	50.9%	11.4%	17.9%	19.9%
Multi-task	35.0%	47.1%	50.4%	9.1%	17.3%	19.7%
Fine-tuned	40.5%	51.5%	54.5%	11.1%	19.0%	21.7%
Mixed fine-tuned	36.5%	48.6%	51.6%	10.1%	17.6%	19.8%
Ensemble decoding	36.3%	48.5%	52.7%	7.9%	14.1%	16.1%
Template-based	37.6%	49.2%	53.6%	6.8%	8.8%	9.6%

which is able to achieve reasonable accuracy due to the larger dataset size than that used in previous studies.[174] The same trends are observed for top-3/5 reactant accuracy and top-1/3/5 accuracy, with the top-N accuracy being significantly lower than reactant accuracy for all models due to the difficulty in predicting the exact reagents with many combinations of conditions being viable (Table 3.4).

While the observed improvement for the domain adaptation approaches over *baseline* isn't as high as reported in previous studies (13.6% increase in top-1 reactant accuracy for *fine-tuned* model here vs 27.0% for carbohydrate reactions and 28.6% for Heck reactions),[174, 176] there are two key aspects to note. Firstly, the mentioned studies used transfer learning for forward reaction prediction, not retrosynthesis, which is considered to be a much easier task, only having one "correct" answer. Moreover, heterocycle formations are a much larger and more diverse class of reactions than Heck reactions or even carbohydrate reactions, making it more difficult for the model to learn all the different reactivities.

We then compared our domain adaptation approaches to previously designed approaches for heterocycle retrosynthesis, based on the Ring Breaker model.[272] When considering reactant accuracy, the *fine-tuned* model performs better than the *template-based* model trained only on ring formation data while all other domain

Table 3.5: Precision-based metrics assessing all models' performance on the *Ring* test set for top-1/3/5 predictions.

Model	Round-trip accuracy			Coverage			Ring-breaking r-t accuracy		
	top-1	top-3	top-5	top-1	top-3	top-5	top-1	top-3	top-5
Baseline	75.9%	73.7%	72.0%	75.9%	84.7%	86.9%	51.4%	49.7%	48.2%
Ring-only	63.5%	59.6%	56.6%	63.5%	73.2%	76.4%	59.5%	56.6%	53.9%
Multi-task	77.9%	76.5%	75.1%	77.9%	86.3%	88.2%	58.5%	57.6%	56.5%
Fine-tuned	72.1%	70.4%	69.1%	72.1%	81.7%	84.4%	69.5%	68.1%	67.0%
Mixed fine-tuned	74.6%	73.3%	71.8%	74.6%	84.4%	86.9%	62.1%	61.3%	59.9%
Ensemble decoding	71.8%	69.7%	68.0%	71.8%	82.2%	84.7%	63.9%	62.1%	60.2%
Template-based	53.4%	38.9%	31.8%	53.4%	66.7%	70.1%	51.6%	37.7%	30.8%

adaptation approaches (and the *ring-only* model) are worse than the *template-based* model (Figure 3.6A). While the top-N accuracy of the *template-based* model is lower than for any of the benchmarked domain adaptation approaches, this can be attributed to the applied templates not containing reagents (by definition). However, when considering precision-based metrics (Table 3.5), the *template-based* model performs significantly worse than the other approaches, with the round-trip accuracy and ring-breaking round-trip accuracy sharply decreasing for top-3 and top-5 predictions. This rapid decrease (from 53.4% top-1 to 31.8% top-5 round-trip accuracy) is in contrast to the Transformer-based models which maintain high round-trip accuracy from top-1 to top-5 (e.g., 74.6% vs 71.8% for the *mixed fine-tuned* model). The low round-trip accuracy could indicate the *template-based* model's inability to apply multiple templates to one molecule. Hence, it is likely that the Transformer-based models learn a wider range of chemistry than the *template-based* model, which is limited in diversity when it comes to disconnection strategies. However, it is important to note that the forward reaction prediction model used for calculating round-trip accuracy is also a Transformer model and is trained on the same reaction data (but with reversed labels). This could be biasing the metric towards the our domain adaptation approaches and mean that the difference in round-trip accuracy between them and

the *template-based* model is not as significant as it seems. A more objective way of calculating metrics such as round-trip accuracy could be to use a different model to predict reaction viability instead of the forward reaction prediction model, however, we were not able to train such a model for this work due to lack of negative reaction data.

Interestingly, even though each of the domain adaptation approaches increases the ring-breaking round-trip accuracy by at least 7% when compared to the *baseline* model, the same trend is not observed when considering just the round-trip accuracy or coverage of the predictions (Table 3.5). For example, for the *mixed fine-tuned* model, the top-1 ring-breaking round-trip accuracy increases by over 10%, while the top-1 round-trip accuracy decreases by 1%. The same trend can be observed for all other approaches apart from the *multi-task* model, where the round-trip accuracy increases but not as much as the ring-breaking round-trip accuracy. This indicates that the main improvement between the various models trained using transfer learning and the *baseline* model is in the type of disconnection suggested, i.e. ring-breaking versus more common reaction types, and not in turning chemically invalid disconnections into valid ones. It also suggests that while the molecules in the *Ring* test set were synthesized using ring formation reactions, there are other chemically viable disconnections available.

Indeed, comparing the predictions of the *baseline* and *mixed fine-tuned* model revealed that the former often suggested more common reaction types, such as functional group interconversions (FGIs) or protection/deprotections, instead of the ground-truth heterocycle formation predicted by the *mixed fine-tuned* (Figure 3.7). For instance, in example 3.7A, the *mixed fine-tuned* model correctly identifies a click reaction to generate the triazole from two fragments of similar complexity. In contrast, the *baseline* model only suggests a more trivial N-alkylation reaction. Similarly, for 3.7B. the *mixed fine-tuned* model suggests a condensation reaction to form the central

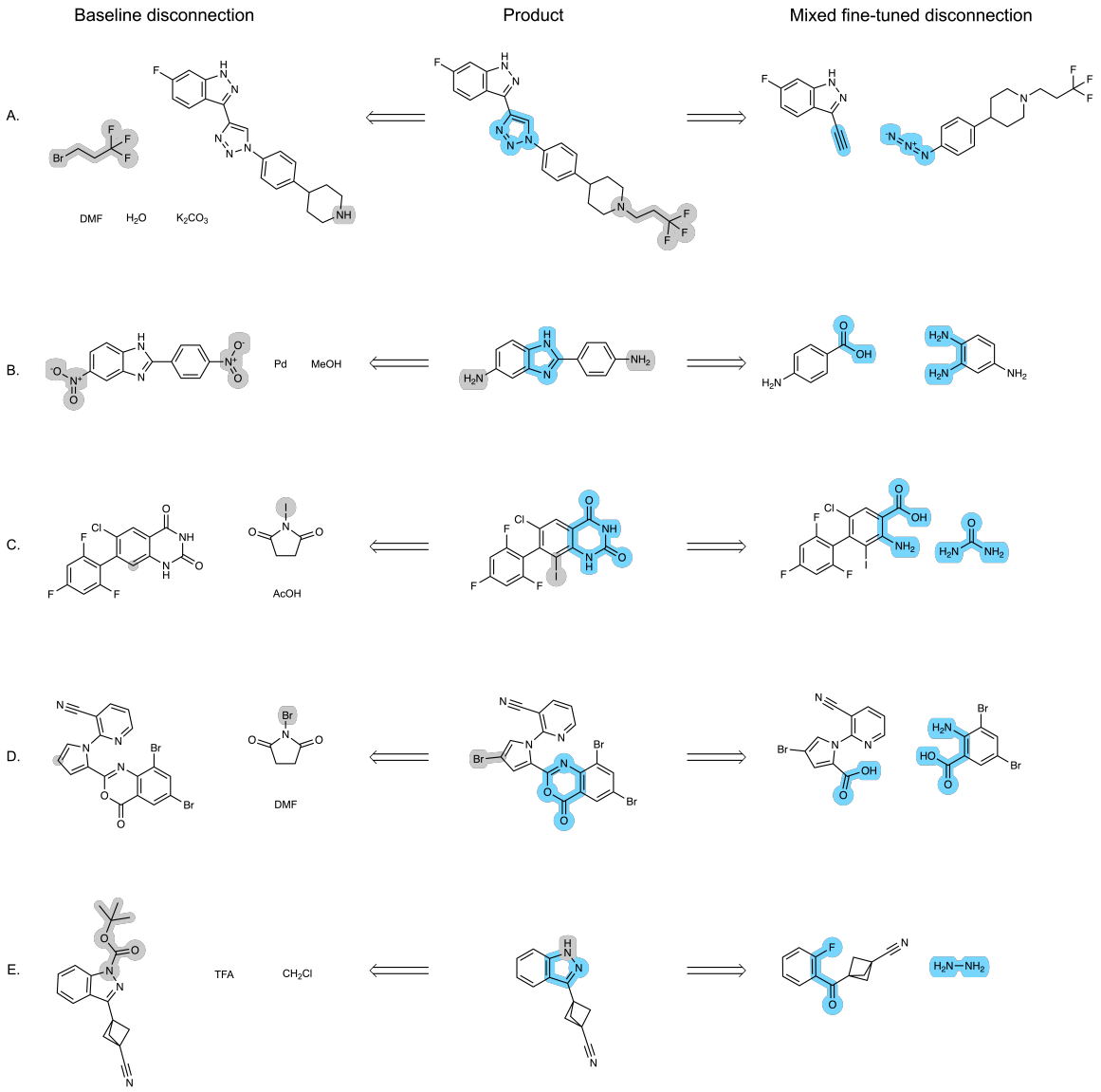


Figure 3.7: Example top-1 predictions of the *mixed fine-tuned* and *baseline* models for *Ring* test set molecules. All predicted reactants and reagents are shown. For all the examples shown, the *mixed fine-tuned* prediction was accurate, whilst the *baseline* prediction was valid but not ring breaking. The disconnections suggested by the *mixed fine-tuned* model are highlighted in blue, while the disconnections suggested by the *baseline* model are highlighted in grey.

benzimidazole ring, while the *baseline* model suggests a functional group interconversion, which would be more suitable earlier in the synthetic route. In 3.7C. and 3.7D. the *baseline* model predicts simple halogenation reactions rather than ring disconnections. Interestingly, although the *mixed fine-tuned* model's prediction is accurate for 3.7D., it was not counted as round-trip accurate due to the forward model predicting a condensation reaction with both the carboxylic acid and the nitro group instead of just a single condensation with the former. This highlights a limitation of metrics based on round-trip accuracy, where the model's prediction is only assessed by another model that is not 100% accurate instead of comparing the prediction to those reported in the literature or assessed by skilled organic chemists. Finally, in 3.7E. the *mixed fine-tuned* model correctly predicts the disconnection of indazole, while the *baseline* model suggests a Boc protection of the indazole nitrogen without simplifying the molecule. While the ability of the model to suggest protection reactions is notable, as they are crucial parts of synthetic routes, this specific protection is unnecessary and might lead the model to predict a cycle of protection/deprotection reactions, preventing further disconnections of the molecule.

3.3.1.2 Performance for other reaction classes

When tested on the *General* test set, the domain adaptation approaches exhibit almost the opposite trend to the performance on the *Ring* test set (Figure 3.8, Table 3.6 and Table 3.7). Performance of the *fine-tuned* model drastically decreases compared to the *baseline* model, with the top-1 reactant accuracy dropping from 26.4% to 11.4% and top-1 round-trip accuracy from 87.4% to 52.6% (Figure 3.8). The performance of the *ring-only* baseline is even poorer, with only 2.0% top-1 reactant accuracy and 21.8% top-1 round-trip accuracy. Meanwhile, the metrics for the *mixed fine-tuned* and *multi-task* model only change marginally, dropping by at most 2%. Ensemble decoding falls in between, with a top-1 reactant accuracy of 22.7% and top-1 round-

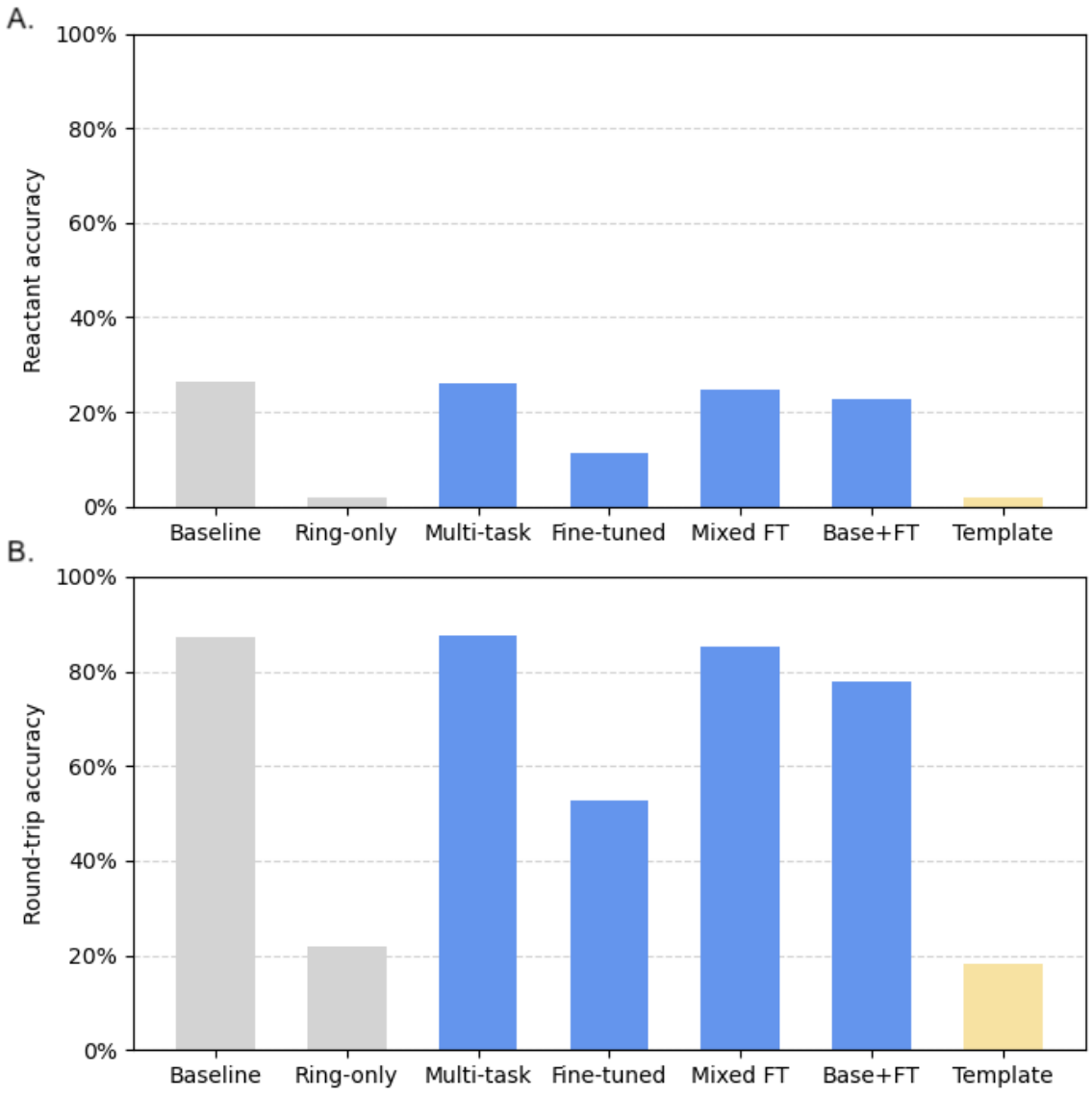


Figure 3.8: Comparison of model performance for all reaction classes. (A) Top-1 reactant accuracy and (B) top-1 round-trip accuracy are shown for the *General* test set. All domain adaptation approaches (blue); *multi-task, fine-tuned, mixed fine-tuned* (Mixed FT) models and ensemble decoding (Base+FT); are compared to the two baselines (grey); *baseline* and *ring-only* model; and the *template-based* (Template) model (yellow).

Table 3.6: Recall-based metrics assessing all models' performance on the *General* test set for top-1/3/5 predictions.

Model	Reactant accuracy			Top-N accuracy		
	top-1	top-3	top-5	top-1	top-3	top-5
Baseline	26.4%	34.3%	36.8%	14.5%	21.5%	23.7%
Ring-only	2.0%	2.7%	3.0%	0.4%	0.7%	0.8%
Multi-task	26.1%	33.9%	36.5%	13.6%	20.6%	22.7%
Fine-tuned	11.4%	16.1%	17.8%	2.7%	4.6%	5.3%
Mixed fine-tuned	24.8%	32.7%	35.2%	12.5%	19.1%	21.0%
Ensemble decoding	22.7%	30.1%	32.3%	9.3%	14.2%	15.6%
Template-based	2.1%	2.4%	2.4%	0.5%	0.6%	0.6%

trip accuracy of 77.9%. The same trends are observed when considering the metrics for top-3/5 predictions (Table 3.6 and Table 3.7).

The *template-based* model performs similarly to the *ring-only* model and much worse than any of the domain adaptation approaches, with top-1 reactant accuracy of 2.1% and top-1 round-trip accuracy of 18.1% (Figure 3.8). With the training data for this model including only ring-forming templates, such low performance is to be expected. Interestingly, when considering top-3 and top-5 metrics, the increase in reactant accuracy (Table 3.6) and coverage (Table 3.7) are lower for the *template-based* model when compared to the *ring-only* model and the decrease in round-trip accuracy (Table 3.7) is higher. This is similar to the trends observed when testing on the *Ring* dataset and supports the conclusion that the sequence-to-sequence models are capable of providing more diverse disconnections.

While the low accuracy of the *ring-only* and *template-based* models can be easily explained by the lack of non-ring forming reactions in their training sets, the drop in performance observed with the *fine-tuned* model can most likely be attributed to catastrophic forgetting - the tendency of neural networks to forget previously learned information when trained on new data. This drop can be disregarded if the models are only intended for one-step ring disconnection. However, it becomes problematic

Table 3.7: Precision-based metrics assessing all models' performance on the *General* test set for top-1/3/5 predictions.

Model	Round-trip accuracy			Coverage		
	top-1	top-3	top-5	top-1	top-3	top-5
Baseline	87.4%	85.4%	83.7%	87.4%	93.4%	94.7%
Ring-only	21.8%	20.0%	18.9%	21.8%	29.7%	33.5%
Multi-task	87.5%	86.0%	84.5%	87.5%	93.4%	94.7%
Fine-tuned	52.6%	49.7%	48.4%	52.6%	67.1%	71.7%
Mixed fine-tuned	85.4%	83.2%	81.3%	85.4%	92.2%	93.7%
Ensemble decoding	77.9%	75.7%	74.1%	77.9%	87.5%	89.9%
Template-based	18.1%	13.2%	10.6%	18.1%	24.7%	27.2%

for multi-step retrosynthesis as the *fine-tuned*, *ring-only* or *template-based* models will not be able to disconnect the linear intermediates obtained after disconnecting the ring. In that case, either the *mixed fine-tuned* or *multi-task* model would be more suitable.

3.3.1.3 Suitability for use in multi-step retrosynthesis tools

Overall, both multi-task learning and mixed fine-tuning show improved performance for ring-breaking disconnections while retaining the ability to predict other reaction classes. While the *fine-tuned* model performs best for heterocycle disconnections, it is not suitable for multi-step retrosynthesis due to catastrophic forgetting. Ensemble decoding ranks in the middle, not being as good at ring disconnections as the *fine-tuned* model, but also performing worse for other reaction classes than the *mixed fine-tuned* model. While both the *ring-only* and the *template-based* models perform better for ring disconnections than all domain adaptation approaches but fine-tuning, their complete inability to predict other reaction classes prevents them from being used for multi-step retrosynthesis on their own.

When considering time and computational resources, fine-tuning and mixed fine-tuning are the two most optimal domain adaptation approaches. With the assumption

that a pre-trained model is available, the fine-tuning time is ∼40 times shorter than the training time for multi-task learning (∼40 minutes vs ∼28 hours on 1 x NVIDIA Tesla V100 GPU in our case). While the same training time would be needed for ensemble decoding as for fine-tuning or mixed fine-tuning, the inference time is doubled due to two models being employed. While this difference is not as impactful when making predictions at scales of hundreds of molecules (such as might be the case in one academic lab), it becomes significant in higher throughput scenarios.

With all the above considerations in mind, mixed fine-tuning appears preferable for use in multi-step retrosynthesis tools due to its good performance for both ring disconnections and other reaction classes and significantly shorter training time compared to multi-task learning (which is especially significant if planning to frequently retrain the model as new data becomes available). Due to this, we perform all further experiments and comparisons with the *mixed fine-tuned* model, as the most versatile and best performing one.

3.3.2 Further fine-tuning on recent heterocycle formations

While the *mixed fine-tuned* model has shown good performance on the *Ring* test set, many reactions present there are based on common, textbook chemistry and would be obvious disconnection ideas for a trained synthetic chemist. To evaluate whether the *mixed fine-tuned* model could predict newer reactions and extrapolate to unknown disconnections, we extracted 1.5k heterocycle ring-forming reactions from 47 papers published in 2022 detailing new heterocycle formation methodologies (here referred to as *Recent* dataset). While the model was, unsurprisingly, unable to predict the exact reported reactions, it provided ring-breaking round-trip accurate top-1 predictions for 30.4% of the molecules. This indicates that while the reported reactions are new, potentially more efficient or greener routes than those reported already, many of the heterocycles formed were already synthetically accessible (Figure 3.9A). Inter-

estingly, the routes suggested by our model often resembled the ground truth (3.9Ai.-iii.). For example, both the *mixed fine-tuned* model and literature suggested the same Friedländer synthesis for quinoline (Figure 3.9Ai.). In the literature synthesis, there is an additional oxime intermediate; however, the *mixed fine-tuned* model's prediction follows the direct approach previously taken for trifluoromethane-substituted quinolines by Jiang *et al.*[282] On the other hand, the reaction predicted in Figure 3.9Aii, while at the same site as the literature disconnection, does not have similar examples in the literature and does not look chemically viable. This is a case where the model prediction might serve as an idea generator, but probably couldn't be executed directly.

Although the *mixed fine-tuned* model found valid ring-breaking disconnections for almost a third of the molecules in the *Recent* test set, when compared to the top-1 ring-breaking round-trip accuracy on the *Ring* test set, this proportion is lower by 30%. Therefore, this indicates that the *Recent* test set includes a higher number of heterocycles unknown to our model and therefore considered synthetically inaccessible. If the model was trained on those new heterocycle formations, it could potentially explore a new region of the chemical space. To address this, we further trained the *mixed fine-tuned* model using the *Recent* dataset. This updated *2022 fine-tuned* model was trained on the three datasets: *General*, *Ring* and *Recent* for another 6,000 steps starting from the *mixed fine-tuned* model, with a dataset weight ration of 4:4:1 (Figure 3.9B). The top-1 accuracy of this *2022 fine-tuned* model is shown in Figure 3.9C. This updated *2022 fine-tuned* model exhibited only a slight decrease in accuracy on the *General* and *Ring* test sets while showing an increased top-1 reactant accuracy on the *Recent* test set (89.9%). This illustrates that the model can be fine-tuned to incorporate new reaction data without significantly compromising performance on previously learned tasks. While we used a small dataset of heterocycle formations here, this approach could be applied to a larger dataset or

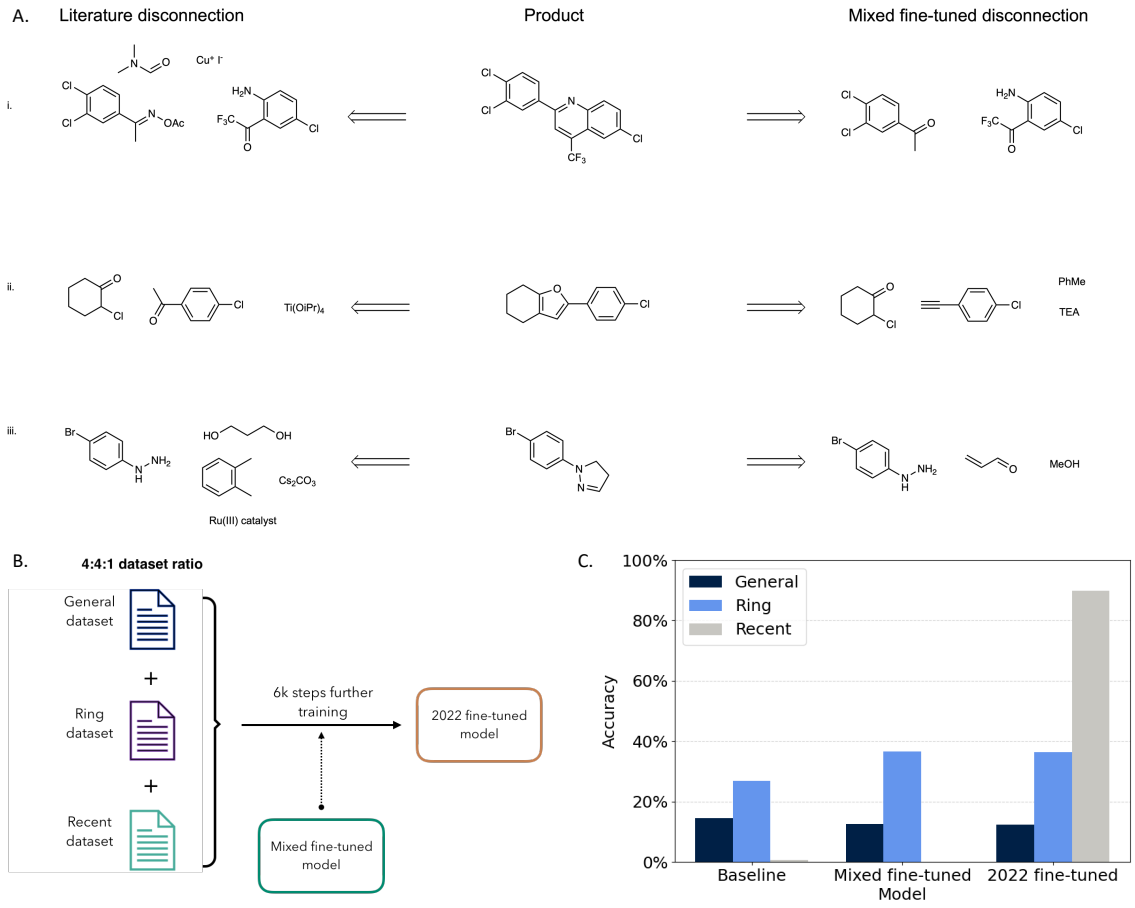


Figure 3.9: Recent reaction prediction. (A) Example valid predictions of the *mixed fine-tuned* model on the *Recent* test set. All predicted reactants and reagents are shown. (B) The further fine-tuning approach: the *mixed fine-tuned* model is further trained on all three datasets. (C) Top-1 accuracy for the *baseline*, *mixed fine-tuned* and *further fine-tuned* model on *General*, *Ring* and *Recent* test sets. Reactant accuracy is reported for the *Ring* and *Recent* test sets.

reaction data for different reaction classes of interest.

3.3.3 Multi-step case studies

The previous sections demonstrated the *mixed fine-tuned* model's performance on single-step retrosynthesis tasks, however, in most real-life scenarios, multi-step retrosynthesis prediction is much more applicable than single-step retrosynthesis. Therefore, to assess the practical use of the *mixed fine-tuned* model in synthesis planning for drug-like targets, we constructed a multi-step retrosynthesis prediction tool using neural-guided A* Search, based on the algorithm used in Retro*.[245] Two drug-like targets were chosen as case studies: **CZS-035** and **ADD** (Figure 3.10), for which syntheses were reported in 2023. The exact reactions employed in these syntheses are therefore absent in our training set, which contains reactions from patents and literature up to 2022. For comparison, we also built an analogous multi-step retrosynthesis tool employing the baseline single-step model, maintaining identical search settings.

The first case study, **CZS-035**, is a ligand for polo-like kinase 4 (PLK4) and a warhead component used to synthesise a therapeutic PROTAC for breast cancer treatment, discovered by Sun *et al.* [283] (Figure 3.10A). Both the *baseline* and *mixed fine-tuned* multi-step models successfully identify retrosynthetic routes for **CZS-035** from purchasable precursors in our stock molecule database. Both models accurately reproduce the protection of nitrogen with Boc (**A1)** as seen in the literature synthesis.[283] Both models also correctly identify the two S_NAr disconnections used in the literature to reproduce **B1** and **B2**. However, the *mixed fine-tuned* model uniquely identifies the final ring disconnection of pyrazole in **B1** to **C1** and **C2**, which aligns with the literature approach. In contrast, the *baseline* model suggests the more complex and more expensive pyrazolopyrimidine **C3** as the final purchasable precursor. This result showcases the enhanced performance of the *mixed fine-tuned* model for predicting key ring disconnections for multi-step routes, overcoming catastrophic for-

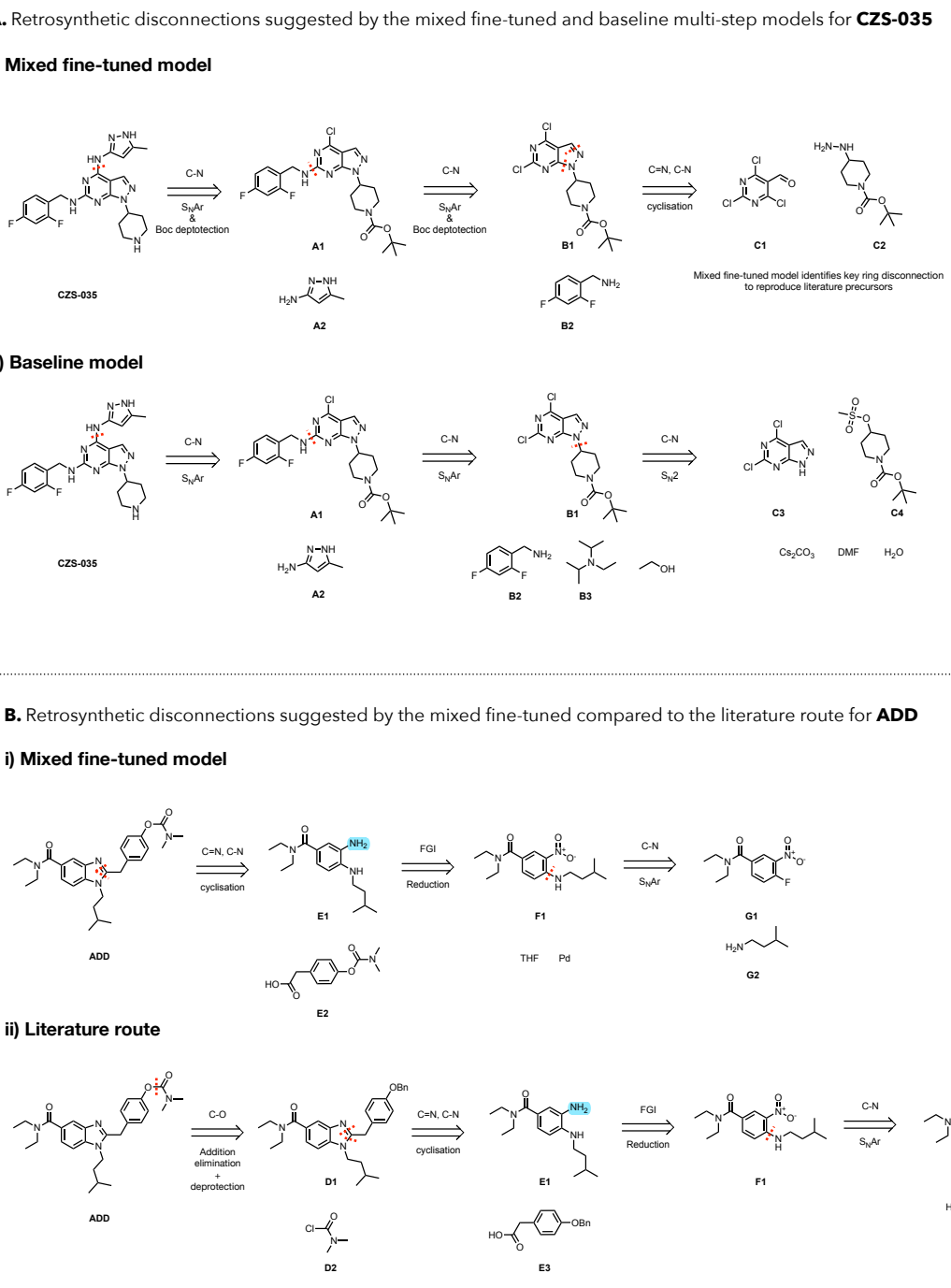


Figure 3.10: Example synthetic routes found by the *mixed fine-tuned* model for molecules of clinical interest. (A) Comparison of the retrosynthetic routes for **CZS-035** predicted by (i) *mixed fine-tuned* and (ii) *baseline* models. (B) The retrosynthetic route for **ADD** (i) predicted by the *mixed fine-tuned* model compared to (ii) the literature route. The *baseline* model failed to predict a complete route for this compound.

getting and correctly identifying all non-ring breaking disconnections of **CZS-035**. We note that the ability of seq2seq models over template-based models to simultaneously suggest protections and S_NAr disconnections in different sites as in **A1** is a unique advantage.

The second case study was **ADD** (compound 15d in ref [284]), a merged human butyrylcholinesterase (hBChE) inhibitor/cannabinoid receptor 2 (hCB2R) ligand and a therapeutic target for preventing learning impairments in Alzheimer's disease (Figure 3.10B).[284] The *baseline* multi-step model failed to identify a synthetic route, while the *mixed fine-tuned* model predicts retrosynthetic disconnections similar to the literature route (Figure 3.10). Reagents were omitted from the literature route to focus on the core synthons. While the *mixed fine-tuned* model deviated by not reproducing the carbamate disconnection of **ADD** to benzyl-protected phenol **D1**, instead using the pre-synthesised phenyl carbamate **E2**, it proposed subsequent disconnections featuring the same cyclisation, reduction, and S_NAr as the literature route to mutually predicted reactants **E1**, **F1**, **G1**, and **G2**. This further reaffirms the improved ring-breaking performance in multi-step retrosynthesis of the *mixed fine-tuned* model, where the *baseline* model failed for the benzoimidazole scaffold in **ADD**.

These results demonstrate the capability of the *mixed fine-tuned* multi-step model in suggesting tractable synthetic routes for newly-discovered, complex drug-like targets containing heterocycles. This highlights its potential as a tool for synthetic chemists, aiding them in designing synthetic routes towards novel heterocycle-containing therapeutics.

3.4 Conclusions

In this chapter, we compared four different transfer learning approaches: fine-tuning, multi-task learning, mixed fine-tuning, and ensemble decoding, with the aim to im-

prove the performance of seq2seq retrosynthesis prediction models for ring-breaking disconnections. We have found that mixed fine-tuning performs best overall, with short training time, top-1 reactant accuracy for ring formations increased by 10% compared to the *baseline* model, and a barely decreased accuracy on other reaction classes. The accuracy for ring formations is comparable to the template-based model we trained based on Ring Breaker; however, the *mixed fine-tuned* model vastly outperforms the template-based model in other reaction classes. While the *fine-tuned* model performs best for ring formations, with top-1 reactant accuracy of 40.5%, its performance significantly drops for other reaction classes due to catastrophic forgetting. This makes it unusable for multi-step retrosynthesis, which requires disconnection of both rings and linear intermediates.

We have also introduced a new metric, the "ring-breaking round-trip accuracy", to assess the performance of the models for ring-breaking disconnections. By comparing the round-trip accuracy and ring-breaking round-trip accuracy of the *baseline* and *mixed fine-tuned* models, we have shown that both models suggested viable disconnections for a similar proportion of molecules. However, the key improvement in the *mixed fine-tuned* model was the type of disconnection suggested. While the *baseline* model suggests common reactions, such as protections/deprotections or functional group interconversions, which were either unnecessary or better suitable earlier in the synthetic route, the *mixed fine-tuned* model favoured ring formation reactions, with 62.1% of disconnections being ring-breaking round-trip accurate.

We then introduced a method for further fine-tuning the model on additional reaction data. By using this further mixed fine-tuning we have substantially improved the model's top-1 reactant accuracy on ring formation reactions published in 2022 from 0% to 89.9% without significantly compromising performance for older ring formation reactions or other reaction classes. While this approach has been applied to a small dataset of less than 1.5k heterocycle formations, it has the potential to be scaled up

for a larger dataset or a different reaction class. The limitation of the established further fine-tuning workflow is the need for access to all the datasets the model has been previously trained on, which might not always be available in the case of proprietary datasets. Future work could focus on exploring continual learning approaches that don't require access to data, for example generative or feature replay.[285]

Finally, we showcased the practical utility of the *mixed fine-tuned* model by using it for multi-step retrosynthesis of two newly-discovered, complex drug-like compounds containing heterocycles. This illustrates how the model can be used to assist synthetic and medicinal chemists, aiding them in designing synthetic routes towards novel heterocycle-containing therapeutics. While the above qualitative analysis indicates that the *mixed fine-tuned* model outperforms the *baseline* model when applied in multi-step retrosynthesis, conducting a quantitative analysis on a larger set of heterocycle-containing molecules is crucial to establish whether the differences in the performance of the single-step models fully translate to the multi-step predictions.

Chapter 4

Exploiting validated synthesis plans with active learning

Recent computational advancements have significantly accelerated the drug discovery process; however, *in silico* design of molecules often results in compounds that, despite possessing the desired properties, are not synthetically feasible. One approach to address this issue involves integrating retrosynthesis prediction tools, such as the one discussed in Chapter 3, into the molecular design workflow to filter out molecules that cannot be synthesised. Nonetheless, the computational cost and time required to perform retrosynthetic analysis on every generated molecule are substantial, highlighting the inefficiencies of this approach. A more effective strategy involves considering synthetic feasibility during the molecule generation phase rather than as a post-hoc filter. Recently, deep learning methods have emerged that incorporate synthetic accessibility metrics directly into the objective function, optimising for both desired properties and feasibility simultaneously. Despite these advances, traditional enumeration-based approaches remain widely used, wherein molecules are generated from predefined molecular fragments and reaction transformations, providing a practical and more reliable means of incorporating synthetic constraints into the drug

design process.

This chapter discusses the design and application of *retro-active*, a framework to generate synthesisable molecules optimised for user-defined scores. *Retro-active* generates molecules based on a provided synthesis route and a provided stock of building blocks (usually a commercial or in-house molecule library) available for enumeration. This approach facilitates the exploitation of the established chemistry in hit-to-lead and lead optimisation stages of drug discovery projects to rapidly access and test new compounds. The use of active learning for building block selection allows *retro-active* to efficiently optimise the product molecule analogues for user-defined objective functions. In validation studies, *retro-active* successfully recovered 59% and 69% of the top-scoring ground truth molecules when employing objective functions relevant to ligand-based and structure-based drug discovery, respectively, while evaluating only approximately 7% of the total chemical space. Furthermore, *retro-active* demonstrated robust performance across various synthesis routes, generating drug-like molecules tailored to a range of optimisation tasks, including multi-parameter optimisation, thereby underscoring its versatility and potential utility in drug discovery applications.

A manuscript detailing the work presented in this chapter was in preparation at the time of submission.

4.1 Introduction

Computational methods such as virtual screening or *de novo* design have emerged as significant contributors to the drug discovery process by substantially accelerating the Design-Make-Test-Analyse (DMTA) cycles and expanding the chemical space of molecules that can be explored (as discussed in Section 1.2).[4, 286] Synthesisability of the *in silico* designed molecules is crucial for progressing them to experimental

validation, however it is often considered as an afterthought or overlooked. Gao *et al.* recently highlighted this issue, noting the low proportion of synthetically feasible compounds produced by *de novo* generative models.[15] Since then, synthetic accessibility has gained increased attention within the field of molecular generation, with a range of strategies employed to address this challenge.[287] These strategies include the incorporation of synthesisability filters and the implementation of biases or constraints within the generative models (and other tools) to favour the production of synthesisable molecules (Figure 4.1). An overview of those general strategies for synthesisable molecule generation, together with their strengths and limitations, has been provided below.

The most straightforward strategy for obtaining synthesisable molecules involves post-hoc filtering (Figure 4.1A), where synthesis routes for previously generated molecules are predicted using a retrosynthesis prediction tool[220–222, 225] or, for a quicker but less reliable filter, a synthesisability score[142, 149, 153] is calculated (as described in Sections 1.5 and 1.3 respectively). However, this approach can prove slow and inefficient, with computational resources being wasted at both the generation and scoring stage, when only ∼30% of the molecules generated by deep learning models are synthesisable.[15] The alternative that is steadily gaining popularity involves integrating a bias towards synthesisability directly into the generative model (Figure 4.1B). This can be achieved either through careful curation of the training set, ensuring only synthetically accessible molecules with desirable properties are present, or by incorporating a synthesisability metric into the model's objective function. Such metrics may include synthesisability scores or simpler heuristics describing molecular complexity, such as SMILES length, the number of stereocenters, or the substructures present. The recent work by Guo *et al.* has shown that, for sufficiently sample-efficient generative models, even a retrosynthesis model can be included in the objective function.[288]

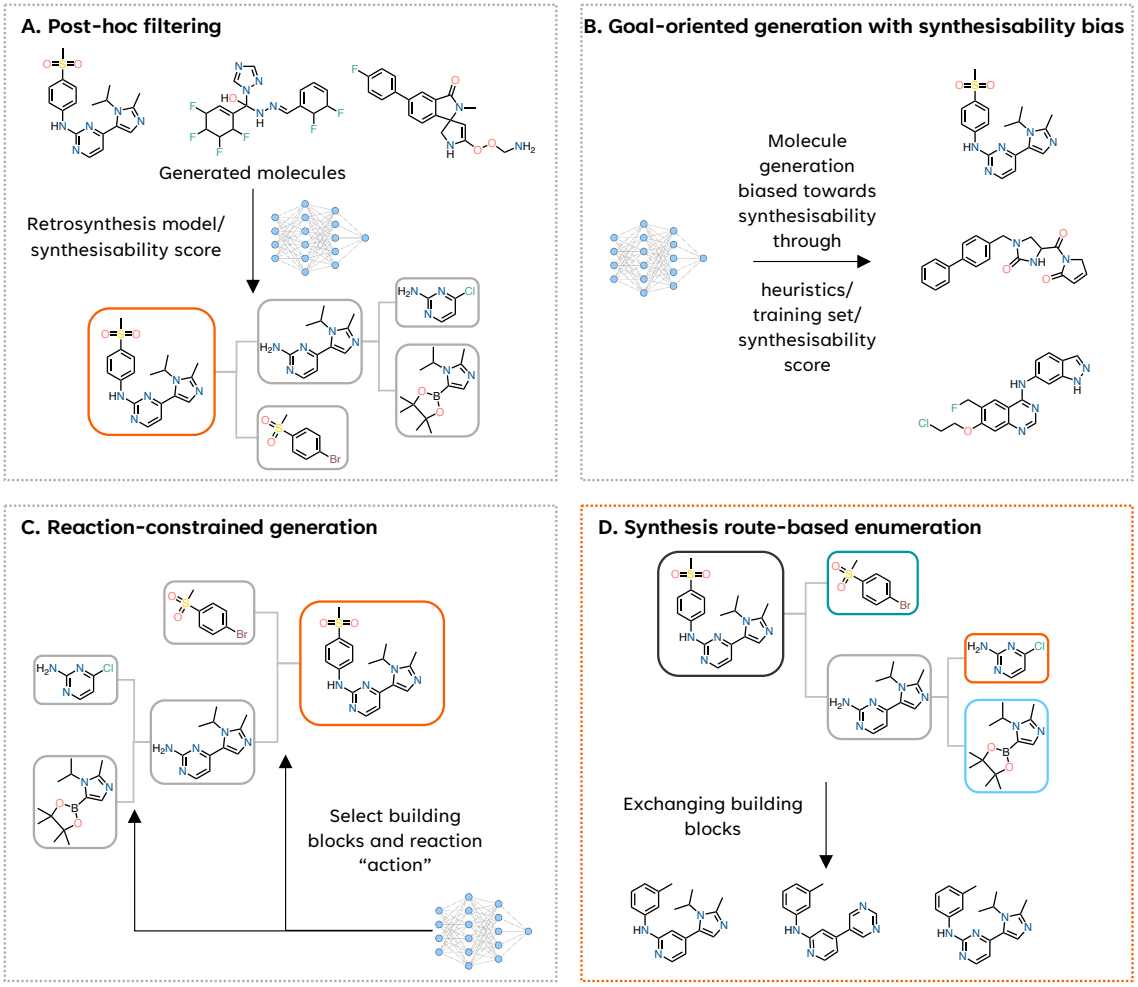


Figure 4.1: Strategies for synthesisable molecule generation. (A) In post-hoc filtering, molecules obtained from a generative model are filtered based on the predictions of a retrosynthesis tool or a synthesisability score. (B) Generative models can also be biased towards synthesisable molecules, either through curation of the training set or by including synthesisability metrics in the objective function. (C) In reaction-constrained generation, molecules are obtained by constructing a synthesis route from bottom up, through selection of building blocks and reaction actions. (D) In route-based enumeration, molecules are constructed based on provided synthesis routes, by exchanging the building blocks and propagating the changes from the bottom up to arrive at product molecule analogues.

Another recently emerging strategy for synthesisable molecule generation is the use of reaction-constrained generative models (Figure 4.1C).[77, 173, 289–297] In this approach product molecules are obtained based on a forward synthesis route, theoretically guaranteeing synthesisability. The route is generated from the bottom up, with building blocks selected by the model and combined through the application of pre-defined reaction templates or a forward reaction prediction model. While this approach shows promise, some concerns remain about the validity of the generated synthesis routes and the breadth of chemical space explored by those models.[287]

A simpler, non-machine learning approach involves enumerating analogues of product molecules based on a predefined synthesis route and available building block stock (Figure 4.1D).[32, 298–300] This method restricts the generated compound space to theoretically synthetically accessible molecules; however, it lacks an inherent mechanism for optimising the molecules against a specific objective function. Instead, desired properties can either be achieved through careful selection of the building blocks or post-hoc filtering of the product molecules with the objective function. However, due to the combinatorial explosion associated with the exhaustive enumeration of all possible building block combinations, scoring the entire possible product space becomes impractical even for relatively simple systems. This is why most work to date has focused on enumerating either one building block at a time or significantly restricting the number of building blocks enumerated at each position.

Active learning (AL) is a commonly employed strategy in drug discovery when exploring prohibitively large chemical spaces and employing expensive objective functions.[58, 261, 264, 300, 301] This method involves an iterative cycle where molecules are scored in small batches, and these scores are used to train a machine learning model that serves as a computationally efficient surrogate for the objective function. The surrogate model is then employed to guide the selection of the next set of molecules to evaluate. Active learning has been previously integrated with route-

based enumeration, for generation of cyclin-dependent kinase 2 inhibitors.[300] In this approach, product molecules were first enumerated using the PathFinder tool, followed by the use of active learning to select the best molecules based on predicted potency. While this approach has improved the computational efficiency of the filtering stage, the explored chemical space was still restricted by the number of product molecules that could be enumerated and stored in memory.

Here we introduce *retro-active*, a novel route-based enumeration framework for molecular generation that employs active learning to identify and prioritise the most promising building blocks for enumeration. Unlike previous approaches that apply active learning at the product molecule level, *retro-active* leverages active learning during the building block acquisition phase, allowing for the efficient exploration of significantly larger chemical spaces. *Retro-active* can optimise product molecules for a range of objective functions, including ligand-based and structure-based metrics, as well as multi-parameter optimisation (MPO).

4.2 *Retro-active*

The *retro-active* tool was developed during my work at Exscientia to generate synthesisable molecules optimised for user-defined objective functions. *Retro-active* designs molecules based on a provided synthesis route, by exchanging the building blocks in the route and enumerating product molecule analogues. It also simultaneously optimises the product molecules for the provided objective function in a computationally efficient way through the use of active learning. A detailed description of the *retro-active* workflow and the implemented building block acquisiton and product enumeration methods are included below.

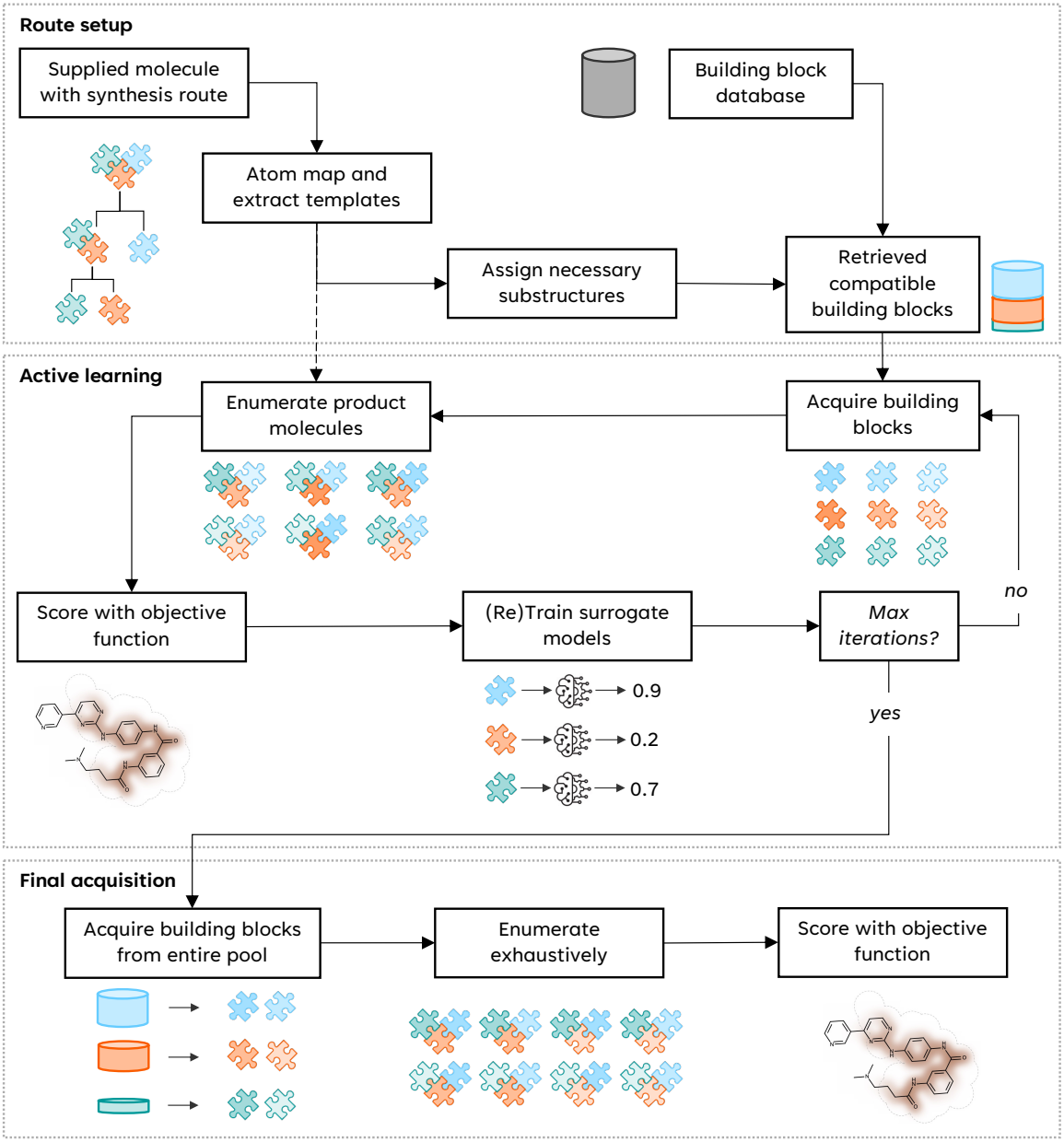


Figure 4.2: Overview of the *retro-active* workflow. The workflow is split into three stages: route setup, active learning and final acquisition. Given a synthesis route and a building block stock as input, *retro-active* returns a pool of product molecules obtained using the given synthesis route but starting from different building blocks.

4.2.1 Overview of workflow

The *retro-active* workflow (Figure 4.2) can be split into three distinct stages:

- **Route setup**, where the synthesis route is prepared for later stages and compatible building block pools are obtained.
- **Active learning**, the exploration-focused stage, where models are trained to predict scores for building blocks in an iterative cycle of building block acquisition, product enumeration, and scoring.
- **Final acquisition**, the exploitation-focused stage, where the trained models are used to select a large batch of highest scoring building blocks that are enumerated to form the final product molecule pool together with molecules scored in the previous stage.

The expected input for *retro-active* is a synthesis route together with a database of available building blocks (a stock of molecules, usually from a commercial catalogue or available in-house). In the first stage, all reactions in the synthesis route are atom-mapped in order to extract reaction templates. Then, for each of the starting materials in the synthesis route, the reaction templates are used to assign the substructures (functional groups) that need to be present in this building block for it to be able to undergo all relevant reactions in the synthesis route. Those assigned necessary substructures are then used to extract all compatible building blocks from the building block database through substructure matching. This step completes the route setup stage.

The active learning stage begins with a random acquisition of building blocks from each compatible building block pool. Then, combinations of those building blocks are enumerated and the new product molecules are obtained after forward enumeration through the prepared synthesis route. Next, each product molecule is scored with the objective function. The obtained scores are used to train a machine learning model

for each of the starting materials in the route, that given a building block predicts the score of a product molecule containing that building block. Those models are then used to acquire the next batch of building blocks based on their predicted scores. This process is repeated a set number of iterations (also referred to later as loops), with the surrogate models being retrained every iteration on all previously scored molecules. The models trained during the last iteration are then used for the final acquisition stage.

In the final acquisition stage the previously trained models are used to acquire a large batch of building blocks, typically much larger than the batches acquired during the active learning iterations. The building blocks are always acquired from the entire compatible building block pool, regardless of the acquisition method chosen for the active learning stage (see section 4.2.2). Then, all combinations of the acquired building blocks are enumerated and the product molecules are obtained after forward enumeration through the prepared synthesis route. Finally, all product molecules are scored with the objective scoring function and returned together with the product molecules scored during the active learning stage.

4.2.2 Acquisition and enumeration strategies

Three building block acquisition strategies and two product enumeration strategies were implemented as part of *retro-active*. A brief description of each strategy is included below.

Acquisition:

- **Random acquisition** - Building blocks are selected through random sampling from the whole building block pool. This strategy is only used for the initial loop of active learning.

- **Greedy acquisition** - Building blocks are selected based on the score predicted

by the surrogate models, with the highest scoring building blocks selected. If selected, building blocks remain in the building block pool and can be selected again in a later AL loop.

- **Explorative acquisition** - Building blocks are selected based on the score predicted by the surrogate models, with the highest scoring building blocks selected. Once selected, building blocks are removed from the building block pool and cannot be selected again in a later AL loop.

Enumeration:

- **Exhaustive enumeration** - A cartesian product of the acquired building block sets, *i.e.* all possible combinations of acquired building blocks are enumerated.

- **Hypercube enumeration** - A Latin hypercube sample of the Cartesian product of the acquired building block sets, *i.e.* a sample from the exhaustively enumerated product space that aims for the most equal coverage of the acquired building blocks in the sampled product molecules. In all experiments $number\,of\,sampled\,product\,molecules \geq number\,of\,acquired\,building\,blocks$, so each building block is guaranteed to contribute to at least one selected product molecule.

4.3 Materials and Methods

In this study, *retro-active* was benchmarked on three synthesis routes towards active molecules for different target proteins, using a variety of objective functions. The methodology associated with those experiments is included below.

4.3.1 Protein and molecule preparation

Protein/ligand systems: In this study, three different protein/ligand systems were used for experiments: AZD5438 bound to CDK2 (PDB: 6GUH),[302] infigratinib bound to FGFR1 (PDB: 3TT0)[303] and lasmiditan bound to 5-HT1F (PDB: 7EXD).[304]

Docking receptors: The relevant protein structures were prepared for docking using Exscientia's internal pipeline based on OpenEye's Spruce tool.[305]

Reference ligands: The reference ligand conformers used to calculate ROCS overlay were taken directly from the relevant co-crystal structures.

Conformer generation: The enumerated product molecules went through a series of preparation and conformer generation steps before docking/ROCS overlay. The preparation step used an internal pipeline based on OpenEye tools to standardise the molecules, enumerate tautomers (using Oequacpac) and enumerate unspecified stereochemistry centres (using Omega). The conformer generation step used OpenEye tool Omega to get a maximum of 5 conformations for the 1M space benchmark experiments, 10 conformations for ROCS overlay in other experiments and 20 conformations for docking in other experiments.

4.3.2 Synthesis route preparation

The synthesis route for AZD5438 was predicted using IBM RXN.[225] For infigratinib and lasmiditan, literature routes were used with minimal modifications.[306, 307] All reactions in the synthesis routes were atom-mapped using rxnmapper[194] and reaction templates were extracted using rdchiral-cpp,[280] setting *no_special_groups* to *True* and *template radius* to 0 (with the exception of the route for lasmiditan where the radius was set to 1).

4.3.3 Building block stock

All building blocks used in this study were extracted from Exscientia's internal version of MCule (accessed Sept. 18, 2023).[308] The building block stock contained 5,613,802 molecules.

4.3.4 Model training

In all experiments the model used was a Random Forest regressor, trained using scikit-learn[309] with the default hyperparameters. The models were trained to predict the value of the relevant objective function for the product molecule, given the Morgan fingerprint of the building block. 2048-bit Morgan fingerprints were used with a radius of 2. These fingerprints were precomputed using RDKit[143] and cached to minimise inference time. During each active learning loop the models were retrained on all molecules scored during this and all previous loops.

4.3.5 Scoring functions

4.3.5.1 ROCS

Both 3D shape similarity and 3D shape and pharmacophore similarity were computed using OpenEye's ROCS tool.[63] The molecules were prepared as described in section 4.3.1. The ROCS score for a molecule was defined as the highest score achieved by any of its conformers: ROCS ShapeTanimoto for 3D shape similarity and ROCS TanimotoCombo for 3D shape and pharmacophore similarity. The ShapeTanimoto score ranges between 0-1, with a higher score indicating more similar molecules. The TanimotoCombo score ranges between 0-2 and is a sum of ShapeTanimoto and ColorTanimoto scores.

4.3.5.2 Docking

Docking was performed using the OpenEye HYBRID tool.[310] The molecules were prepared as described in section 4.3.1. Five poses were docked for each molecule and the docking score for a molecule was defined as the Chemgauss4 score for the lowest scoring pose. For MPO experiments, the negative of the docking score was used and divided by 20.

4.3.5.3 Multi-parameter optimisation

Each multi-parameter optimisation (MPO) objective function was defined as a weighted geometric mean of ROCS/docking score (referred to below as Shape score), Quantitative Estimate of Drug-likeness (QED),[311] and various ADMET properties:

$$\sqrt[i+5]{Shape\ score^i \times QED \times hERGI\ score \times PXR\ score \times MDCK\ score \times logD\ score}$$

The Shape scores were obtained as described in sections 4.3.5.1 and 4.3.5.2 with the remaining scores described below.

QED: The QED score was calculated using RDKit. No transformations were necessary as the score is in the 0-1 range, with a higher score being more desirable.

hERGI: Exscientia's internal regression model was used to predict the pIC50 value for hERG. The pIC50 values were transformed into the hERGI score:

$$hERGI\ score = (pIC50 - 8)/(4 - 8)$$

so that typically obtained pIC50 values corresponded to a hERGI score in the 0-1 range, with the molecules likely to inhibit hERG obtaining a lower hERGI score.

PXR: Exscientia's internal classification model was used to predict pregnane X receptor (PXR) activation. A PXR score value of 0 was assigned for molecules that

were predicted to activate PXR and a value of 1 was assigned for molecules that did not, so that a low score corresponded to molecules likely to upregulate the undesired metabolism.

MDCK: Exscientia's internal regression model was used to predict log_{10}(AB Papp) for Madin-Darby canine kidney (MDCK) cell permeability. The output was transformed into the MDCK score:

$$MDCK\ score = (log_{10}(ABPapp) + 1)/(2 + 1)$$

so that typically obtained permeability values corresponded to a MDCK score in the 0-1 range, with the molecules likely to be permeable obtaining a higher MDCK score.

logD: Exscientia's internal regression model was used to predict logD. The predicted logD value was transformed into the logD score:

$$logD\ score = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-\frac{(logD-1)^2}{2}} & \text{if } logD < 1 \\ 1 & \text{if } 1 \leq logD \leq 3 \\ \frac{1}{\sqrt{2\pi}} e^{-\frac{(logD-3)^2}{2}} & \text{if } 3 < logD \end{cases}$$

so that all values were in the 0-1 range with values closer to the optimal logD range receiving a higher logD score.

4.4 Results

Initial benchmarking of *retro-active* was conducted on the synthesis route towards AZD5438, a potent inhibitor of cyclin-dependent kinase 1,2 and 9 (CDK1/2/9), which is shown in Figure 4.3.[312] Out of the four starting materials in the route, three were considered for replacement (highlighted in green, orange and blue) as only they affected the structure of the final product. For each of those starting materials a pool

of compatible building blocks was extracted from the building block stock, containing 720,920, 150,128 and 720,920 molecules respectively and leading to a product molecule space of over 10^{16} molecules.

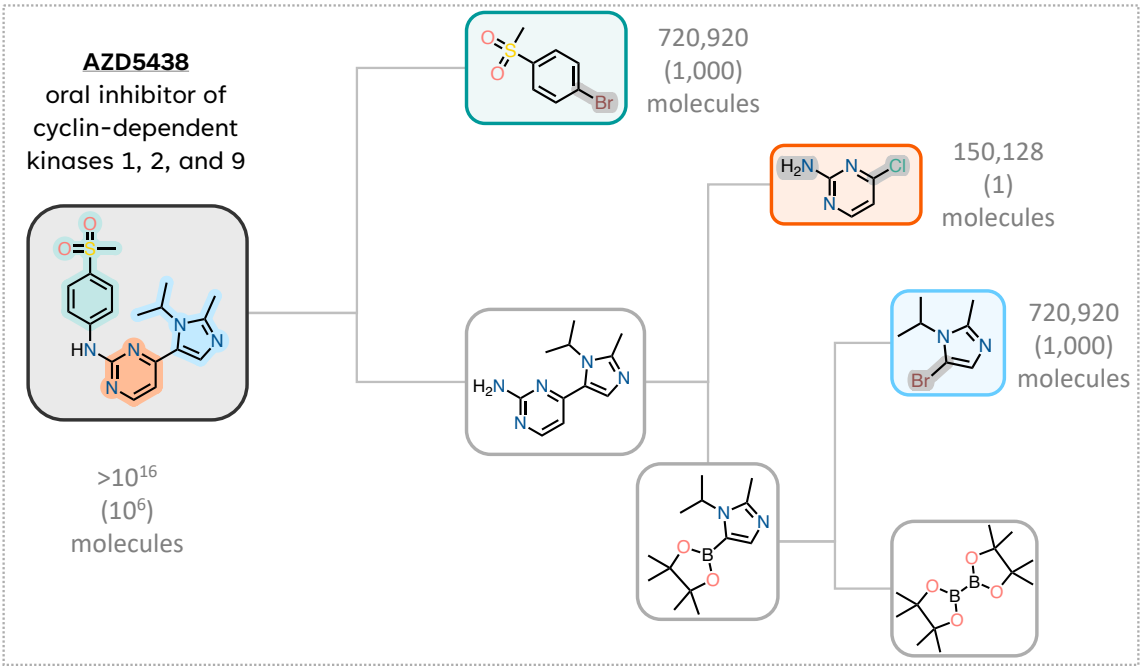


Figure 4.3: Synthesis route towards AZD5438. The green, orange and blue boxes contain building blocks considered for replacement, with the corresponding substructure highlighted in the product molecule in the same colour. The size of the compatible building block pool and the enumerated product space is included next to the molecules; numbers in brackets are for the 1M product space experiments. The compatible building blocks were selected based on containing the substructures highlighted in gray.

4.4.1 Benchmarking *retro-active* on 1M product space

For the first stage of benchmarking, to allow for full enumeration of product space and comparison to ground truth scoring, only two building blocks in the route were considered for replacement: the 5-bromo-1-isopropyl-2-methyl-1H-imidazole (blue) and the 1-bromo-4-(methylsulfonyl)benzene (green). For each of those starting materials, 1,000 building blocks were randomly sampled from their respective compatible building block pools, creating a product space of 1M molecules. This section describes the

results of benchmarking *retro-active* conducted on that 1M product space, focusing on its ability to recover ground-truth top-scoring molecules and the distributions of scores for the molecules it selects. In each of the benchmarking experiments, five loops of active learning, with ∼2,000 molecules scored per loop, were followed by a final acquisition of 250x250 building blocks (corresponding to 62,500 product molecules). This means that in each run of *retro-active* ∼7% of the enumerated product space was scored.

4.4.1.1 Comparison of acquisition and enumeration strategies

To begin with, all combinations of the available building block acquisition methods (explorative and greedy acquisition) and product enumeration methods (hypercube and exhaustive enumeration) were benchmarked with two objective scoring functions: ROCS TanimotoCombo score and the docking score for CDK2 (Figure 4.4). For experiments with hypercube enumeration, 100x100 building blocks were acquired every loop but only 2,000 product molecules were enumerated for scoring. For experiments with exhaustive enumeration, to obtain a comparable number of molecules to score per loop, only 45x45 building blocks were acquired, leading to 2,025 molecules being scored per loop. Six *retro-active* runs were performed for each method combination, with every run starting from a different random state.

Overall, the combination of explorative acquisition and hypercube enumeration performed the best, with, on average, 59% of the top-scoring molecules recovered for ROCS and 69% for docking (Figure 4.4A). The difference in performance between the methods was more pronounced in the case of docking than ROCS. In each case explorative acquisition + hypercube enumeration was followed by greedy acquisition + hypercube enumeration, and explorative acquisition + exhaustive enumeration, with greedy acquisition + exhaustive enumeration performing the worst. However, for ROCS scoring, greedy acquisition + hypercube enumeration and explorative ac-

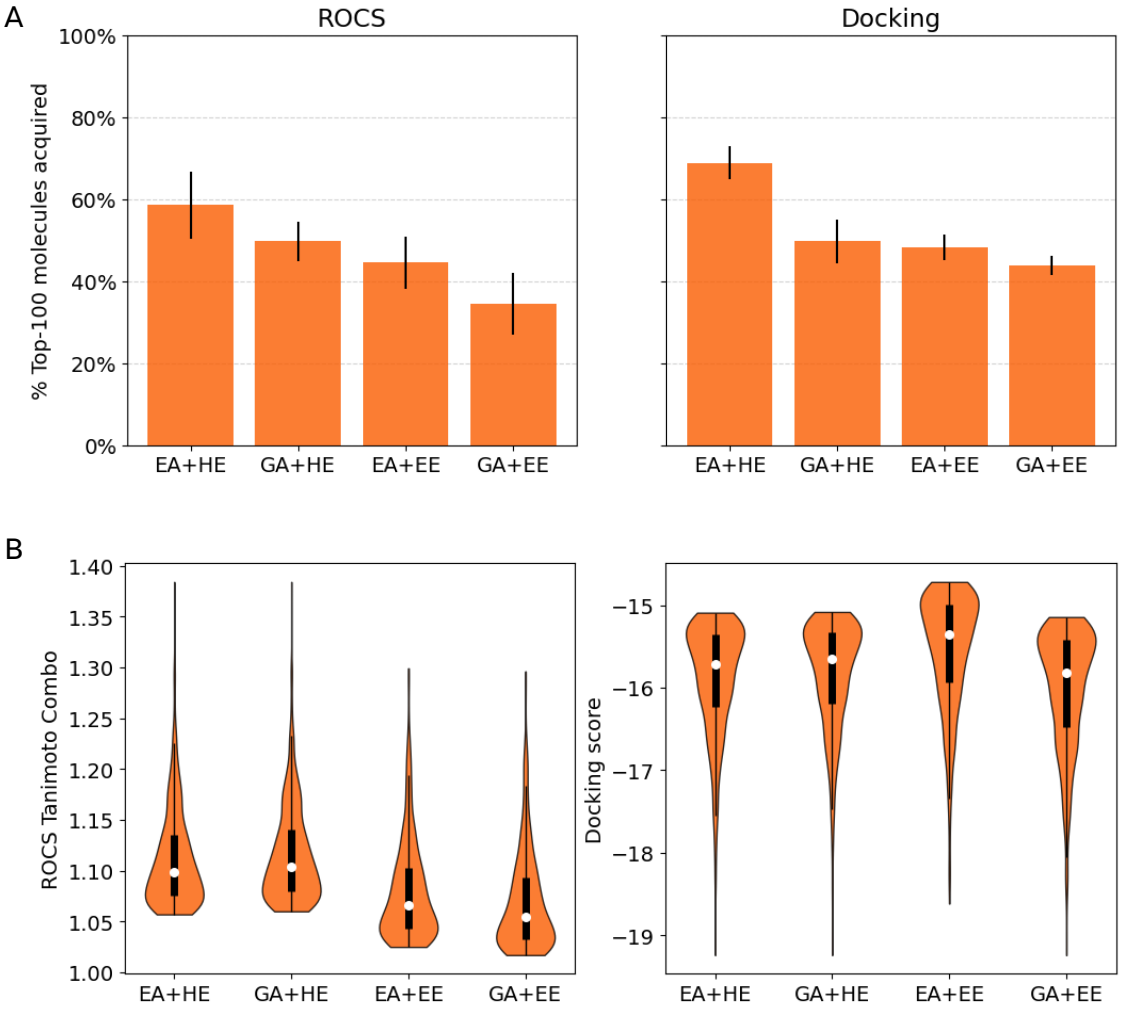


Figure 4.4: Comparison of all combinations of available acquisition and enumeration methods: explorative acquisition (EA), greedy acquisition (GA), hypercube enumeration (HE) and exhaustive enumeration (EE) for scoring with ROCS TanimotoCombo and docking score. (A) Proportion of ground-truth 100 top-scoring molecules acquired by *retro-active*. The average of six runs is shown, with the error bars representing the standard error. (B) Violin plots showing the distribution of scores for the top-1000 molecules selected by *retro-active* for a single randomly selected run.

quisition + exhaustive enumeration were both within error of explorative acquisition + hypercube enumeration. Meanwhile, for docking, explorative acquisition + hypercube enumeration performed significantly better than the other three methods, which all recovered between 44% and 50% of the top scoring molecules. It is noteworthy that the initial random state can significantly affect the outcome of active learning (see Appendix D). For example, for the combination of explorative acquisition and hypercube enumeration scored with ROCS, in four out of the six runs between 68% and 75% of the top scoring molecules were found but in the other two runs this number was twice as low with only 32% and 34% of the top molecules recovered. While less variability was observed in the case of docking, there were still significant differences between runs with 52% of top-scoring molecules recovered in the worst run and 81% in the best run for explorative acquisition + hypercube enumeration. The overall better performance and less variance for docking score vs ROCS is unexpected and might be an artifact of the sub-sampled building block space - repeating the experiments with different sampling of the building block pools could be beneficial to establish if that is the case.

Figure 4.4B shows the distributions of ROCS and docking scores for the top-1000 molecules selected by *retro-active*. For ROCS, the molecules obtained using hypercube enumeration score significantly better than those obtained with exhaustive enumeration, however less difference is observed between the two acquisition methods. Even though explorative acquisition + hypercube enumeration recover 70% of ground-truth top-scoring molecules in the run visualised here, while greedy acquisition + hypercube enumeration only recover 61%, this difference is not reflected in the score distribution. For docking, molecules obtained with explorative acquisition + exhaustive enumeration score significantly worse than the other three methods. This is not surprising, as the active learning run visualised in Figure 4.4B was the worst one out of the six repeats for explorative acquisition + exhaustive enumeration with only 36% of top-

scoring molecules recovered. However, yet again, the score distributions do not reflect that explorative acquisition + hypercube enumeration recovered 14 more top-scoring molecules than the next best performing method. This shows that *retro-active* is still able to obtain good scoring molecules even when it doesn't recover all the ground truth best molecules.

The above results establish that the combination of explorative acquisition and hypercube enumeration leads to best, or at least comparable to other methods, performance for *retro-active*. This is not surprising, as this combination of acquisition and enumeration methods should lead to most different building blocks being acquired and present in molecules scored during active learning, allowing the surrogate models to learn more of the building block chemical space. In light of those results, explorative acquisition and hypercube enumeration were used in all subsequent experiments.

4.4.1.2 Effect of number of AL iterations

While five iterations of active learning were used in the previous experiment following the results from earlier internal studies, we wanted to observe whether the same number of iterations would also be most optimal for *retro-active*. To study this, *retro-active* was used with the parameters described in the previous section for explorative acquisition and hypercube enumeration but with different numbers of active learning iterations, covering each value between zero and five. The distributions of scores for the top-1000 molecules and the proportion of ground truth top-scoring molecules that were recovered after each loop are shown in Figure 4.5 for scoring with ROCS. The greatest improvement in the obtained molecules is seen after the first loop of active learning, with each subsequent loop bringing a smaller improvement than the previous one. The difference between four and five iterations is minimal, both in the distribution of scores for the top-1000 molecules and the proportion of ground truth top-scoring molecules recovered (72% vs 73%). While the score distribution differs

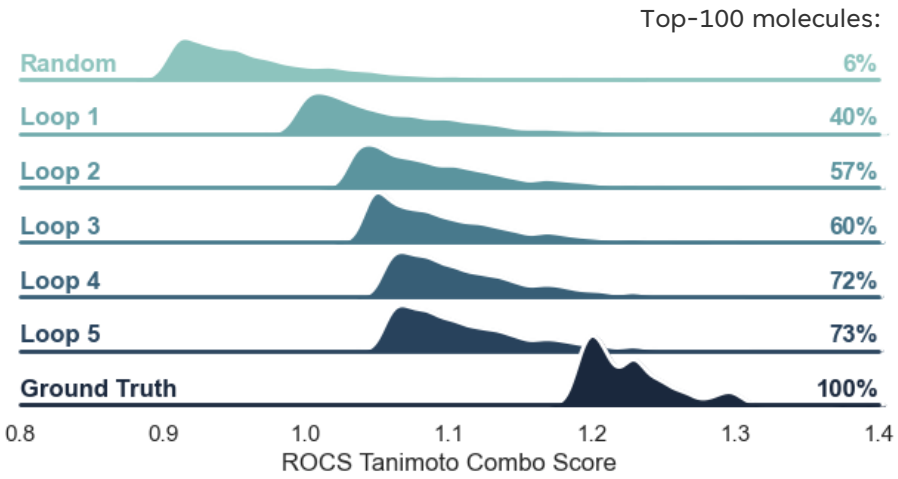


Figure 4.5: Effect of loop number on the performance of *retro-active*. The distributions of ROCS TanimotoCombo score are shown for the top-1000 molecules selected after 0-5 loops of active learning together with the ground truth top-1000 molecules. The proportion of ground truth top-100 molecules recovered is provided on the right.

greatly from the ground truth top-1000 molecules, running *retro-active* for as many as ten loops did not improve the score distribution or allow for recovery of more than 80% of ground truth top-scoring molecules. This indicated that there is a limit to how well the surrogate models can learn the mapping between the building block and the resulting product molecule score, or even how well that mapping can be described, with the molecule score being also affected by building blocks at other positions in the route. In light of those results, and with the same trend observed for docking, it was established that five active learning iterations were optimal for *retro-active*, balancing good performance with reasonable run time.

4.4.1.3 What is *retro-active* missing?

To determine why *retro-active* was not able to retrieve all ground truth top-scoring molecules, we analysed the building blocks present in those molecules and the effect their structural differences had on whether they were acquired by *retro-active* or not. The ground truth top-100 molecules contained 12 different building blocks re-

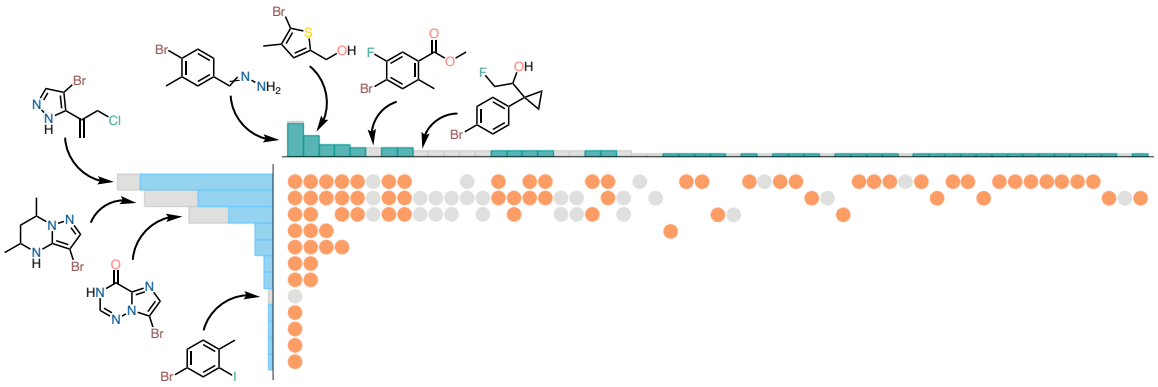


Figure 4.6: Analysis of building blocks present in ground truth top-100 product molecules. The bar charts represent the frequency with which each building block appears in the top-100 products, with the coloured portion indicating how often that block was present in one of the 73 molecules recovered by *retro-active*. Building blocks replacing 5-bromo-1-isopropyl-2-methyl-1H-imidazole (blue) are shown on the left and building blocks replacing 1-bromo-4-(methylsulfonyl)benzene (green) are shown on the top. The top-100 molecules are represented as circles placed at the intersection of the building blocks they contain: orange if the molecule was recovered by *retro-active* and grey if not. Structures of selected building blocks are shown.

placing 5-bromo-1-isopropyl-2-methyl-1H-imidazole and 55 building blocks replacing 1-bromo-4-(methylsulfonyl)benzene (Figure 4.6). The 5-bromo-1-isopropyl-2-methyl-1H-imidazole position was mostly substituted by three building blocks which together covered 83% of the top-product space. On the other hand, the distribution for 1-bromo-4-(methylsulfonyl)benzene replacements was more even, with the majority of building blocks appearing in the top-scoring products only once or twice and the most frequent building block appearing 12 times.

Out of the 12 building block analogues of 5-bromo-1-isopropyl-2-methyl-1H-imidazole, 11 were acquired by *retro-active* in the final acquisition step. All of those 11 building blocks contained a five membered aromatic ring with two nitrogens, either imidazole or pyrazole, mimicking the substructure found in the original building block. Conversely, the building block that was not found by *retro-active* did not contain a heterocyclic core, instead being an iodobenzene derivative. This structural difference could have contributed to the building block not being selected in the final acquisition

stage, especially since it was also not in any of the product molecules scored during the active learning stage.

The top building blocks at the 1-bromo-4-(methylsulfonyl)benzene position were much more poorly recovered. Only 39 out of the 55 building blocks present in the top-scoring products were acquired in the final acquisition step. Unlike for the 5-bromo-1-isopropyl-2-methyl-1H-imidazole replacements, there were no obvious structural differences between the building blocks at this position that were acquired by *retro-active* and those that were not. Moreover, the surrogate model's inability to select those building blocks in the final acquisition stage can't always be attributed to poor exploration of the chemical space surrounding them during the active learning stage, with 7 out of the 16 building blocks present in the scored product molecules.

Overall, *retro-active* recovered all the building blocks appearing at least four times in the top-100 product molecules, but struggled with the less frequently appearing building blocks. With structural differences not accounting for all the building blocks that were not recovered, it is possible that the score distributions for all product molecules containing those building blocks were not sufficiently different from those for other top-scoring building blocks and therefore the surrogate model was not able to prioritise them.

As an aside, even if the surrogate models were able to learn the mapping between the building block structure and the product molecule score perfectly, it would still require at least 55x55=3025 product molecules to be enumerated to recover the top 100. The inefficiency in building block sampling could in theory be mitigated by uneven acquisition of building blocks from each position, for example guided by the building blocks' scores. Even then, at least 12x55=660 product molecules would need to be enumerated. However, for the system explored in this experiment, the distributions of building block scores at the two positions were not sufficiently different to support meaningful asymmetrical sampling, with 243 and 257 building

blocks corresponding to the 5-bromo-1-isopropyl-2-methyl-1H-imidazole and 1-bromo-4-(methylsulfonyl)benzene positions in the overall 500 top-scoring building blocks.

4.4.1.4 Comparison to non-ML based selection

Finally, *Retro-active* was compared to other building block selection methods, ones not based on machine learning models, to establish whether using active learning for building block selection helped *retro-active* obtain better scoring molecules. The benchmarked methods included:

- selection based on only hypercube enumeration - 1,000 product molecules are obtained through hypercube enumeration (so that each building block is sampled exactly once) and scored; building blocks present in the highest scoring product molecules are selected
- selection based on highest tanimoto similarity of the building block fingerprints
- random selection

For each of the non-ML based methods 250x250 building blocks were acquired and enumerated exhaustively to form 62,500 product molecules. Six repeats were performed for random selection and selection based on hypercube enumeration. For *retro-active* the same setup was used as described in section 4.4.1.1 for explorative acquisition + hypercube enumeration.

Figure 4.7 shows the results of benchmarking *retro-active* against these non-ML based methods. For docking, *retro-active* significantly outperforms all other methods, with selection based on only hypercube enumeration coming second and recovering 45.5% ground truth top-scoring molecules. The difference between active learning and selection based on hypercube enumeration is smaller for ROCS, with the latter recovering 51% of the ground truth top-scoring molecules, which is within error of

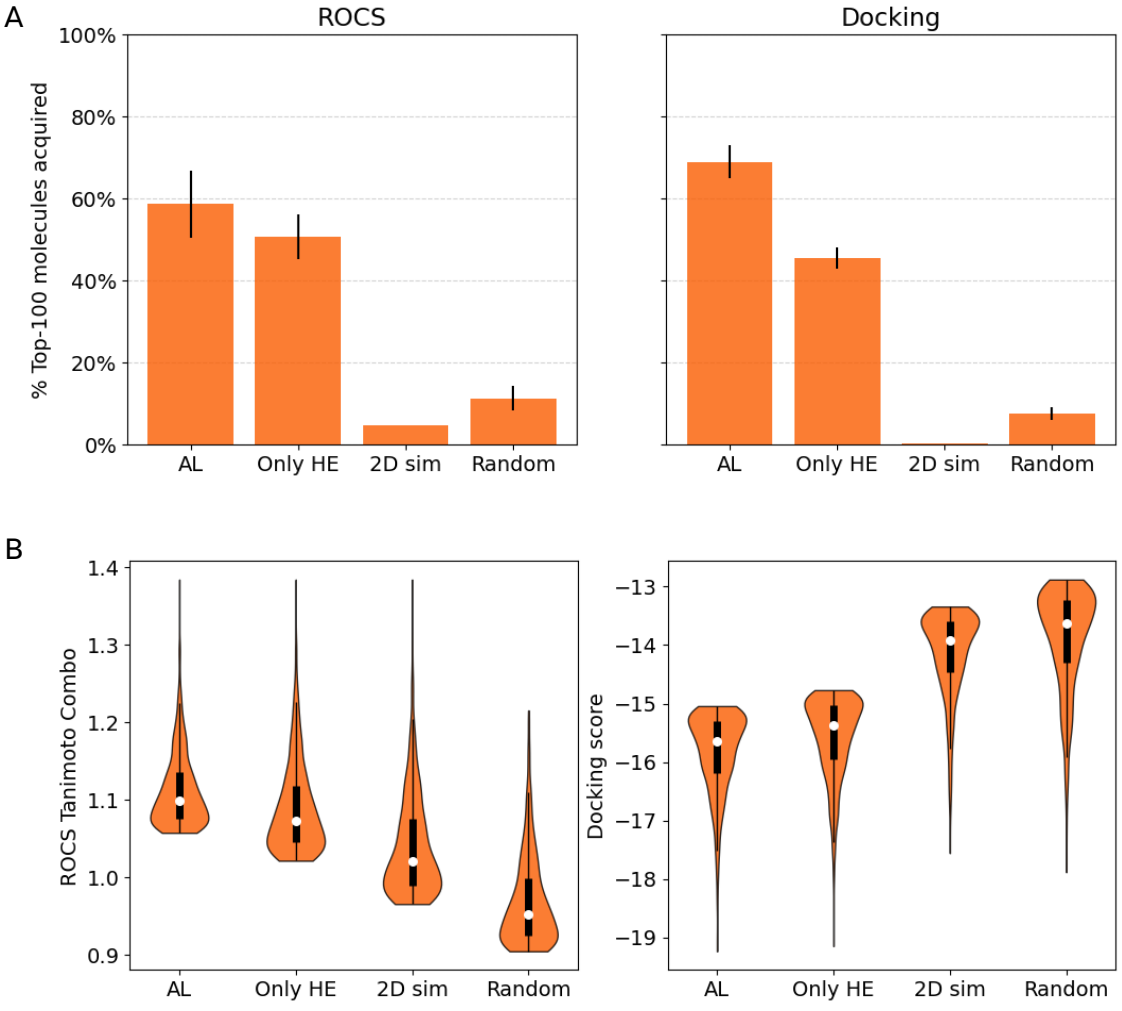


Figure 4.7: Comparison of *retro-active* (AL) to non-ML based building block selection methods: based on scores from one round of hypercube enumeration (Only HE), fingerprint similarity (2D sim) and random selection for scoring with ROCS TanimotoCombo and docking score. (A) Proportion of ground-truth 100 top-scoring molecules acquired by *retro-active*. The average of six runs is shown, with the error bars representing the standard error. (B) Violin plots showing the distribution of scores for the top-1000 molecules selected by each method for a single randomly selected run.

retro-active. However, the distribution of both ROCS and docking scores for the top-1000 molecules is better for *retro-active* than for selection based on only hypercube enumeration. Both random selection and selection based on 2D similarity perform much worse than *retro-active*. Interestingly, in the case of both ROCS and docking, random selection recovers a higher proportion of ground truth top-scoring molecules than selection based on 2D similarity, but has a worse distribution of scores for the top-1000 molecules.

The fact that both *retro-active* and selection based on only hypercube enumeration significantly outperform the other two methods shows the benefit of including product molecule scoring in the building block selection process. Moreover, the better performance of *retro-active* demonstrates that the surrogate models used in active learning can utilise the scores to select even more advantageous building blocks.

4.4.2 Multi-parameter optimisation

Having benchmarked *retro-active* for simple objective scoring functions, we wanted to determine whether it could be used to obtain molecules optimised for multiple objectives. AZD5438 was again used as a starting molecule, with the same synthesis route as used previously (Figure 4.3). However, for the multi-parameter optimisation (MPO) experiments, all three relevant building blocks were considered for replacement, with the entire compatible building block pools used - leading to a possible product space of over 10^{16} molecules.

The MPO objective function used for this experiment was formulated to resemble an example MPO score that would be used in a lead-optimisation stage of a drug discovery project. It included scores describing shape similarity, drug-likeness of the molecules and absorption, distribution, metabolism, excretion and toxicity (ADMET) properties:

- 3D shape similarity as described by the ROCS ShapeTanimoto score

- Quantitative Estimate of Drug-likeness (QED),[311] accounting for properties such as molecular weight, logP, number of hydrogen bond donors and acceptors, etc.
- Inhibition of hERG, an ion channel responsible for the potassium cation flux in the heart muscle cells, which is the most common cause for drug cardiotoxicity
- Activation of pregnane X receptor (PXR), which up-regulates the expression of proteins involved in drug clearance
- logD, a measure of drug lipophilicity
- Cell permeability, calculated for Madin-Darby canine kidney (MDCK) cells

The MPO objective function is a geometric mean of scores for the above properties, each with weight 1 (for a detailed description see section 4.3.5.3).

4.4.2.1 Comparison to other selection methods

Retro-active was first benchmarked against random selection and a set of molecules from ChEMBL.[313] For *retro-active*, explorative acquisition and hypercube enumeration were used per previous experiments. The active learning was run for five iterations, with 20,000x20,000x20,000 building blocks acquired each loop and 40,000 product molecules enumerated and scored. In the final acquisition stage, 60x60x60 building blocks were acquired and enumerated into 216,000 product molecules. This corresponded to 416,000 product molecules being scored in total by *retro-active*. To allocate a similar number of calls to the objective function for random selection, 75x75x75 building blocks were randomly sampled and enumerated exhaustively to obtain 421,875 product molecules. For the ChEMBL molecules, all molecules with associated activities for CDK2 were extracted and scored.

Figure 4.8A shows the distribution of scores for the top-1000 highest scoring molecules (according to the MPO score) obtained by each of those methods. The PXR

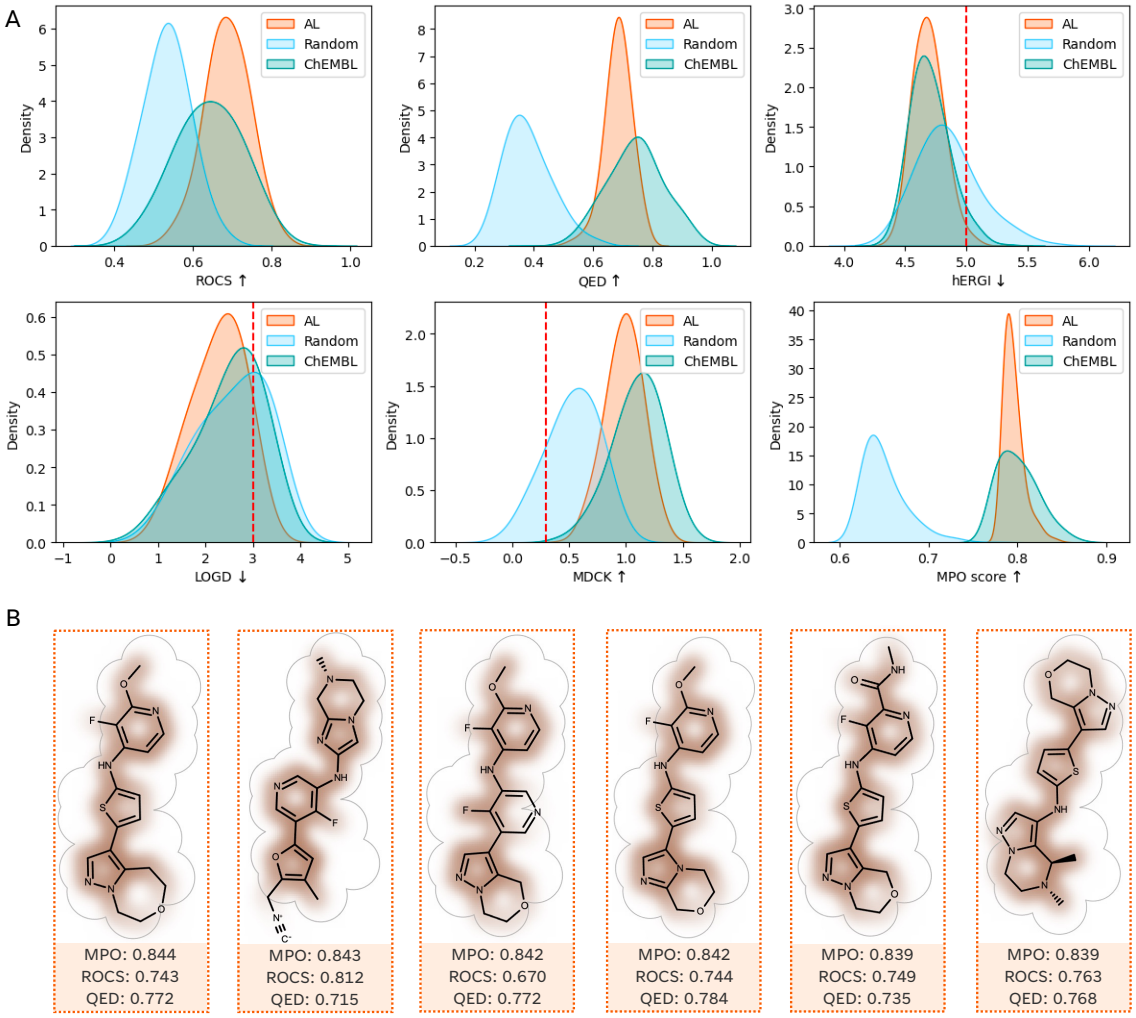


Figure 4.8: Performance of *retro-active* for multi-parameter optimisation. (A) Comparison of score distribution for the 1,000 highest scoring molecules obtained by *retro-active* (AL, in orange), random selection of building blocks (Random, in blue) and molecules from ChEMBL with associated activities for CDK2 (ChEMBL, in green). The shown scores are ROCS ShapeTanimoto, QED, hERG pIC50, logD, MDCK log_{10}(AB Papp) and the MPO score. The PXR activation score distribution is not provided as all top-1000 molecules were predicted to be inactive. The red lines represent commonly acknowledged thresholds for the scores where relevant, while the arrows represent whether a higher (up) or lower (down) score is more desirable. (B) Example molecules from top-10 obtained after active learning. The ROCS shape overlay with AZD5438 is visualised, with the QED, ROCS ShapeTanimoto and MPO scores provided below.

activation score distribution is not provided as all top-1000 molecules were predicted to be inactive. The molecules selected by *retro-active* significantly outscore those obtained through random selection according to the MPO score and all of its other components. For the three ADMET properties, hERGI, logD and MDCK permeability, the majority of molecules returned by *retro-active* are below (or above for MDCK) the commonly acknowledged acceptable threshold: logD<3, hERG pIC50<5 and MDCK log_{10}(AB Papp)>0.3. When compared to the ChEMBL molecules, molecules obtained by *retro-active* have a similar, but sharper, distribution of the MPO score. For some of the MPO components, such as QED, molecules from ChEMBL tend to have a better score, while for others, most notably ROCS, the molecules obtained by *retro-active* score better. However, overall it can be concluded that *retro-active* can optimise molecules for multiple objectives at the same time and return molecules with drug-like properties, resembling those present in ChEMBL.

Examples of top-scoring molecules obtained by *retro-active* are shown in Figure 4.8B. When looking at the top-10 scoring molecules, the main pyrimidine core of AZD5438 is replaced by only two building blocks, thiophene and 3-fluoropyridine. There is more variety in the building blocks selected at the other two positions in the route, although the imidazole-based building block is most often replaced by other imidazole or pyrrazole derivatives. While most of the acquired product molecules look drug-like, a more extensive building block database curation process could be implemented in the future to remove building blocks with potentially unsuitable properties or functionality, such as the isonitrile group present in the second molecule in Figure 4.8B.

Finally, *retro-active* was compared to hypercube enumeration, used as a more difficult benchmark than random selection (see section 4.4.1.4). For hypercube enumeration 720,920 product molecules were first scored and then a further 65x65x65 building blocks present in highest scoring molecules were selected and enumerated

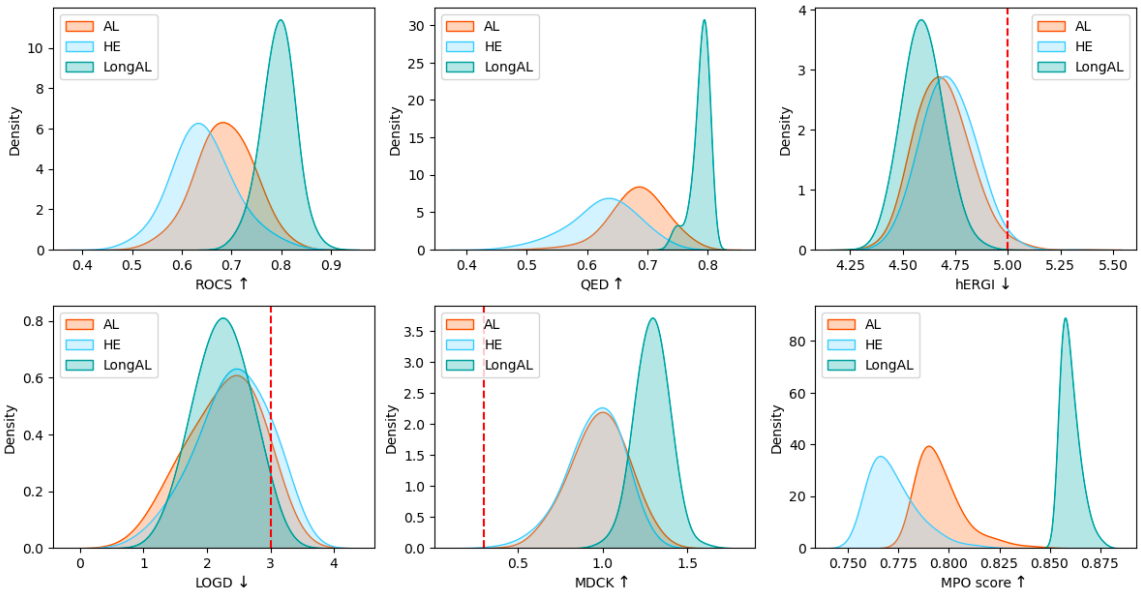


Figure 4.9: Benchmarking multi-parameter optimisation with *retro-active* vs selection based on only hypercube enumeration. Comparison of score distribution for the 1,000 highest scoring molecules obtained after a standard *retro-active* run (AL, in orange), building block selection based on scores from one round of hypercube enumeration (HE, in blue) and a longer *retro-active* run, with the same scoring budget as only hypercube enumeration (LongAL, in green). The shown scores are ROCS ShapeTanimoto, QED, hERG pIC50, logD, MDCK log_{10}(AB Papp) and the MPO score. The PXR activation score distribution is not provided as all top-1000 molecules were predicted to be inactive.

to form 274,625 molecules, with 995,545 molecules scored in total. For *retro-active*, in addition to the standard run discussed above, a longer run was performed with a similar number of molecules being scored when compared to scoring based on only hypercube enumeration. For the long *retro-active* run, twelve iterations of active learning were performed with 20,000x20,000x20,000 building blocks acquired each iteration (using greedy acquisition) and 40,000 product molecules enumerated, with a final acquisition of 80x80x80=512,000 product molecules and 992,000 molecules scored in total. The distribution of scores for the top-1000 highest scoring molecules obtained by each of those methods are shown in Figure 4.9. Even though over twice as many molecules were scored in selection based on hypercube enumeration than in the standard *retro-active* run, the molecules obtained by *retro-active* had, on average,

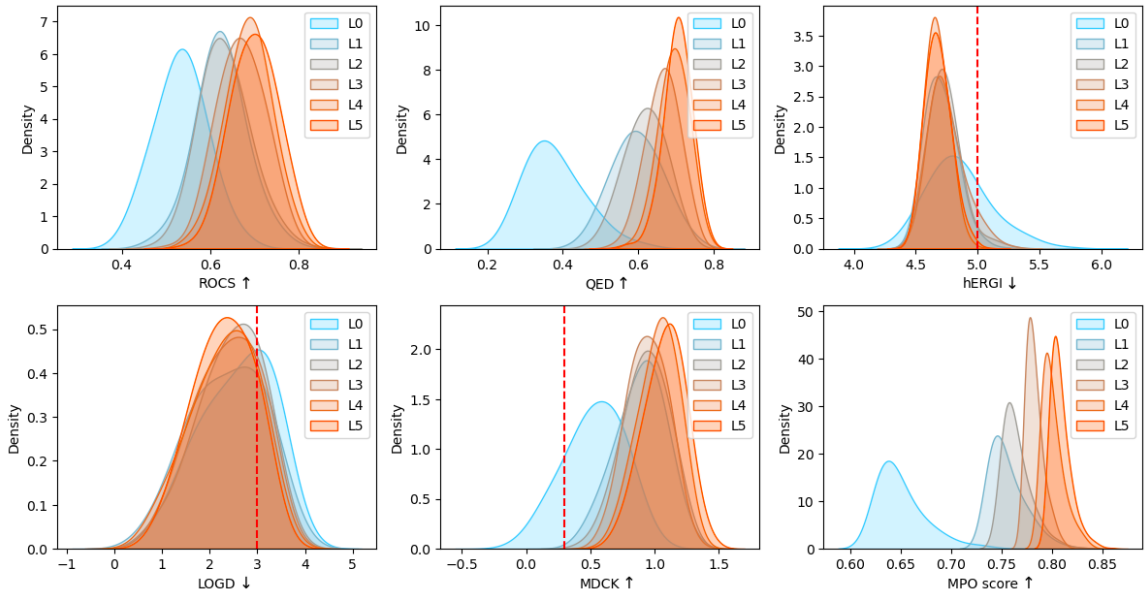


Figure 4.10: Effect of number of active learning iterations on multi-parameter optimisation. The distributions of ROCS ShapeTanimoto, QED, hERG pIC50, logD, MDCK log_{10}(AB Papp) and MPO score are shown for top-1000 molecules obtained from *retro-active* after 0-5 loops of active learning (L0 to L5).

a higher MPO score. The distributions of scores for all the MPO components were also either better for *retro-active* (ROCS ShapeTanimoto and QED) or comparable (ADMET properties) to hypercube enumeration. The molecules obtained after the longer *retro-active* run had even better properties, with the MPO score and all of its components much improved when compared to selection based off only hypercube enumeration. However, the high computational cost of obtaining the scores for all six components of the MPO for almost 1M molecules means that the longer *retro-active* runs are less feasible to use and were not adopted for further experiments.

4.4.2.2 Effect of number of AL iterations

Similarly to the experiments performed for single-objective functions, the effect of the number of active learning iterations on multi-parameter optimisation was also studied. *Retro-active* was run with the same configuration as described in the previous section but with a final acquisition step performed every iteration (as in section 4.4.1.2).

Figure 4.10 shows the distributions of scores for the top-1000 molecules that would be obtained if *retro-active* was run with between zero and five active learning iterations. The MPO score distribution improves with every loop, with the mean score increasing and the distribution becoming sharper. As in previous experiments (section 4.4.1.2), the greatest improvement is seen during the first iteration, with the changes becoming smaller each subsequent iteration. The distributions of scores for the components of the MPO score also improve with every iteration. For some, such as hERG inhibition or MDCK permeability, the scores after only one iteration are already very good and further improvements are marginal. For other components, such as QED or ROCS ShapeTanimoto, the improvement is more consistent across iterations, potentially because they have a more complex relationship to the building block structure and the surrogate models require more data to learn it. Even with the different trends in how the scores are optimised, the above results demonstrate that both the MPO score and all its components can be optimised by *retro-active* in five iterations of active learning, similarly to what was previously shown for single-objective scoring functions.

4.4.3 Case studies

Having benchmarked *retro-active* on AZD5438, we wanted to observe whether the same performance would be obtained for other systems. Two synthesis routes towards known therapeutics were chosen with different route topologies: one linear route and one convergent route. In both cases, *retro-active* was used with five iterations of active learning, with 5,000x5,000x5,000 building blocks acquired and 10,000 product molecules enumerated and scored every loop. At the final acquisition stage 77x77x77 building blocks were acquired and enumerated to form 456,533 product molecules, which in total amounted to 506,533 molecules being scored. For both case studies *retro-active* was compared to random selection of 80x80x80=512,000 product

molecules and a set of molecules from ChEMBL with known activities for the relevant target protein. The effect of active learning iterations was also studied as in section 4.4.2.2.

4.4.3.1 Infigratinib

The first case study was infigratinib, a fibroblast growth factor receptor (FGFR) inhibitor that was initially marketed for cholangiocarcinoma.[314] The synthesis route used is shown in Figure 4.11 and is a shortened version of the literature route towards this molecule.[307] Only three of the building blocks were considered for replacement to restrict the possible product molecule space and the number of molecules scored during each active learning iteration. The MPO function used for scoring was a weighted geometric mean of QED and the ADMET properties described in section 4.4.2 (each with weight=1) and the ROCS TanimotoCombo score (with weight=2).

The distribution of the MPO scores and its components for the top-1000 highest scoring molecules obtained by *retro-active* compared to random selection and known binders of FGFR1 from ChEMBL are shown in Figure 4.12A. Similarly to the results obtained for AZD5438, *retro-active* is able to retrieve molecules with better score distributions for the MPO score and all its components than random selection for the synthesis route towards infigratinib. Almost all of the top-scoring molecules are below the threshold for hERG pIC50 and above the threshold for MDCK permeability. However, the average logD of the obtained molecules is higher than it was for AZD5438, potentially due to the larger size of infigratinib. Despite this, the molecules obtained by *retro-active* still have a better MPO score than the known binders of FGFR1 from ChEMBL, mostly aided by the higher ROCS TanimotoCombo scores. The distribution of hERG pIC50 and logD is similar, but sharper, for molecules obtained by *retro-active* when compared to the molecules from ChEMBL. The ChEMBL molecules have a better QED score and MDCK permeability, as was the case for AZD5438, but

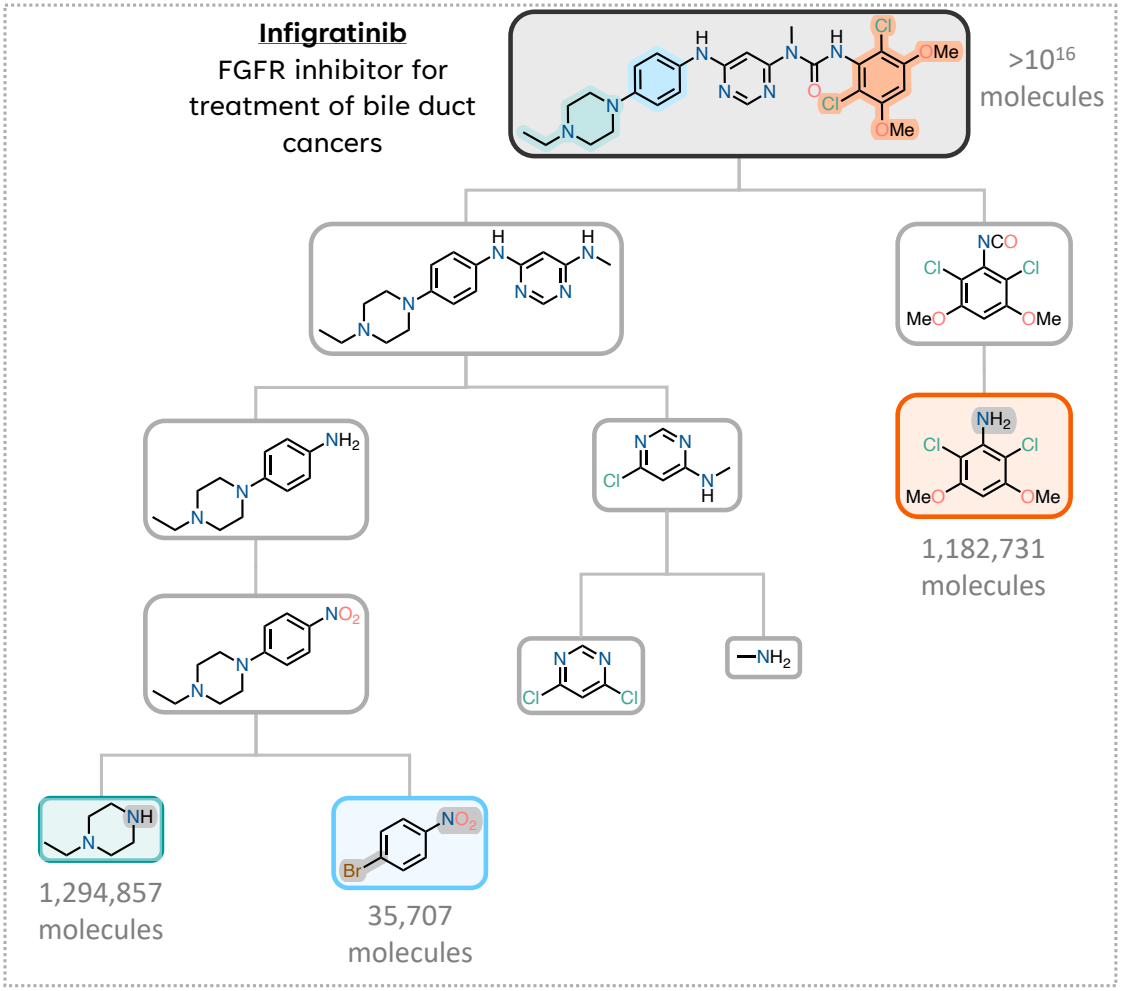


Figure 4.11: Synthesis route towards infigratinib. The green, orange and blue boxes contain building blocks considered for replacement, with the corresponding substructure highlighted in the product molecule in the same colour. The size of the compatible building block pool and the enumerated product space is included next to the molecules. The compatible building blocks were selected based on containing the substructures highlighted in gray.

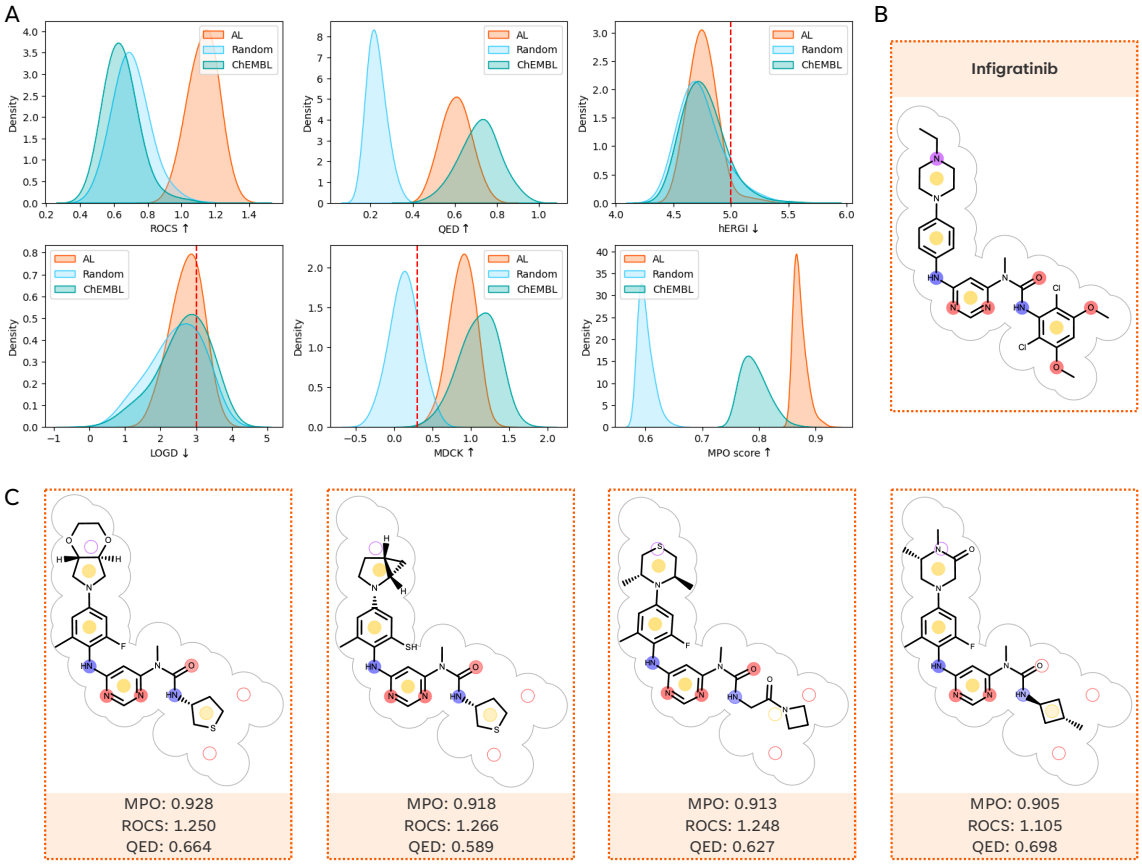


Figure 4.12: Performance of *retro-active* for infigratinib. (A) Comparison of score distribution for the 1,000 highest scoring molecules obtained by *retro-active* (AL, in orange), random selection of building blocks (Random, in blue) and molecules from ChEMBL with associated activities for FGFR1 (ChEMBL, in green). The shown scores are ROCS TanimotoCombo, QED, hERG pIC50, logD, MDCK log_{10}(AB Papp) and the MPO score. The PXR activation score distribution is not provided as all top-1000 molecules were predicted to be inactive. The red lines represent commonly acknowledged thresholds for the scores where relevant, while the arrows represent whether a higher (up) or lower (down) score is more desirable. (B) The reference compound with pharmacophores considered by ROCS highlighted in colour. (C) Example molecules from top-20 obtained after active learning. The ROCS shape and pharmacophore overlay with infigratinib is visualised, with the QED, ROCS TanimotoCombo and MPO scores provided below.

retro-active still improves those scores over random selection.

Selected top scoring molecules retrieved by *retro-active* are shown in Figure 4.12C, with their shape and pharmacophore overlap with infigratinib visualised. The pharmacophores for infigratinib are provided as reference in Figure 4.12B. Within the top-20 molecules obtained by *retro-active*, more variety in the sampled building blocks was seen for the terminal building blocks than for the core, similarly to what was observed for AZD5438. The ShapeTanimoto score for the top molecules was much higher than the ColourTanimoto score ($\sim$0.7 vs $\sim$0.5) showing that optimising for shape similarity was an easier task than for pharmacophore similarity. The low ColourTanimoto score can be mostly attributed to the lack of pharmacophores corresponding to the two oxygens on 1,5-dichloro-2,4-dimethoxybenzene and in some cases the nitrogen on 1-ethylpiperazine. It is possible that because 1,5-dichloro-2,4-dimethoxybenzene is quite a complex and substituted building block, there are not many other compatible building blocks in the pool that would be able to replicate both its shape and pharmacophores, making it difficult for *retro-active* to find them.

In regard to the performance across different active learning iterations, the increase in average MPO score was yet again observed with every AL iteration (as shown in Figure 4.13). Significant improvement was seen during the first three iterations with further changes being more marginal. Interestingly, the same behaviour was not exhibited by all the components of the MPO score. The logD distribution improved after the first AL iteration, but then got worse again, most likely as a trade-off to improve the other components of the MPO score.

4.4.3.2 Lasmiditan

The second case study was lasmiditan, a serotonin receptor agonist marketed for the treatment of migraines.[315] The synthesis route used is shown in Figure 4.14 and is a longer version of the route provided in the patent for this molecule.[306] Two steps

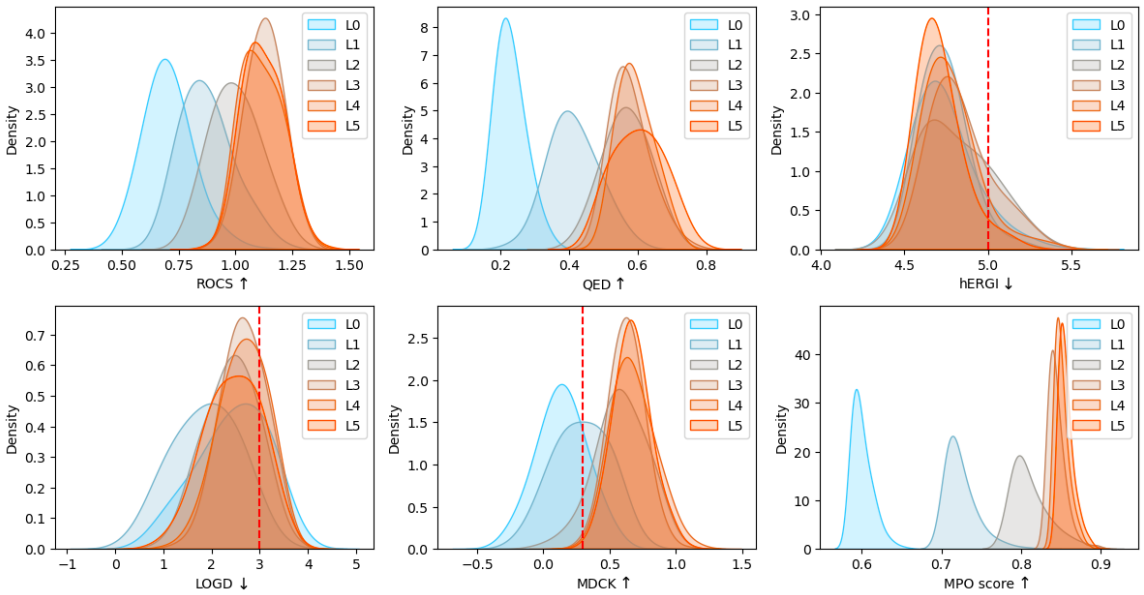


Figure 4.13: Effect of number of active learning iterations on optimisation of infigratinib. The distributions of ROCS TanimotoCombo, QED, hERG pIC50, logD, MDCK log_{10}(AB Papp) and MPO score are shown for top-1000 molecules obtained from *retro-active* after 0-5 loops of active learning (L0 to L5).

were added for the synthesis of the Weinreb amide intermediate from the carboxylic acid to expand the size of the compatible building block pool. The MPO function used for scoring was a weighted geometric mean of QED and the ADMET properties described in section 4.4.2 (each with weight=1) and the docking score for 5-HT1F (with weight=3).

Figure 4.15A shows the distribution of MPO score and its components for the top-1000 molecules obtained by *retro-active* compared to the top-1000 molecules after random selection of building blocks and all 175 molecules with associated activities for 5-HT1F found in ChEMBL. A similar performance for *retro-active* was observed as in the case of the two previous systems. The molecules obtained by *retro-active* had, on average, a better MPO score than those from random selection or ChEMBL. The distribution of scores for all the MPO components was also improved when compared to random, or, in the case of logD, comparable to random. The molecules obtained by *retro-active* are drug-like, with better or almost comparable properties

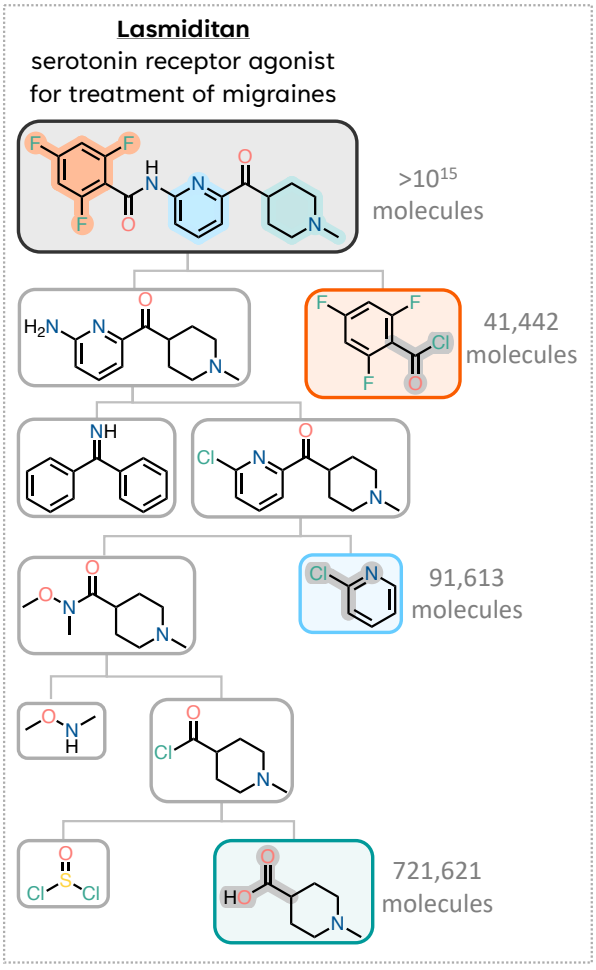


Figure 4.14: Synthesis route towards lasmiditan. The green, orange and blue boxes contain building blocks considered for replacement, with the corresponding substructure highlighted in the product molecule in the same colour. The size of the compatible building block pool and the enumerated product space is included next to the molecules. The compatible building blocks were selected based on containing the substructures highlighted in gray.

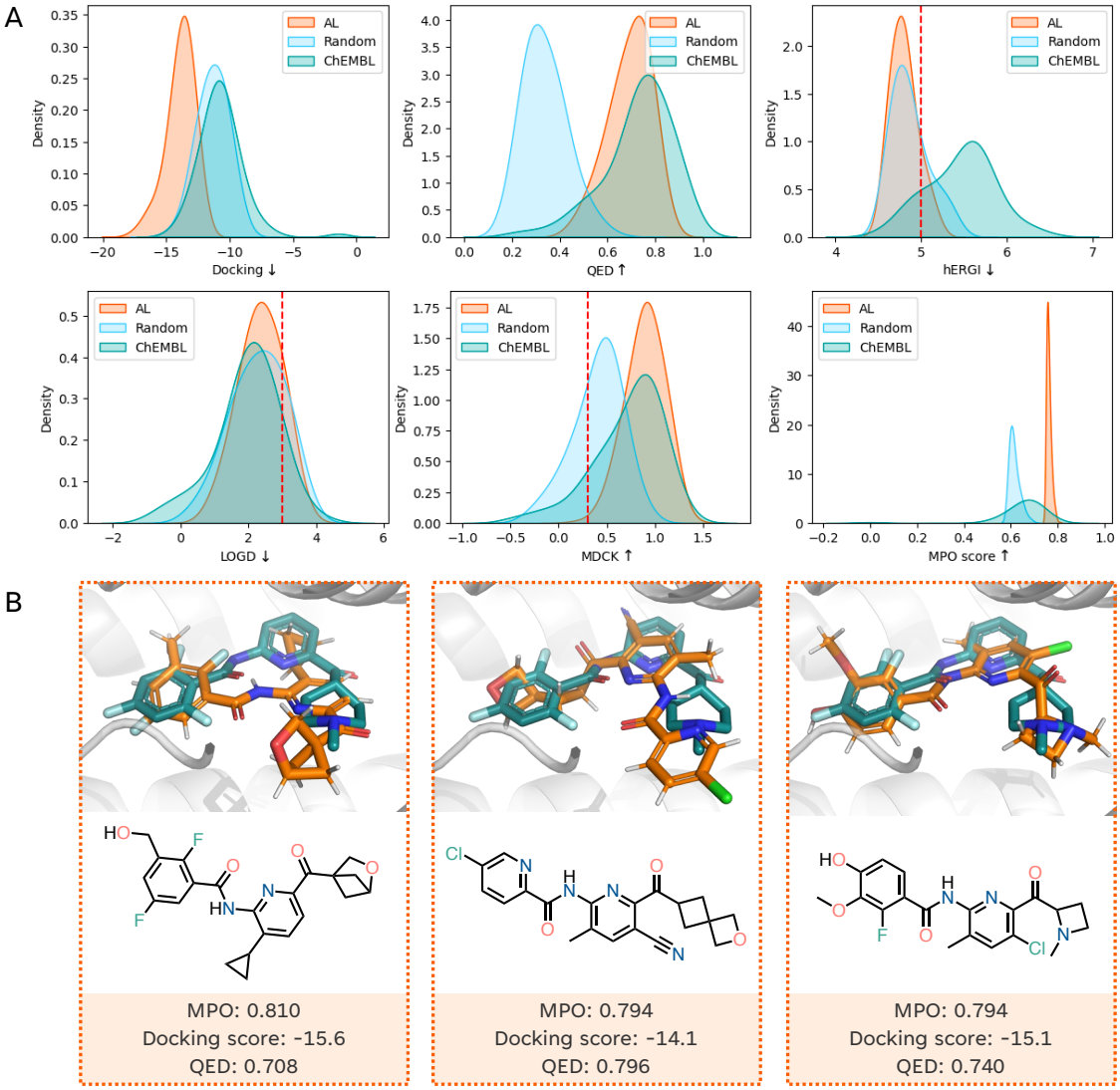


Figure 4.15: Performance of *retro-active* for lasmiditan. (A) Comparison of score distribution for the 1,000 highest scoring molecules obtained by *retro-active* (AL, in orange), random selection of building blocks (Random, in blue) and molecules from ChEMBL with associated activities for FGFR1 (ChEMBL, in green). The shown scores are docking score, QED, hERG pIC50, logD, MDCK log_{10}(AB Papp) and the MPO score. The PXR activation score distribution is not provided as all top-1000 molecules obtained by *retro-active* were predicted to be inactive. The red lines represent commonly acknowledged thresholds for the scores where relevant, while the arrows represent whether a higher (up) or lower (down) score is more desirable. (B) Example molecules from top-10 obtained after active learning. The docked pose of the molecule (orange) together with lasmiditan (green, for reference) is shown, with the QED, ROCS ShapeTanimoto and MPO scores provided below.

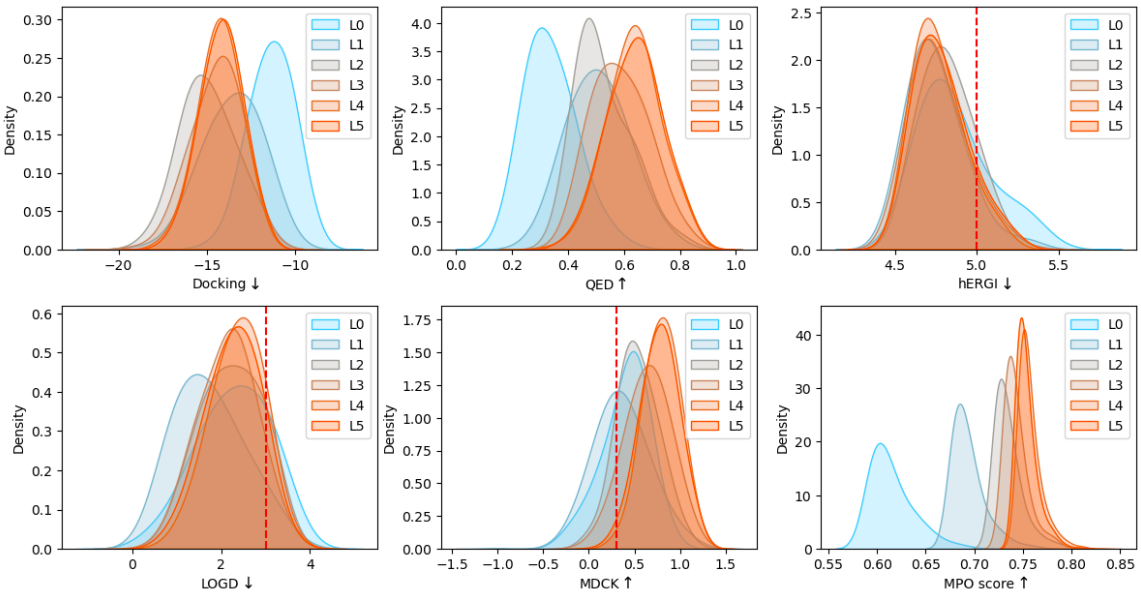


Figure 4.16: Effect of number of active learning iterations on optimisation of lasmiditan. The distributions of docking score, QED, hERG pIC50, logD, MDCK log_{10}(AB Papp) and MPO score are shown for top-1000 molecules obtained from *retro-active* after 0-5 loops of active learning (L0 to L5).

to the molecules from ChEMBL.

Figure 4.15B contains example top-10 molecules obtained by *retro-active*. The docked poses are shown, together with lasmiditan as a reference. Similarly to the analogues obtained for infigratinib and AZD5438, more variety was seen for the terminal building blocks than the core in the top scoring product molecules. All the core building blocks included in the top-10 molecules are substituted pyridines, while various heterocycles, including spirocycles and bridged rings, are sampled for the piperidine-based building block.

The MPO score distribution for the top-1000 molecules obtained by *retro-active* improved with each active learning iteration until four loops were reached, with a comparable distribution for the molecules obtained after the fifth iteration (Figure 4.16). This is different from the experiments with infigratinib and AZD5438, where a marginal improvement was still obtained during the fifth iteration. The step-wise improvement during every loop was not seen for every component of the MPO score,

similarly to the study with infigratinib. Both logD and the docking score had a better distribution after one of the earlier iterations, but some of this improvement was sacrificed to boost the other scores (likely QED).

4.5 Conclusions

This chapter introduced *retro-active*, a new method for generating synthesisable molecules based off known synthesis routes while also optimising them for user-defined scoring functions. The benchmarking experiments established that explorative acquisition and hypercube enumeration for five active learning iterations were the most optimal configuration for *retro-active*. *Retro-active* was able to recover 59% of the ground truth best molecules when using ROCS as the scoring function and 69% when using docking by only scoring $\sim 7\%$ of the total product molecule space. This demonstrated the applicability of *retro-active* in both ligand-based and structure-based drug design. *Retro-active* was also shown to perform better than product molecule enumeration from building blocks selected based on 2D similarity, scores from a subset of products or random selection.

Retro-active was proven to be suitable for a variety of synthesis routes with different reactions and topologies, structurally different target molecules and a variety of scoring functions, including multi-parameter optimisation. The multi-parameter objectives included both physics-based (ROCS and docking) and data-based (ML prediction of ADMET) scores, the latter provided by both regression and classification models. This demonstrates that *retro-active* is capable of combining scores from different sources and optimising the product molecules for all of them.

The molecules returned by *retro-active* had drug-like properties and were comparable to the active molecules retrieved from ChEMBL. To account for properties not included in the multi-parameter objective, such as number of rotatable bonds

or presence of undesirable functional groups, a more extensive building block stock filtering can be carried out before it is supplied to *retro-active*.

The modularity of *retro-active* allows for it to be easily extended to other acquisition or enumeration methods, as well as to the use of different scoring functions and surrogate models. Possible future directions include exploring uncertainty-based building block acquisition methods that would be better able to capture the relationship between the building block and the distribution of scores of its product molecules. While a Random Forest was used as the surrogate model in this study due to its low computational cost, the use of deep learning models, such as message-passing neural networks, could be considered in the future to boost performance.

Chapter 5

Conclusions and Future work

This thesis explores the use of machine learning in drug discovery to ensure the synthesisability of the designed compounds. Two complimentary approaches are explored: post-hoc retrosynthesis prediction and synthesisability-constrained molecule generation.

Chapter 3 addressed the synthesisability issue for molecules containing heterocyclic scaffolds, which are important motifs in drug discovery. Various transfer learning approaches were benchmarked to improve the performance of a sequence-to-sequence single-step retrosynthesis prediction model for ring-breaking disconnections. Among them, mixed fine-tuning was found to be most optimal, with a short training time and improved performance for heterocycle formation prediction, while retaining acceptable performance for other reaction classes. While the improvement in model accuracy was modest compared to previous work for forward reaction prediction,[174] this likely reflects the greater diversity in heterocycle formation reactions. The transfer learning methodology proposed by us can be adopted for other reaction datasets of interest and we provided a workflow for further fine-tuning of the models on additional reaction datasets. Finally, the mixed fine-tuned model has been applied for multi-step retrosynthesis of two recently published drug-like probe molecules, showcasing how

the model can be used in drug discovery scenarios for novel molecules.

Chapter 4 introduced *retro-active*, a framework for synthesisable molecule generation from known synthesis routes based on building block enumeration. Unlike previous work in synthesis route-based enumeration which either fully enumerated the building blocks or filtered them using heuristics,[298, 299] *retro-active* uses active learning with surrogate models to select the building blocks that will optimise the synthesis route product for a user-defined scoring function. Benchmarking of *retro-active* revealed that by scoring only 7% of the combinatorial space it was able to recover respectively 59% and 69% of ground truth best molecules when using a ligand-based and structure-based scoring function, and that it outperformed other non-ML based selection methods. While this is a noteworthy improvement in efficiency, some of the ground truth top molecules were never recovered even when increasing the number of active learning iterations. This is likely an intrinsic limitation of building block (instead of product molecule) selection, when sometimes a high scoring molecule might be made from an unlikely combination of low scoring building blocks.

The use cases of *retro-active* were further expanded to multi-parameter optimisation to mimic real-life drug discovery settings. We demonstrated that *retro-active* can simultaneously optimise molecules for docking or shape similarity and physicochemical and ADMET properties in a combinatorial space of over 10^{15} product molecules. The breadth of application was shown with regards to both the synthesis route topology (linear or convergent) and the disease and protein target. While the highest scoring output molecules had drug-like properties and appearance, the compounds were only tested with *in silico* scoring functions and true experimental validation is needed to fully confirm *retro-active's* applicability in drug discovery.

5.1 Future directions

5.1.1 Retrosynthesis prediction

Despite the many recent developments in single-step retrosynthesis, most real-world applications of Computer-Aided Synthesis Planning in drug discovery are still limited to the use of template-based models with Monte Carlo Tree Search. This may just be due to the slow speed of adoption of new methods by pharmaceutical industry, but could also stem from the disconnect between the academic benchmarks and the actual use cases of retrosynthesis in drug discovery. While most academic work focuses on the improvement of single-step retrosynthesis prediction models and benchmarks them through top-N accuracy (as an easily computable proxy), it is unclear whether those incremental increases in top-N accuracy actually translate to better performance for predicting full synthesis routes, and the only true way to test this would be through extensive prospective experimental validation, which is costly and time-consuming. There is a need for more relevant "proxy" evaluation metrics that would reflect the common use cases of those models in industry: either as synthesisability filters or to predict routes to molecules that can be made with only a handful of pharmaceutically relevant reactions (e.g., Suzuki, amide coupling). An overview of possible future work to improve the adoption of academic retrosynthesis models (both those presented in this thesis and more general) and bridge the gap between single-step and multi-step retrosynthesis is included below.

5.1.1.1 Data availability and quality

Most academic models have been trained on patent reactions, which cover a different chemical space to literature data.[242] This can bias them towards making predictions that resemble "old-school" chemistry, which is illustrated by the abundance of ketone-based condensations predicted by our heterocycle-focused retrosynthesis

models. This limits the models' applicability and can be off-putting for experienced synthetic chemists, who expect the models to suggest more innovative chemistry. It is hoped that initiatives such as the Open Reaction Database (ORD)[316] will improve the diversity and quality of publicly available reaction data and therefore also the academic model performance. If more heterocycle formation data is deposited to ORD, it is recommended that our heterocycle-focused models be retrained with it using the same transfer learning approaches. The application of our transfer learning methods for other reaction types is also another avenue of research that would be enabled by access to more reaction data.

Additionally, the quality of reaction data in the automatically extracted datasets has often been criticised, with errors during extraction leading to unfeasible reactions present. This would naturally limit the performance of the models trained on those datasets. Manual inspection of our ring formation dataset revealed that there are still unfeasible reactions present, despite the data cleaning undertaken. Future work could focus on implementing more rigorous data cleaning pipelines, either by defining a library of hand encoded rules or training a reaction correctness prediction model. However, the latter approach would require the creation of the training dataset, most likely through time-consuming manual labelling of the extracted data.

5.1.1.2 Evaluation metrics

The issues with current single-step retrosynthesis evaluation metrics have already been highlighted in this thesis (see Section 1.5.1.3) and no perfect metric has been developed yet. While this thesis introduced a new evaluation metric for heterocycle disconnections (ring-breaking round-trip accuracy), it suffers from the same problem of reliance on a forward reaction prediction model as round-trip accuracy. The ideal reaction viability metric would use a reaction classification model trained on positive and negative reaction data, however public datasets, and in general reaction reporting,

are biased towards positive outcomes. Future work could make use of electronic lab notebook (ELN) data or curate a synthetic negative reaction dataset based on a set of expert-defined rules to train such a viability model.

5.1.1.3 Continual learning

With new reactions being published every day, it is important to be able to consistently and iteratively update the retrosynthesis prediction models. The simplest approach would involve retraining the models on the full dataset from scratch, however it is time consuming and requires access to all previous training data. In this thesis we have introduced a further fine-tuning workflow for seq2seq models that improves the time efficiency over retraining from scratch. However, access to all training datasets is still needed, with all our attempts at removing the *Ring* dataset or its subsets from the further fine-tuning step resulting in significantly reduced accuracy for those reactions. Alternative approaches can be explored in the future that have been used in other machine learning applications for continual learning, such as regularisation-based or replay-based approaches.[285]

5.1.1.4 Multi-step retrosynthesis

Recent studies have showcased that the performance differences between single-step retrosynthesis models do not directly translate to multi-step retrosynthesis.[317] While a small qualitative comparison of the *baseline* and *mixed fine-tuned* model was performed in this thesis, the need for a larger quantitative benchmark between all the models presented in this thesis is indicated. Such a benchmark should be performed on a set of heterocycle-containing drug-like molecules, whose synthesis involved a ring formation step. Moreover, apart from standard evaluation metrics such as success rate, search time, prediction diversity and tree edit distance from reference route, heterocycle-specific metrics should be introduced, for example ring-breaking success

rate.

Additionally, our preliminary analysis has shown that single-step retrosynthesis models trained on the same dataset but with different model architectures (template-based vs seq2seq) cover slightly different chemical space. Together with the recent work of Maziarz *et al.* on ensembling models in multi-step predictions,[318] this shows promise for an adoption of both models in multi-step planning to improve performance.

5.1.2 Synthesisability-focused molecule generation

While most of the current work in molecule generation focuses on deep learning based approaches, with *retro-active* we have demonstrated that a combination of simple machine learning models with cheminformatics-based methods can also be an effective and computationally cheaper method for synthesisable molecule generation. However, there is still room for improvement in this method, both on the side of cheminformatics, with reaction-based enumeration, and the algorithms used for building block selection and molecule optimisation.

5.1.2.1 Reaction-based enumeration

The quality and chemical validity of product molecules created through reaction-based enumeration is directly dependent on the reaction templates applied. While improvements have been made to template extraction in recent years, there is usually a trade-off between generalisability and accuracy of the automatically extracted templates. In particular, issues such as regio- or chemo-selectivity are rarely captured. This means that while all molecules generated by reaction-based enumeration tools, including *retro-active*, are supposed to be theoretically synthesisable, in practice the proportion will be significantly lower than 100%. While development of more reliable methods for template extraction would solve this problem, they are outside

of the scope of the topics covered by this thesis. A simpler approach that can be applied to *retro-active* is the inclusion of hand-curated SMARTS-based filters that would detect substructure incompatibilities with specific reaction types or guide the correct regiochemistry. This has already been partially implemented in *retro-active*, with molecules filtered out if they have multiples of the same template substructure present to avoid chemoselectivity issues.

Another improvement that can be implemented in *retro-active* is the pre-filtering of building blocks with simple numerical descriptors. While *retro-active* has demonstrated the ability for multi-parameter optimisation, some properties, such as molecular weight or number of rotatable bonds, are reasonably additive and very quick to calculate, so using them as filters instead of optimising them through active learning (explicitly or implicitly through for example QED) would make the process more efficient.

5.1.2.2 Optimisation methods

While Random Forests have been used in this thesis as the surrogate models in *retro-active*, this design choice was made mostly due to their quick training time. The modularity of *retro-active* allows for any models to be employed in the framework and future work could focus on benchmarking more complex deep learning algorithms, such as the message-passing neural network which is frequently used for molecular property prediction,[319] with the hope of improving performance. The use of models with an intrinsic uncertainty quantification would also allow for easy inclusion of alternative acquisition methods, geared more towards exploration than exploitation. However based on previous benchmarks in active learning, the utility of those uncertainty-based acquisition methods might be limited.[264]

An alternative approach to explore is the use of Bayesian methods instead of active learning with ML surrogate models. Thompson Sampling in particular has recently

been shown to perform well in searching large on-demand libraries without the need for full enumeration,[320] and *retro-active* should at least be benchmarked against it.

5.1.3 Experimental validation

While both computational approaches presented in this thesis, the heterocycle retrosynthesis prediction models and *retro-active*, have been extensively validated *in silico*, this can only be treated as an estimation of their real-life applicability. True prospective experimental validation is necessary to fully establish the effects these approaches can have on early-stage drug discovery. The heterocycle retrosynthesis prediction models could be used to either attempt the synthesis of a novel heterocyclic scaffold (a difficult task) or a novel medicinally-relevant molecule that contains a known heterocyclic scaffold (an easier task). The applicability of *retro-active* would be most easily tested in an ongoing drug discovery campaign for hit or lead optimisation, with the scoring function including an experimental assay component.

References

1. Wouters, O. J., McKee, M. & Luyten, J. Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009-2018. *JAMA* **323,** 844–853 (2020).

2. Scannell, J. W., Blanckley, A., Boldon, H. & Warrington, B. Diagnosing the decline in pharmaceutical R&D efficiency. *Nature Reviews Drug Discovery* **11,** 191–200 (2012).

3. Cook, D. *et al.* Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework. *Nature Reviews Drug Discovery* **13,** 419–431 (2014).

4. Sadybekov, A. V. & Katritch, V. Computational approaches streamlining drug discovery. *Nature* **616,** 673–685 (2023).

5. Pandey, M. *et al.* The transformational role of GPU computing and deep learning in drug discovery. *Nature Machine Intelligence* **4,** 211–221 (2022).

6. Catacutan, D. B., Alexander, J., Arnold, A. & Stokes, J. M. Machine learning in preclinical drug discovery. *Nature Chemical Biology* **20,** 960–973 (2024).

7. Vamathevan, J. *et al.* Applications of machine learning in drug discovery and development. *Nature Reviews Drug Discovery* **18,** 463–477 (2019).

8. Grygorenko, O. O. *et al.* Generating Multibillion Chemical Space of Readily Accessible Screening Compounds. *iScience* **23** (2020).

9. Jones, N. Crystallography: Atomic secrets. *Nature* **505,** 602–603 (2014).

10. Renaud, J.-P. *et al.* Cryo-EM in drug discovery: achievements, limitations and prospects. *Nature Reviews Drug Discovery* **17,** 471–492 (2018).

11. Fernandez-Leiro, R. & Scheres, S. H. W. Unravelling biological macromolecules with cryo-electron microscopy. *Nature* **537,** 339–346 (2016).

12. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630,** 493–500 (2024).

13. Wohlwend, J. *et al.* Boltz-1 Democratizing Biomolecular Interaction Modeling. *bioRxiv.* eprint: `https://www.biorxiv.org/content/early/2024/11/20/2024.11.19.624167.full.pdf` (2024).

14. team, C. D. *et al.* Chai-1: Decoding the molecular interactions of life. *bioRxiv.* eprint: `https://www.biorxiv.org/content/early/2024/10/11/2024.10.10.615955.full.pdf` (2024).

15. Gao, W. & Coley, C. W. The Synthesizability of Molecules Proposed by Generative Models. *Journal of Chemical Information and Modeling* **60,** 5714–5723 (28, 2020).

16. Hughes, J., Rees, S., Kalindjian, S. & Philpott, K. Principles of early drug discovery. *British Journal of Pharmacology* **162,** 1239–1249 (2011).

17. Lu, R.-M. *et al.* Development of therapeutic antibodies for the treatment of diseases. *Journal of Biomedical Science* **27,** 1 (2020).

18. Wang, L. *et al.* Therapeutic peptides: current applications and future directions. *Signal Transduction and Targeted Therapy* **7,** 1–27 (2022).

19. Arnold, C. PROTAC protein degraders to drug the undruggable enter phase 3 trials. *Nature Medicine* **30,** 3030–3031 (2024).

20. Tabana, Y., Babu, D., Fahlman, R., Siraki, A. G. & Barakat, K. Target identification of small molecules: an overview of the current applications in drug discovery. *BMC Biotechnology* **23,** 44 (2023).

21. Schenone, M., Dančík, V., Wagner, B. K. & Clemons, P. A. Target identification and mechanism of action in chemical biology and drug discovery. *Nature Chemical Biology* **9,** 232–240 (2013).

22. Macarron, R. *et al.* Impact of high-throughput screening in biomedical research. *Nature Reviews Drug Discovery* **10,** 188–195 (2011).

23. Lyu, J. *et al.* Ultra-large library docking for discovering new chemotypes. *Nature* **566,** 224–229 (2019).

24. Stein, R. M. *et al.* Virtual discovery of melatonin receptor ligands to modulate circadian rhythms. *Nature* **579,** 609–614 (2020).

25. Murray, C. W. & Rees, D. C. The rise of fragment-based drug discovery. *Nature Chemistry* **1,** 187–192 (2009).

26. Schuffenhauer, A. *et al.* Library Design for Fragment Based Screening. *Current Topics in Medicinal Chemistry* **5,** 751–762 (2005).

27. Kirsch, P., Hartman, A. M., Hirsch, A. K. H. & Empting, M. Concepts and Core Principles of Fragment-Based Drug Design. *Molecules* **24,** 4309 (2019).

28. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings1. *Advanced Drug Delivery Reviews. Special issue dedicated to Dr. Eric Tomlinson, Advanced Drug Delivery Reviews, A Selection of the Most Highly Cited Articles, 1991-1998* **46,** 3–26 (2001).

29. Plowright, A. T. *et al.* Hypothesis driven drug design: improving quality and effectiveness of the design-make-test-analyse cycle. *Drug Discovery Today* **17,** 56–62 (2012).

30. Kandi, V., Vadakedath, S., Kandi, V. & Vadakedath, S. Clinical Trials and Clinical Research: A Comprehensive Review. *Cureus* **15** (2023).

31. Bajorath, J. Computer-aided drug discovery. *F1000Research* **4** (2015).

32. Levin, I., Fortunato, M. E., Tan, K. L. & Coley, C. W. Computer-aided evaluation and exploration of chemical spaces constrained by reaction pathways. *AIChE Journal* **69,** e18234 (2023).

33. Niazi, S. K. & Mariam, Z. Computer-Aided Drug Design and Drug Discovery: A Prospective Analysis. *Pharmaceuticals* **17,** 22 (2024).

34. Sabe, V. T. *et al.* Current trends in computer aided drug design and a highlight of drugs discovered via computational techniques: A review. *European Journal of Medicinal Chemistry* **224,** 113705 (2021).

35. Drie, J. H. V. Computer-aided drug design: the next 20 years. *Journal of Computer-Aided Molecular Design* **21** (2007).

36. Braga, R. C. *et al.* Virtual Screening Strategies in Medicinal Chemistry: The State of the Art and Current Challenges. *Current Topics in Medicinal Chemistry* **14,** 1899–1912 (2014).

37. Sun, H. Pharmacophore-Based Virtual Screening. *Current Medicinal Chemistry* **15,** 1018–1024 (2008).

38. Lionta, E., Spyrou, G., Vassilatis, D. K. & Cournia, Z. Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. *Current Topics in Medicinal Chemistry* **14,** 1923–1938.

39. Alon, A. *et al.* Structures of the 2 receptor enable docking for bioactive ligand discovery. *Nature* **600,** 759–764 (2021).

40. Gorgulla, C. *et al.* An open-source drug discovery platform enables ultra-large virtual screens. *Nature* **580,** 663–668 (2020).

41. Friesner, R. A. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry* **47,** 1739–1749 (2004).

42. Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray, C. W. & Taylor, R. D. Improved protein–ligand docking using GOLD. *Proteins: Structure, Function, and Bioinformatics* **52,** 609–623 (2003).

43. Eberhardt, J., Santos-Martins, D., Tillack, A. F. & Forli, S. AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *Journal of Chemical Information and Modeling* **61,** 3891–3898 (2021).

44. Brooks, B. R. *et al.* CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry* **4,** 187–217 (1983).

45. Case, D. A. *et al.* The Amber biomolecular simulation programs. *Journal of Computational Chemistry* **26,** 1668–1688 (2005).

46. Jones, G., Willett, P., Glen, R. C., Leach, A. R. & Taylor, R. Development and validation of a genetic algorithm for flexible docking. *Journal of Molecular Biology* **267,** 727–748 (1997).

47. Velec, H. F. G., Gohlke, H. & Klebe, G. DrugScoreCSDKnowledge-Based Scoring Function Derived from Small Molecule Crystal Data with Superior Recognition Rate of Near-Native Ligand Poses and Better Affinity Prediction. *Journal of Medicinal Chemistry* **48,** 6296–6303 (2005).

48. Shen, Q. *et al.* Knowledge-Based Scoring Functions in Drug Design: 2. Can the Knowledge Base Be Enriched? *Journal of Chemical Information and Modeling* **51,** 386–397 (2011).

49. Böhm, H.-J. The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. *Journal of Computer-Aided Molecular Design* **8,** 243–256 (1994).

50. Durrant, J. D. & McCammon, J. A. NNScore 2.0: A Neural-Network Receptor–Ligand Scoring Function. *Journal of Chemical Information and Modeling* **51,** 2897–2903 (2011).

51. Ballester, P. J. & Mitchell, J. B. A machine learning approach to predicting protein-ligand binding affinity with applications to molecular docking. *Bioinformatics* **26,** 1169–1175 (2010).

52. Rao, S. N., Head, M. S., Kulkarni, A. & LaLonde, J. M. Validation Studies of the Site-Directed Docking Program LibDock. *Journal of Chemical Information and Modeling* **47,** 2159–2171 (2007).

53. Lu, W. *et al.* TANKBind: Trigonometry-Aware Neural NetworKs for Drug-Protein Binding Structure Prediction. *Advances in Neural Information Processing Systems* **35,** 7236–7249 (2022).

54. Méndez-Lucio, O., Ahmad, M., del Rio-Chanona, E. A. & Wegner, J. K. A geometric deep learning approach to predict binding conformations of bioactive molecules. *Nature Machine Intelligence* **3,** 1033–1039 (2021).

55. Corso, G., Stärk, H., Jing, B., Barzilay, R. & Jaakkola, T. *DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking* 2023. arXiv: 2210.01776 [q-bio.BM].

56. Stärk, H., Ganea, O.-E., Pattanaik, L., Barzilay, R. & Jaakkola, T. *EquiBind: Geometric Deep Learning for Drug Binding Structure Prediction* 2022. arXiv: 2202.05146 [q-bio.BM].

57. Buttenschoen, M., M. Morris, G. & M. Deane, C. PoseBusters: AI-based docking methods fail to generate physically valid poses or generalise to novel sequences. *Chemical Science* **15,** 3130–3139 (2024).

58. Graff, D. E., Shakhnovich, E. I. & Coley, C. W. Accelerating high-throughput virtual screening through molecular pool-based active learning. *Chemical Science* **12,** 7866–7881 (9, 2021).

59. Sadybekov, A. A. *et al.* Synthon-based ligand discovery in virtual libraries of over 11 billion compounds. *Nature* **601,** 452–459 (2022).

60. Beroza, P. *et al.* Chemical space docking enables large-scale structure-based virtual screening to discover ROCK1 kinase inhibitors. *Nature Communications* **13,** 6447 (2022).

61. Klebe, G. Virtual ligand screening: strategies, perspectives and limitations. *Drug Discovery Today* **11,** 580 (2006).

62. Jung, S., Vatheuer, H. & Czodrowski, P. VSFlow: an open-source ligand-based virtual screening tool. *Journal of Cheminformatics* **15,** 40 (2023).

63. *ROCS 3.5.0. OpenEye, Cadence Molecular Sciences, Santa Fe, NM.*

64. Walters, W. P. Virtual Chemical Libraries. *Journal of Medicinal Chemistry* **62,** 1116–1124 (2019).

65. Van Hilten, N., Chevillard, F. & Kolb, P. Virtual Compound Libraries in Computer-Assisted Drug Discovery. *Journal of Chemical Information and Modeling* **59,** 644–651 (2019).

66. *REAL Database - Enamine*

67. Schneider, G. *De novo* design – hop(p)ing against hope. *Drug Discovery Today: Technologies. 10_4* **10,** e453–e460 (2013).

68. Du, Y. *et al.* Machine learning-aided generative molecular design. *Nature Machine Intelligence* **6,** 589–604 (2024).

69. Schneider, G. & Fechner, U. Computer-based de novo design of drug-like molecules. *Nature Reviews Drug Discovery* **4,** 649–663 (2005).

70. Polishchuk, P. G., Madzhidov, T. I. & Varnek, A. Estimation of the size of drug-like chemical space based on GDB-17 data. *Journal of Computer-Aided Molecular Design* **27,** 675–679 (2013).

71. Rotstein, S. H. & Murcko, M. A. GenStar: a method for de novo drug design. *Journal of Computer-Aided Molecular Design* **7,** 23–43 (1993).

72. Nishibata, Y. & Itai, A. Automatic creation of drug candidate structures based on receptor structure. Starting point for artificial lead generation. *Tetrahedron* **47,** 8985–8990 (1991).

73. Böhm, H.-J. The computer program LUDI: A new method for the de novo design of enzyme inhibitors. *Journal of Computer-Aided Molecular Design* **6,** 61–78 (1992).

74. Clark, D. E. *et al.* PRO_LIGAND: An approach to de novo molecular design. 1. Application to the design of organic molecules. *Journal of Computer-Aided Molecular Design* **9,** 13–32 (1995).

75. Rotstein, S. H. & Murcko, M. A. GroupBuild: a fragment-based method for de novo drug design. *Journal of Medicinal Chemistry* **36,** 1700–1710 (1993).

76. Wang, R., Gao, Y. & Lai, L. LigBuilder: A Multi-Purpose Program for Structure-Based Drug Design. *Molecular modeling annual* **6,** 498–516 (2000).

77. Hartenfeller, M. *et al.* DOGS: Reaction-Driven de novo Design of Bioactive Compounds. *PLOS Computational Biology* **8,** e1002380 (2012).

78. Brown, N., McKay, B., Gilardoni, F. & Gasteiger, J. A Graph-Based Genetic Algorithm and Its Application to the Multiobjective Evolution of Median Molecules. *Journal of Chemical Information and Computer Sciences* **44,** 1079–1087 (2004).

79. Besnard, J. *et al.* Automated design of ligands to polypharmacological profiles. *Nature* **492,** 215–220 (2012).

80. Dalke, A., Hert, J. & Kramer, C. mmpdb: An Open-Source Matched Molecular Pair Platform for Large Multiproperty Data Sets. *Journal of Chemical Information and Modeling* **58,** 902–910 (2018).

81. Yang, Z. *et al.* Matched Molecular Pair Analysis in Drug Discovery: Methods and Recent Applications. *Journal of Medicinal Chemistry* **66,** 4361–4377 (2023).

82. Hussain, J. & Rea, C. Computationally Efficient Algorithm to Identify Matched Molecular Pairs (MMPs) in Large Data Sets. *Journal of Chemical Information and Modeling* **50,** 339–348 (2010).

83. Hajduk, P. J. & Sauer, D. R. Statistical Analysis of the Effects of Common Chemical Substituents on Ligand Potency. *Journal of Medicinal Chemistry* **51,** 553–564 (2008).

84. Gómez-Bombarelli, R. *et al.* Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules. *ACS Central Science* **4,** 268–276 (2018).

85. Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *Journal of Chemical Information and Computer Sciences* **28,** 31–36 (1988).

86. Krenn, M., Häse, F., Nigam, A., Friederich, P. & Aspuru-Guzik, A. Self-referencing embedded strings (SELFIES): A 100% robust molecular string representation. *Machine Learning: Science and Technology* **1,** 045024 (2020).

87. Simonovsky, M. & Komodakis, N. *GraphVAE: Towards Generation of Small Graphs Using Variational Autoencoders* 2018. arXiv: 1802.03480 [cs.LG].

88. Lim, J., Ryu, S., Kim, J. W. & Kim, W. Y. Molecular generative model based on conditional variational autoencoder for de novo molecular design. *Journal of Cheminformatics* **10,** 31 (2018).

89. Olivecrona, M., Blaschke, T., Engkvist, O. & Chen, H. Molecular de-novo design through deep reinforcement learning. *Journal of Cheminformatics* **9,** 48 (2017).

90. Segler, M. H. S., Kogej, T., Tyrchan, C. & Waller, M. P. Generating Focused Molecule Libraries for Drug Discovery with Recurrent Neural Networks. *ACS Central Science* **4,** 120–131 (2018).

91. Cao, N. D. & Kipf, T. *MolGAN: An implicit generative model for small molecular graphs* 2022. arXiv: 1805.11973 [stat.ML].

92. Guimaraes, G. L., Sanchez-Lengeling, B., Outeiral, C., Farias, P. L. C. & Aspuru-Guzik, A. *Objective-Reinforced Generative Adversarial Networks (ORGAN) for Sequence Generation Models* 2018.

93. Shi, C. *et al. GraphAF: a Flow-based Autoregressive Model for Molecular Graph Generation* 2020. arXiv: 2001.09382 [cs.LG].

94. Popova, M., Shvets, M., Oliva, J. & Isayev, O. *MolecularRNN: Generating realistic molecular graphs with optimized properties* 2019. arXiv: 1905.13372 [cs.LG].

95. Gebauer, N. W. A., Gastegger, M. & Schütt, K. T. *Symmetry-adapted generation of 3d point sets for the targeted discovery of molecules* 2020. arXiv: 1906.00957 [stat.ML].

96. Schneuing, A. *et al. Structure-based Drug Design with Equivariant Diffusion Models* 2024. arXiv: 2210.13695 [q-bio.BM].

97. Hoogeboom, E., Satorras, V. G., Vignac, C. & Welling, M. *Equivariant Diffusion for Molecule Generation in 3D* 2022. arXiv: 2203.17003 [cs.LG].

98. Brown, N., Fiscato, M., Segler, M. H. & Vaucher, A. C. GuacaMol: Benchmarking Models for de Novo Molecular Design. *Journal of Chemical Information and Modeling* **59,** 1096–1108 (2019).

99. Loeffler, H. H. *et al.* Reinvent 4: Modern AI–driven generative molecule design. *Journal of Cheminformatics* **16,** 20 (2024).

100. Yang, X., Wang, Y., Byrne, R., Schneider, G. & Yang, S. Concepts of Artificial Intelligence for Computer-Assisted Drug Discovery. *Chemical Reviews* **119,** 10520–10594 (2019).

101. N. Muratov, E. *et al.* QSAR without borders. *Chemical Society Reviews* **49,** 3525–3564 (2020).

102. Tropsha, A., Isayev, O., Varnek, A., Schneider, G. & Cherkasov, A. Integrating QSAR modelling and deep learning in drug discovery: the emergence of deep QSAR. *Nature Reviews Drug Discovery* **23,** 141–155 (2024).

103. Hansch, C., Maloney, P. P., Fujita, T. & Muir, R. M. Correlation of Biological Activity of Phenoxyacetic Acids with Hammett Substituent Constants and Partition Coefficients. *Nature* **194,** 178–180 (1962).

104. Marchese Robinson, R. L., Palczewska, A., Palczewski, J. & Kidley, N. Comparison of the Predictive Performance and Interpretability of Random Forest and Linear Models on Benchmark Data Sets. *Journal of Chemical Information and Modeling* **57,** 1773–1792 (2017).

105. Martin, E. J., Polyakov, V. R., Tian, L. & Perez, R. C. Profile-QSAR 2.0: Kinase Virtual Screening Accuracy Comparable to Four-Concentration IC50s for Realistically Novel Compounds. *Journal of Chemical Information and Modeling* **57,** 2077–2088 (2017).

106. Svetnik, V. *et al.* Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling. *Journal of Chemical Information and Computer Sciences* **43,** 1947–1958 (2003).

107. Ma, J., Sheridan, R. P., Liaw, A., Dahl, G. E. & Svetnik, V. Deep Neural Nets as a Method for Quantitative Structure–Activity Relationships. *Journal of Chemical Information and Modeling* **55,** 263–274 (2015).

108. Xu, Y., Ma, J., Liaw, A., Sheridan, R. P. & Svetnik, V. Demystifying Multitask Deep Neural Networks for Quantitative Structure–Activity Relationships. *Journal of Chemical Information and Modeling* **57,** 2490–2504 (2017).

109. Lenselink, E. B. *et al.* Beyond the hype: deep neural networks outperform established methods using a ChEMBL bioactivity benchmark set. *Journal of Cheminformatics* **9,** 45 (2017).

110. Dablander, M., Hanser, T., Lambiotte, R. & Morris, G. M. Exploring QSAR models for activity-cliff prediction. *Journal of Cheminformatics* **15,** 47 (2023).

111. Van Tilborg, D., Alenicheva, A. & Grisoni, F. Exposing the Limitations of Molecular Machine Learning with Activity Cliffs. *Journal of Chemical Information and Modeling* **62,** 5938–5951 (2022).

112. Tetko, I. V. & Tanchuk, V. Y. Application of Associative Neural Networks for Prediction of Lipophilicity in ALOGPS 2.1 Program. *Journal of Chemical Information and Computer Sciences* **42,** 1136–1145 (2002).

113. Tetko, I. V. & Poda, G. I. Application of ALOGPS 2.1 to Predict logD Distribution Coefficient for Pfizer Proprietary Compounds. *Journal of Medicinal Chemistry* **47,** 5601–5604 (2004).

114. Yang, S.-S., Lu, W.-C., Gu, T.-H., Yan, L.-M. & Li, G.-Z. QSPR Study of n-Octanol/Water Partition Coefficient of Some Aromatic Compounds Using Support Vector Regression. *QSAR & Combinatorial Science* **28,** 175–182 (2009).

115. Cheng, T., Li, Q., Wang, Y. & Bryant, S. H. Binary Classification of Aqueous Solubility Using Support Vector Machines with Reduction and Recombination Feature Selection. *Journal of Chemical Information and Modeling* **51,** 229–236 (2011).

116. Lusci, A., Pollastri, G. & Baldi, P. Deep Architectures and Deep Learning in Chemoinformatics: The Prediction of Aqueous Solubility for Drug-Like Molecules. *Journal of Chemical Information and Modeling* **53,** 1563–1575 (2013).

117. Coley, C. W., Barzilay, R., Green, W. H., Jaakkola, T. S. & Jensen, K. F. Convolutional Embedding of Attributed Molecular Graphs for Physical Property Prediction. *Journal of Chemical Information and Modeling* **57,** 1757–1772 (2017).

118. Hopfinger, A. J., Esposito, E. X., Llinàs, A., Glen, R. C. & Goodman, J. M. Findings of the Challenge To Predict Aqueous Solubility. *Journal of Chemical Information and Modeling* **49,** 1–5 (2009).

119. Broccatelli, F. *et al.* Predicting Passive Permeability of Drug-like Molecules from Chemical Structure: Where Are We? *Molecular Pharmaceutics* **13,** 4199–4208 (2016).

120. Sun, H. *et al.* Highly predictive and interpretable models for PAMPA permeability. *Bioorganic & Medicinal Chemistry* **25,** 1266–1276 (2017).

121. Dixon, S. L. *et al.* Autoqsar: An Automated Machine Learning Tool for Best-Practice Quantitative Structure–Activity Relationship Modeling. *Future Medicinal Chemistry* **8,** 1825–1839 (2016).

122. Kumar, R., Sharma, A., Siddiqui, M. H. & Tiwari, R. K. Prediction of Human Intestinal Absorption of Compounds Using Artificial Intelligence Techniques. *Current Drug Discovery Technologies* **14,** 244–254.

123. Wang, N.-N. *et al.* Predicting human intestinal absorption with modified random forest approach: a comprehensive evaluation of molecular representation, unbalanced data, and applicability domain issues. *RSC Advances* **7,** 19007–19018 (2017).

124. Ingle, B. L., Veber, B. C., Nichols, J. W. & Tornero-Velez, R. Informing the Human Plasma Protein Binding of Environmental Chemicals by Machine Learning in the Pharmaceutical Space: Applicability Domain and Limits of Predictability. *Journal of Chemical Information and Modeling* **56,** 2243–2252 (2016).

125. Votano, J. R. *et al.* QSAR Modeling of Human Serum Protein Binding with Several Modeling Techniques Utilizing StructureInformation Representation. *Journal of Medicinal Chemistry* **49,** 7169–7181 (2006).

126. Gao, Z., Chen, Y., Cai, X. & Xu, R. Predict drug permeability to blood–brain-barrier from clinical phenotypes: drug side effects and drug indications. *Bioinformatics* **33,** 901–908 (2017).

127. Šícho, M., de Bruyn Kops, C., Stork, C., Svozil, D. & Kirchmair, J. FAME 2: Simple and Effective Machine Learning Model of Cytochrome P450 Regioselectivity. *Journal of Chemical Information and Modeling* **57,** 1832–1846 (2017).

128. Zaretzki, J. *et al.* RS-Predictor: A New Tool for Predicting Sites of Cytochrome P450-Mediated Metabolism Applied to CYP 3A4. *Journal of Chemical Information and Modeling* **51,** 1667–1689 (2011).

129. Mishra, N. K., Agarwal, S. & Raghava, G. P. Prediction of cytochrome P450 isoform responsible for metabolizing a drug molecule. *BMC Pharmacology* **10,** 8 (2010).

130. Lombardo, F. & Jing, Y. In Silico Prediction of Volume of Distribution in Humans. Extensive Data Set and the Exploration of Linear and Nonlinear

Methods Coupled with Molecular Interaction Fields Descriptors. *Journal of Chemical Information and Modeling* **56,** 2042–2052 (2016).

131. Berellini, G., Waters, N. J. & Lombardo, F. In silico Prediction of Total Human Plasma Clearance. *Journal of Chemical Information and Modeling* **52,** 2069–2078 (2012).

132. Mayr, A., Klambauer, G., Unterthiner, T. & Hochreiter, S. DeepTox: Toxicity Prediction using Deep Learning. *Frontiers in Environmental Science* **3** (2016).

133. Ash, J. R. *et al.* Practically significant method comparison protocols for machine learning in small molecule drug discovery. *ChemRxiv* (2024).

134. Wu, Z. *et al.* MoleculeNet: a benchmark for molecular machine learning. *Chemical Science* **9,** 513–530 (2018).

135. Walters, W. P. & Barzilay, R. Critical assessment of AI in drug discovery. *Expert Opinion on Drug Discovery* **16,** 937–947 (2021).

136. Wognum, C. *et al.* A call for an industry-led initiative to critically assess machine learning for real-world drug discovery. *Nature Machine Intelligence* **6,** 1120–1121 (2024).

137. Crusius, D., Cipcigan, F. & C. Biggin, P. Are we fitting data or noise? Analysing the predictive power of commonly used datasets in drug-, materials-, and molecular-discovery. *Faraday Discussions* **256,** 304–321 (2025).

138. Barone, R. & Chanon, M. A New and Simple Approach to Chemical Complexity. Application to the Synthesis of Natural Products. *Journal of Chemical Information and Computer Sciences* **41,** 269–272 (2001).

139. Bertz, S. H. On the complexity of graphs and molecules. *Bulletin of Mathematical Biology* **45,** 849–855 (1983).

140. Bertz, S. H. *The first general index of molecular complexity* 1981.

141. PubChem. *https://pubchem.ncbi.nlm.nih.gov/*

142. Ertl, P. & Schuffenhauer, A. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *Journal of Cheminformatics* **1,** 8 (2009).

143. Landrum, G. *RDKit: Open-source cheminformatics*

144. Takaoka, Y. *et al.* Development of a Method for Evaluating Drug-Likeness and Ease of Synthesis Using a Data Set in Which Compounds Are Assigned Scores Based on Chemists' Intuition. *Journal of Chemical Information and Computer Sciences* **43,** 1269–1275 (2003).

145. Boda, K., Seidel, T. & Gasteiger, J. Structure and reaction based evaluation of synthetic accessibility. *Journal of Computer-Aided Molecular Design* **21,** 311–325 (2007).

146. Yu, J. *et al.* Organic Compound Synthetic Accessibility Prediction Based on the Graph Attention Mechanism. *Journal of Chemical Information and Modeling* **62,** 2973–2986 (2022).

147. Voršilák, M., Kolář, M., Čmelo, I. & Svozil, D. SYBA: Bayesian estimation of synthetic accessibility of organic compounds. *Journal of Cheminformatics* **12,** 35 (2020).

148. Wang, S., Wang, L., Li, F. & Bai, F. DeepSA: a deep-learning driven predictor of compound synthesis accessibility. *Journal of Cheminformatics* **15,** 103 (2023).

149. Thakkar, A., Chadimová, V., Jannik Bjerrum, E., Engkvist, O. & Reymond, J.-L. Retrosynthetic accessibility score (RAscore) – rapid machine learned synthesizability classification from AI driven retrosynthetic planning. *Chemical Science* **12,** 3339–3349 (2021).

150. Calvi, A., Gaudin, T., Miketa, D., Sydow, D. & Wilbraham, L. *Leap: molecular synthesisability scoring with intermediates* 2024. arXiv: 2403.13005 [q-bio.BM].

151. Liu, C.-H. *et al.* RetroGNN: Fast Estimation of Synthesizability for Virtual Screening and De Novo Design by Learning from Slow Retrosynthesis Software. *Journal of Chemical Information and Modeling* **62,** 2293–2300 (2022).

152. Huang, Q., Li, L.-L. & Yang, S.-Y. RASA: A Rapid Retrosynthesis-Based Scoring Method for the Assessment of Synthetic Accessibility of Drug-like Molecules. *Journal of Chemical Information and Modeling* **51,** 2768–2777 (2011).

153. Coley, C. W., Rogers, L., Green, W. H. & Jensen, K. F. SCScore: Synthetic Complexity Learned from a Reaction Corpus. *Journal of Chemical Information and Modeling* **58,** 252–261 (2018).

154. Kim, H., Lee, K., Kim, C., Lim, J. & Kim, W. Y. DFRscore: Deep Learning-Based Scoring of Synthetic Complexity with Drug-Focused Retrosynthetic Analysis for High-Throughput Virtual Screening. *Journal of Chemical Information and Modeling* **64,** 2432–2444 (2024).

155. Neeser, R. M., Correia, B. & Schwaller, P. *FSscore: A Machine Learning-based Synthetic Feasibility Score Leveraging Human Expertise* 2024. arXiv: 2312.12737 [cs.LG].

156. Reaxys. *https://www.reaxys.com/*

157. Skoraczyński, G., Kitlas, M., Miasojedow, B. & Gambin, A. Critical assessment of synthetic accessibility scores in computer-assisted synthesis planning. *Journal of Cheminformatics* **15,** 6 (2023).

158. Schwaller, P. *et al.* Machine intelligence for chemical reaction space. *WIREs Computational Molecular Science* **12,** e1604 (2022).

159. Tu, Z., Stuyver, T. & Coley, C. W. Predictive chemistry: machine learning for reaction deployment, reaction development, and reaction discovery. *Chemical Science* **14,** 226–244 (2023).

160. Madzhidov, T. I. *et al.* Machine learning modelling of chemical reaction characteristics: yesterday, today, tomorrow. *Mendeleev Communications* **31,** 769–780 (2021).

161. Wei, J. N., Duvenaud, D. & Aspuru-Guzik, A. Neural Networks for the Prediction of Organic Chemistry Reactions. *ACS Central Science* **2,** 725–732 (2016).

162. Segler, M. H. S. & Waller, M. P. Neural-Symbolic Machine Learning for Retrosynthesis and Reaction Prediction. *Chemistry – A European Journal* **23,** 5966–5971 (2017).

163. Coley, C. W., Barzilay, R., Jaakkola, T. S., Green, W. H. & Jensen, K. F. Prediction of Organic Reaction Outcomes Using Machine Learning. *ACS Central Science* **3,** 434–443 (2017).

164. Schwaller, P., Gaudin, T., Lányi, D., Bekas, C. & Laino, T. "Found in Translation": predicting outcomes of complex organic chemistry reactions using neural sequence-to-sequence models. *Chemical Science* **9,** 6091–6098 (2018).

165. Nam, J. & Kim, J. *Linking the Neural Machine Translation and the Prediction of Organic Chemistry Reactions* 2016. arXiv: `1612.09529 [cs.LG]`.

166. Schwaller, P. *et al.* Molecular Transformer: A Model for Uncertainty-Calibrated Chemical Reaction Prediction. *ACS Central Science* **5,** 1572–1583 (2019).

167. Irwin, R., Dimitriadis, S., He, J. & Bjerrum, E. J. Chemformer: a pre-trained transformer for computational chemistry. *Machine Learning: Science and Technology* **3,** 015022 (2022).

168. Zhong, Z. *et al.* Root-aligned SMILES: a tight representation for chemical reaction prediction. *Chemical Science* **13,** 9023–9034 (2022).

169. Tu, Z. & Coley, C. W. Permutation Invariant Graph-to-Sequence Model for Template-Free Retrosynthesis and Reaction Prediction. *Journal of Chemical Information and Modeling* **62,** 3503–3513 (2022).

170. Yoo, S. *et al. Graph-Aware Transformer: Is Attention All Graphs Need?* 2020. arXiv: `2006.05213 [cs.LG]`.

171. Coley, C. W. *et al.* A graph-convolutional neural network model for the prediction of chemical reactivity. *Chemical Science* **10,** 370–377 (2019).

172. Sacha, M. *et al.* Molecule Edit Graph Attention Network: Modeling Chemical Reactions as Sequences of Graph Edits. *Journal of Chemical Information and Modeling* **61,** 3273–3284 (2021).

173. Bradshaw, J., Paige, B., Kusner, M. J., Segler, M. & Hernández-Lobato, J. M. *A Model to Search for Synthesizable Molecules* in *Advances in Neural Information Processing Systems* **32** (Curran Associates, Inc., 2019).

174. Pesciullesi, G., Schwaller, P., Laino, T. & Reymond, J.-L. Transfer learning enables the molecular transformer to predict regio- and stereoselective reactions on carbohydrates. *Nature Communications* **11,** 4874 (2020).

175. Zhang, Y. *et al.* Data augmentation and transfer learning strategies for reaction prediction in low chemical data regimes. *Organic Chemistry Frontiers* **8,** 1415–1423 (2021).

176. Wang, L., Zhang, C., Bai, R., Li, J. & Duan, H. Heck reaction prediction using a transformer model based on a transfer learning strategy. *Chemical Communications* **56,** 9368–9371 (2020).

177. Kreutter, D., Schwaller, P. & Reymond, J.-L. Predicting enzymatic reactions with a molecular transformer. *Chemical Science* **12,** 8648–8659 (2021).

178. Probst, D. *et al.* Biocatalysed synthesis planning using data-driven learning. *Nature Communications* **13,** 964 (2022).

179. Coley, C. W., Green, W. H. & Jensen, K. F. Machine Learning in Computer-Aided Synthesis Planning. *Accounts of Chemical Research* **51,** 1281–1289 (2018).

180. Thakkar, A. *et al.* Artificial intelligence and automation in computer aided synthesis planning. *Reaction Chemistry & Engineering* **6,** 27–51 (2021).

181. Shen, Y. *et al.* Automation and computer-assisted planning for chemical synthesis. *Nature Reviews Methods Primers* **1,** 1–23 (2021).

182. Maser, M. R. *et al.* Multilabel Classification Models for the Prediction of Cross-Coupling Reaction Conditions. *Journal of Chemical Information and Modeling* **61,** 156–166 (2021).

183. Gao, H. *et al.* Using Machine Learning To Predict Suitable Conditions for Organic Reactions. *ACS Central Science* **4,** 1465–1476 (2018).

184. Li, J. & Eastgate, M. D. Making better decisions during synthetic route design: leveraging prediction to achieve greenness-by-design. *Reaction Chemistry & Engineering* **4,** 1595–1607 (2019).

185. Marcou, G. *et al.* Expert System for Predicting Reaction Conditions: The Michael Reaction Case. *Journal of Chemical Information and Modeling* **55,** 239–250 (2015).

186. Beker, W. *et al.* Machine Learning May Sometimes Simply Capture Literature Popularity Trends: A Case Study of Heterocyclic Suzuki–Miyaura Coupling. *Journal of the American Chemical Society* **144,** 4819–4827 (2022).

187. Vaucher, A. C. *et al.* Automated extraction of chemical synthesis actions from experimental procedures. *Nature Communications* **11,** 3601 (2020).

188. Häse, F., Roch, L. M., Kreisbeck, C. & Aspuru-Guzik, A. Phoenics: A Bayesian Optimizer for Chemistry. *ACS Central Science* **4,** 1134–1145 (2018).

189. Shields, B. J. *et al.* Bayesian reaction optimization as a tool for chemical synthesis. *Nature* **590,** 89–96 (2021).

190. Reker, D., Hoyt, E. A., Bernardes, G. J. L. & Rodrigues, T. Adaptive Optimization of Chemical Reactions with Minimal Experimental Information. *Cell Reports Physical Science* **1** (2020).

191. Zhou, Z., Li, X. & Zare, R. N. Optimizing Chemical Reactions with Deep Reinforcement Learning. *ACS Central Science* **3,** 1337–1344 (2017).

192. Chen, W. L., Chen, D. Z. & Taylor, K. T. Automatic reaction mapping and reaction center detection. *WIREs Computational Molecular Science* **3,** 560–593 (2013).

193. *NextMove Software, NameRxn* 2024.

194. Schwaller, P., Hoover, B., Reymond, J.-L., Strobelt, H. & Laino, T. Extraction of organic chemistry grammar from unsupervised learning of chemical reactions. *Science Advances* **7,** eabe4166 (7, 2021).

195. Nugmanov, R., Dyubankova, N., Gedich, A. & Wegner, J. K. Bidirectional Graphormer for Reactivity Understanding: Neural Network Trained to Reaction Atom-to-Atom Mapping Task. *Journal of Chemical Information and Modeling* **62,** 3307–3315 (2022).

196. Chen, S., An, S., Babazade, R. & Jung, Y. Precise atom-to-atom mapping for organic reactions via human-in-the-loop machine learning. *Nature Communications* **15,** 2250 (2024).

197. Schwaller, P., Vaucher, A. C., Laino, T. & Reymond, J.-L. Prediction of chemical reaction yields using deep learning. *Machine Learning: Science and Technology* **2,** 015016 (2021).

198. Huerta, F., Hallinder, S. & Minidis, A. Machine Learning to Reduce Reaction Optimization Lead Time – Proof of Concept with Suzuki, Negishi and Buchwald-Hartwig Cross-Coupling Reactions. *ChemRxiv* (2020).

199. Ahneman, D. T., Estrada, J. G., Lin, S., Dreher, S. D. & Doyle, A. G. Predicting reaction performance in C–N cross-coupling using machine learning. *Science* **360,** 186–190 (2018).

200. Haywood, A. L. *et al.* Kernel Methods for Predicting Yields of Chemical Reactions. *Journal of Chemical Information and Modeling* **62,** 2077–2092 (2022).

201. Fu, Z. *et al.* Optimizing chemical reaction conditions using deep learning: a case study for the Suzuki–Miyaura cross-coupling reaction. *Organic Chemistry Frontiers* **7,** 2269–2277 (2020).

202. Nielsen, M. K., Ahneman, D. T., Riera, O. & Doyle, A. G. Deoxyfluorination with Sulfonyl Fluorides: Navigating Reaction Space with Machine Learning. *Journal of the American Chemical Society* **140,** 5004–5008 (2018).

203. Skoraczyński, G. *et al.* Predicting the outcomes of organic reactions via machine learning: are current descriptors sufficient? *Scientific Reports* **7,** 3582 (2017).

204. Kravtsov, A. A., Karpov, P. V., Baskin, I. I., Palyulin, V. A. & Zefirov, N. S. Prediction of rate constants of SN2 reactions by the multicomponent QSPR method. *Doklady Chemistry* **440,** 299–301 (2011).

205. Kravtsov, A. A., Karpov, P. V., Baskin, I. I., Palyulin, V. A. & Zefirov, N. S. Prediction of the preferable mechanism of nucleophilic substitution at saturated carbon atom and prognosis of SN1 rate constants by means of QSPR. *Doklady Chemistry* **441,** 314–317 (2011).

206. Glavatskikh, M. *et al.* Predictive Models for Kinetic Parameters of Cycloaddition Reactions. *Molecular Informatics* **38,** 1800077 (2019).

207. Li, X., Zhang, S.-Q., Xu, L.-C. & Hong, X. Predicting Regioselectivity in Radical CH Functionalization of Heterocycles through Machine Learning. *Angewandte Chemie International Edition* **59,** 13253–13259 (2020).

208. Struble, T. J., Coley, C. W. & Jensen, K. F. Multitask prediction of site selectivity in aromatic C–H functionalization reactions. *Reaction Chemistry & Engineering* **5,** 896–902 (2020).

209. Guan, Y. *et al.* Regio-selectivity prediction with a machine-learned reaction representation and on-the-fly quantum mechanical descriptors. *Chemical Science* **12,** 2198–2208 (2021).

210. Zahrt, A. F. *et al.* Prediction of higher-selectivity catalysts by computer-driven workflow and machine learning. *Science* **363,** eaau5631 (2019).

211. Moon, S., Chatterjee, S., Seeberger, P. H. & Gilmore, K. Predicting glycosylation stereoselectivity using machine learning. *Chemical Science* **12,** 2931–2939 (2021).

212. Schneider, N., Lowe, D. M., Sayle, R. A. & Landrum, G. A. Development of a Novel Fingerprint for Chemical Reactions and Its Application to Large-Scale Reaction Classification and Similarity. *Journal of Chemical Information and Modeling* **55,** 39–53 (2015).

213. Probst, D., Schwaller, P. & Reymond, J.-L. Reaction classification and yield prediction using the differential reaction fingerprint DRFP. *Digital Discovery* **1,** 91–97 (2022).

214. Schwaller, P. *et al.* Mapping the space of chemical reactions using attention-based neural networks. *Nature Machine Intelligence* **3,** 144–152 (2021).

215. Roughley, S. D. & Jordan, A. M. The Medicinal Chemist's Toolbox: An Analysis of Reactions Used in the Pursuit of Drug Candidates. *Journal of Medicinal Chemistry* **54,** 3451–3479 (2011).

216. Corey, E. J. in *The Chemistry of Natural Products* 19–37 (Butterworth-Heinemann, 1967). ISBN: 978-0-08-020741-4.

217. PENSAK, D. A. & COREY, E. J. in *Computer-Assisted Organic Synthesis ACS Symposium Series* 61, 1–32 (AMERICAN CHEMICAL SOCIETY, 1977). ISBN: 978-0-8412-0394-5.

218. Klucznik, T. *et al.* Efficient Syntheses of Diverse, Medicinally Relevant Targets Planned by Computer and Executed in the Laboratory. *Chem* **4,** 522–532 (2018).

219. Szymkuć, S. *et al.* Computer-Assisted Synthetic Planning: The End of the Beginning. *Angewandte Chemie International Edition* **55,** 5904–5937 (2016).

220. Segler, M. H. S., Preuss, M. & Waller, M. P. Planning chemical syntheses with deep neural networks and symbolic AI. *Nature* **555,** 604–610 (2018).

221. Coley, C. W. *et al.* A robotic platform for flow synthesis of organic compounds informed by AI planning. *Science* **365,** eaax1566 (2019).

222. Genheden, S. *et al.* AiZynthFinder: a fast, robust and flexible open-source software for retrosynthetic planning. *Journal of Cheminformatics* **12,** 70 (2020).

223. Liu, B. *et al.* Retrosynthetic Reaction Prediction Using Neural Sequence-to-Sequence Models. *ACS Central Science* **3,** 1103–1113 (2017).

224. Schwaller, P. *et al.* Predicting retrosynthetic pathways using transformer-based models and a hyper-graph exploration strategy. *Chemical Science* **11,** 3316–3325 (2020).

225. *IBM. Rxn for chemistry.*

226. Fortunato, M. E., Coley, C. W., Barnes, B. C. & Jensen, K. F. Data Augmentation and Pretraining for Template-Based Retrosynthetic Prediction in

Computer-Aided Synthesis Planning. *Journal of Chemical Information and Modeling* **60,** 3398–3407 (2020).

227. Chen, S. & Jung, Y. Deep Retrosynthetic Reaction Prediction using Local Reactivity and Global Attention. *JACS Au* **1,** 1612–1620 (2021).

228. Karpov, P., Godin, G. & Tetko, I. A Transformer Model for Retrosynthesis. *ChemRxiv* (2019).

229. Zheng, S., Rao, J., Zhang, Z., Xu, J. & Yang, Y. Predicting Retrosynthetic Reactions Using Self-Corrected Transformer Neural Networks. *Journal of Chemical Information and Modeling* **60,** 47–55 (2020).

230. Tetko, I. V., Karpov, P., Van Deursen, R. & Godin, G. State-of-the-art augmented NLP transformer models for direct and single-step retrosynthesis. *Nature Communications* **11,** 5575 (2020).

231. Ucak, U. V., Ashyrmamatov, I., Ko, J. & Lee, J. Retrosynthetic reaction pathway prediction through neural machine translation of atomic environments. *Nature Communications* **13,** 1186 (2022).

232. Han, Y. *et al.* Retrosynthesis prediction with an iterative string editing model. *Nature Communications* **15,** 6404 (2024).

233. Zhong, W., Yang, Z. & Chen, C. Y.-C. Retrosynthesis prediction using an end-to-end graph generative architecture for molecular graph editing. *Nature Communications* **14,** 3009 (2023).

234. Somnath, V. R., Bunne, C., Coley, C. W., Krause, A. & Barzilay, R. *Learning Graph Models for Retrosynthesis Prediction* 2021.

235. Chen, Z., Ayinde, O. R., Fuchs, J. R., Sun, H. & Ning, X. $\mathsf{Gˆ2Retro}$ as a Two-Step Graph Generative Models for Retrosynthesis Prediction. *Communications Chemistry* **6,** 102 (2023).

236. Lowe, D. *Chemical reactions from US patents (1976-Sep2016)*

237. Schneider, N., Stiefl, N. & Landrum, G. A. What's What: The (Nearly) Definitive Guide to Reaction Role Assignment. *Journal of Chemical Information and Modeling* **56,** 2336–2346 (2016).

238. Jin, W., Coley, C. W., Barzilay, R. & Jaakkola, T. *Predicting Organic Reaction Outcomes with Weisfeiler-Lehman Network* 2017.

239. Dai, H., Li, C., Coley, C. W., Dai, B. & Song, L. *Retrosynthesis Prediction with Conditional Graph Logic Network* 2020.

240. *NextMove Software, Pistachio* 2022.

241. C. *Chemical Abstracts Service, https://www.cas.org/*

242. Thakkar, A., Kogej, T., Reymond, J.-L., Engkvist, O. & Jannik Bjerrum, E. Datasets and their influence on the development of computer assisted synthesis planning tools in the pharmaceutical domain. *Chemical Science* **11,** 154–168 (2020).

243. Toniato, A., Schwaller, P., Cardinale, A., Geluykens, J. & Laino, T. Unassisted noise reduction of chemical reaction datasets. *Nature Machine Intelligence* **3,** 485–494 (2021).

244. Maziarz, K. *et al.* Re-evaluating Retrosynthesis Algorithms with Syntheseus. *Faraday Discussions* **256,** 568–586 (2025).

245. Chen, B., Li, C., Dai, H. & Song, L. *Retro*: Learning Retrosynthetic Planning with Neural Guided A* Search* 2020. arXiv: `2006.15820 [cs.LG]`.

246. Genheden, S. & Bjerrum, E. PaRoutes: towards a framework for benchmarking retrosynthesis route predictions. *Digital Discovery* **1,** 527–539 (2022).

247. Schneider, N., Sayle, R. A. & Landrum, G. A. Get Your Atoms in Order—An Open-Source Implementation of a Novel and Robust Molecular Canonicalization Algorithm. *Journal of Chemical Information and Modeling* **55,** 2111–2120 (2015).

248. Heller, S. R., McNaught, A., Pletnev, I., Stein, S. & Tchekhovskoi, D. InChI, the IUPAC International Chemical Identifier. *Journal of Cheminformatics* **7,** 23 (2015).

249. *Daylight Theory: SMARTS - A Language for Describing Molecular Patterns*

250. Durant, J. L., Leland, B. A., Henry, D. R. & Nourse, J. G. Reoptimization of MDL Keys for Use in Drug Discovery. *Journal of Chemical Information and Computer Sciences* **42,** 1273–1280 (2002).

251. Rogers, D. & Hahn, M. Extended-Connectivity Fingerprints. *Journal of Chemical Information and Modeling* **50,** 742–754 (2010).

252. Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning* (MIT Press, 2016).

253. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521,** 436–444 (2015).

254. Hecht-nielsen, R. in *Neural Networks for Perception* (ed Wechsler, H.) 65–93 (Academic Press, 1992). ISBN: 978-0-12-741252-8.

255. Breiman, L. Random Forests. *Machine Learning* **45,** 5–32 (2001).

256. Vaswani, A. *et al. Attention Is All You Need* 2017. arXiv: `1706.03762 [cs.CL]`.

257. Bahdanau, D., Cho, K. & Bengio, Y. *Neural Machine Translation by Jointly Learning to Align and Translate* 2014. arXiv: `1409.0473 [cs.CL]`.

258. Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* 2019. arXiv: `1810.04805 [cs.CL]`.

259. Pan, S. J. & Yang, Q. A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering* **22,** 1345–1359 (2010).

260. Crawshaw, M. *Multi-Task Learning with Deep Neural Networks: A Survey* 2020. arXiv: 2009.09796 [cs.LG].

261. Reker, D. & Schneider, G. Active-learning strategies in computer-assisted drug discovery. *Drug Discovery Today* **20,** 458–465 (1, 2015).

262. Lewis, D. D. & Gale, W. A. *A Sequential Algorithm for Training Text Classifiers* 1994. arXiv: cmp-lg/9407020 [cmp-lg].

263. Cohn, D., Atlas, L. & Ladner, R. Improving generalization with active learning. *Machine Learning* **15,** 201–221 (1994).

264. Van Tilborg, D. & Grisoni, F. Traversing chemical space with active deep learning for low-data drug discovery. *Nature Computational Science* **4,** 786–796 (2024).

265. Wieczorek, E. *et al.* Transfer learning for Heterocycle Synthesis Prediction. *ChemRxiv* (2024).

266. McGrath, N. A., Brichacek, M. & Njardarson, J. T. A Graphical Journey of Innovative Organic Architectures That Have Improved Our Lives. *Journal of Chemical Education* **87,** 1348–1349 (2010).

267. Taylor, R. D., MacCoss, M. & Lawson, A. D. G. Rings in Drugs. *Journal of Medicinal Chemistry* **57,** 5845–5859 (2014).

268. Dudkin, V. Y. Bioisosteric equivalence of five-membered heterocycles. *Chemistry of Heterocyclic Compounds* **48,** 27–32 (2012).

269. Meanwell, N. A. Synopsis of Some Recent Tactical Application of Bioisosteres in Drug Design. *Journal of Medicinal Chemistry* **54,** 2529–2591 (2011).

270. Pitt, W. R., Parry, D. M., Perry, B. G. & Groom, C. R. Heteroaromatic Rings of the Future. *Journal of Medicinal Chemistry* **52,** 2952–2963 (2009).

271. Brown, D. G. & Boström, J. Analysis of Past and Present Synthetic Methodologies on Medicinal Chemistry: Where Have All the New Reactions Gone? *Journal of Medicinal Chemistry* **59,** 4443–4458 (2016).

272. Thakkar, A., Selmi, N., Reymond, J.-L., Engkvist, O. & Bjerrum, E. J. "Ring Breaker": Neural Network Driven Synthesis Prediction of the Ring System Chemical Space. *Journal of Medicinal Chemistry* **63,** 8791–8808 (2020).

273. Chu, C., Dabre, R. & Kurohashi, S. *An Empirical Comparison of Domain Adaptation Methods for Neural Machine Translation* in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)* (Association for Computational Linguistics, Vancouver, Canada, 2017), 385–391.

274. Freitag, M. & Al-Onaizan, Y. *Fast Domain Adaptation for Neural Machine Translation* 2016. arXiv: 1612.06897 [cs.CL].

275. Jiang, S. *et al.* When SMILES Smiles, Practicality Judgment and Yield Prediction of Chemical Reaction via Deep Chemical Language Processing. *IEEE Access* **9,** 85071–85083 (2021).

276. Probst, D. & Reymond, J.-L. Visualization of very large high-dimensional data sets as minimum spanning trees. *Journal of Cheminformatics* **12,** 12 (2020).

277. Kovács, D. P., McCorkindale, W. & Lee, A. A. Quantitative interpretation explains machine learning models for chemical reaction prediction and uncovers bias. *Nature Communications* **12,** 1695 (2021).

278. Ramsundar, B. *et al. Deep Learning for the Life Sciences* (O'Reilly Media, 2019).

279. Klein, G., Kim, Y., Deng, Y., Senellart, J. & Rush, A. M. *OpenNMT: Open-Source Toolkit for Neural Machine Translation* in *Proc. ACL* (2017).

280. Coley, C. W., Green, W. H. & Jensen, K. F. RDChiral: An RDKit Wrapper for Handling Stereochemistry in Retrosynthetic Template Extraction and Application. *Journal of Chemical Information and Modeling* **59,** 2529–2537 (24, 2019).

281. Martín Abadi *et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems* 2015.

282. Jiang, B., Dong, J.-j., Jin, Y., Du, X.-l. & Xu, M. The First Proline-Catalyzed Friedlander Annulation: Regioselective Synthesis of 2-Substituted Quinoline Derivatives. *European Journal of Organic Chemistry* **2008,** 2693–2696 (2008).

283. Sun, Y. *et al.* Discovery of the First Potent, Selective, and In Vivo Efficacious Polo-like Kinase 4 Proteolysis Targeting Chimera Degrader for the Treatment of TRIM37-Amplified Breast Cancer. *Journal of Medicinal Chemistry* **66,** 8200–8221 (2023).

284. Spatz, P. *et al.* Dual-Acting Small Molecules: Subtype-Selective Cannabinoid Receptor 2 Agonist/Butyrylcholinesterase Inhibitor Hybrids Show Neuroprotection in an Alzheimer's Disease Mouse Model. *Journal of Medicinal Chemistry* **66,** 6414–6435 (2023).

285. Wang, L., Zhang, X., Su, H. & Zhu, J. *A Comprehensive Survey of Continual Learning: Theory, Method and Application* 2024. arXiv: `2302.00487 [cs.LG]`.

286. Pang, C., Qiao, J., Zeng, X., Zou, Q. & Wei, L. Deep Generative Models in De Novo Drug Molecule Generation. *Journal of Chemical Information and Modeling* **64,** 2174–2194 (8, 2024).

287. Stanley, M. & Segler, M. Fake it until you make it? Generative de novo design and virtual screening of synthesizable molecules. *Current Opinion in Structural Biology* **82,** 102658 (1, 2023).

288. Guo, J. & Schwaller, P. *Directly Optimizing for Synthesizability in Generative Molecular Design using Retrosynthesis Models* 16, 2024. arXiv: 2407.12186[q-bio].

289. Bradshaw, J., Paige, B., Kusner, M. J., Segler, M. & Hernández-Lobato, J. M. *Barking up the right tree: an approach to search over molecule synthesis DAGs* in *Advances in Neural Information Processing Systems* **33** (Curran Associates, Inc., 2020), 6852–6866.

290. Gao, W., Mercado, R. & Coley, C. W. *Amortized Tree Generation for Bottom-up Synthesis Planning and Synthesizable Molecular Design* 12, 2022. arXiv: 2110.06389[cs,q-bio].

291. Korovina, K. *et al. ChemBO: Bayesian Optimization of Small Organic Molecules with Synthesizable Recommendations* in *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics* International Conference on Artificial Intelligence and Statistics (PMLR, 3, 2020), 3393–3403.

292. Swanson, K. *et al.* Generative AI for designing and validating easily synthesizable and structurally novel antibiotics. *Nature Machine Intelligence* **6,** 338–353 (2024).

293. Gottipati, S. K. *et al. Learning to Navigate The Synthetically Accessible Chemical Space Using Reinforcement Learning* in *Proceedings of the 37th International Conference on Machine Learning* International Conference on Machine Learning (PMLR, 21, 2020), 3668–3679.

294. Horwood, J. & Noutahi, E. Molecular Design in Synthetically Accessible Chemical Space via Deep Reinforcement Learning. *ACS Omega* **5,** 32984–32994 (29, 2020).

295. Koziarski, M. *et al. RGFN: Synthesizable Molecular Generation Using GFlowNets* 1, 2024. arXiv: 2406.08506[physics,q-bio].

296. Cretu, M., Harris, C., Roy, J., Bengio, E. & Liò, P. *SynFlowNet: Towards Molecule Design with Guaranteed Synthesis Pathways* 2, 2024. arXiv: 2405.01155[cs,q-bio].

297. Gao, W., Luo, S. & Coley, C. W. *Generative Artificial Intelligence for Navigating Synthesizable Chemical Space* 2024. arXiv: 2410.03494 [cs.LG].

298. Dolfus, U., Briem, H., Gutermuth, T. & Rarey, M. Full Modification Control over Retrosynthetic Routes for Guided Optimization of Lead Structures. *Journal of Chemical Information and Modeling* **63,** 6587–6597 (13, 2023).

299. Dolfus, U., Briem, H. & Rarey, M. Synthesis-Aware Generation of Structural Analogues. *Journal of Chemical Information and Modeling* **62,** 3565–3576 (8, 2022).

300. Konze, K. D. *et al.* Reaction-Based Enumeration, Active Learning, and Free Energy Calculations To Rapidly Explore Synthetically Tractable Chemical Space and Optimize Potency of Cyclin-Dependent Kinase 2 Inhibitors. *Journal of Chemical Information and Modeling* **59,** 3782–3793 (23, 2019).

301. Gusev, F., Gutkin, E., Kurnikova, M. G. & Isayev, O. Active Learning Guided Drug Design Lead Optimization Based on Relative Binding Free Energy Modeling. *Journal of Chemical Information and Modeling* **63,** 583–594 (23, 2023).

302. Wood, D. J. *et al.* Differences in the Conformational Energy Landscape of CDK1 and CDK2 Suggest a Mechanism for Achieving Selective CDK Inhibition. *Cell Chemical Biology* **26,** 121–130.e5 (17, 2019).

303. Guagnano, V. *et al.* Discovery of 3-(2,6-Dichloro-3,5-dimethoxy-phenyl)-1-{6-[4-(4-ethyl-piperazin-1-yl)-phenylamino]-pyrimidin-4-yl}-1-methyl-urea (NVP-BGJ398), A Potent and Selective Inhibitor of the Fibroblast Growth Factor Receptor Family of Receptor Tyrosine Kinase. *Journal of Medicinal Chemistry* **54,** 7066–7083 (27, 2011).

304. Huang, S. *et al.* Structural basis for recognition of anti-migraine drug lasmiditan by the serotonin receptor 5-HT1F–G protein complex. *Cell Research* **31,** 1036–1038 (2021).

305. *Spruce 1.5.3. OpenEye, Cadence Molecular Sciences, Santa Fe, NM.*

306. Cohen, M. P. *et al. pat.* WO2003084949A1 (2003).

307. McInturff, E. L. *et al.* Synthetic Approaches to the New Drugs Approved During 2021. *Journal of Medicinal Chemistry* **66,** 10150–10201 (10, 2023).

308. *Mcule Database.*

309. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12,** 2825–2830 (2011).

310. *OEDOCKING 4.1.2. OpenEye, Cadence Molecular Sciences, Santa Fe, NM.*

311. Bickerton, G. R., Paolini, G. V., Besnard, J., Muresan, S. & Hopkins, A. L. Quantifying the chemical beauty of drugs. *Nature Chemistry* **4,** 90–98 (2012).

312. Byth, K. F. *et al.* AZD5438, a potent oral inhibitor of cyclin-dependent kinases 1, 2, and 9, leads to pharmacodynamic changes and potent antitumor effects in human tumor xenografts. *Molecular Cancer Therapeutics* **8,** 1856–1866 (14, 2009).

313. Zdrazil, B. *et al.* The ChEMBL Database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic Acids Research* **52,** D1180–D1192 (D1 5, 2024).

314. White, K. *et al.* Infigratinib for the Treatment of Metastatic or Locally Advanced Cholangiocarcinoma With Known FGFR2 Gene Fusions or Rearrangements. *Cureus* **15,** e46792.

315. Berger, A. A. *et al.* Lasmiditan for the Treatment of Migraines With or Without Aura in Adults. *Psychopharmacology Bulletin* **50,** 163–188 (15, 2020).

316. Kearnes, S. M. *et al.* The Open Reaction Database. *Journal of the American Chemical Society* **143,** 18820–18826 (2021).

317. Torren-Peraire, P. *et al.* Models Matter: the impact of single-step retrosynthesis on synthesis planning. *Digital Discovery* **3,** 558–572 (2024).

318. Maziarz, K. *et al. Chimera: Accurate retrosynthesis prediction by ensembling models with diverse inductive biases* 2024. arXiv: `2412.05269 [cs.LG]`.

319. Heid, E. *et al.* Chemprop: A Machine Learning Package for Chemical Property Prediction. *Journal of Chemical Information and Modeling* **64,** 9–17 (2024).

320. Klarich, K., Goldman, B., Kramer, T., Riley, P. & Walters, W. P. Thompson SamplingAn Efficient Method for Searching Ultralarge Synthesis on Demand Databases. *Journal of Chemical Information and Modeling* **64,** 1158–1171 (2024).

Appendix A

Mixed fine-tuning optimisation

We studied the effect of dataset weight ratio and number of fine-tuning steps on the performance of the *mixed fine-tuned* model.

The dataset weights control the proportion of reactions from each dataset (General and Ring) included in each training batch. We trained five different models with dataset ratios ranging from 1 Ring: 9 General to 9 Ring: 1 General (Figure A.1). Increasing the ratio of heterocycle formations improves the model performance on ring forming reactions but decreases it for other reaction classes. The optimal dataset ratio was found to be 1:1, which demonstrates a significant improvement over other partitions, while maintaining a similar accuracy for predicting general reactions (baseline model). Consequently, we utilized the 1:1 dataset ratio for mixed fine-tuning.

Next, we investigated the effect of the number of fine-tuning steps (Figure A.2). Fine-tuning for even 2000 steps shows an improved performance compared to the baseline model. However, further improvements are marginal, with around 1% improvement for every additional 2,000 steps. Beyond 6,000 steps, the reactant accuracy fluctuates around 36.5%, indicating that further fine-tuning does not enhance the model performance. Thus, we conclude that fine-tuning for 6,000 steps with a 1:1 dataset weight ratio yields the most optimal results.

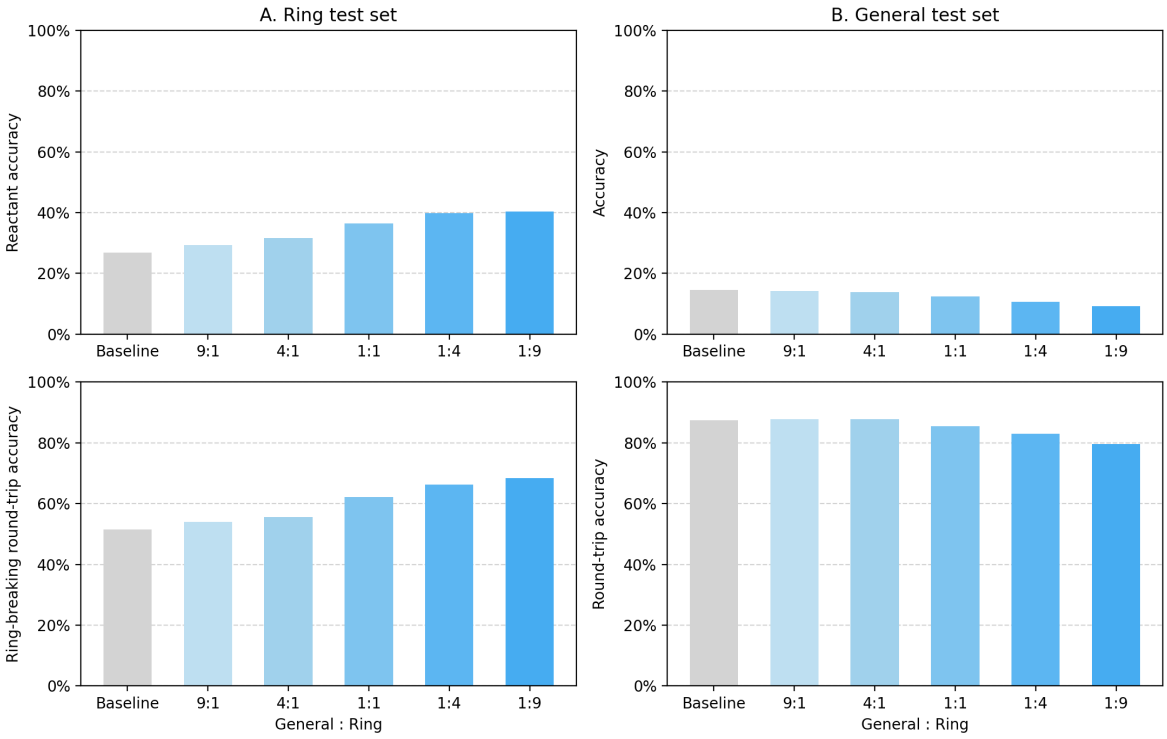


Figure A.1: Effect of dataset weight ratio on model performance in mixed fine-tuning. Top-1 reactant accuracy and proportion of valid ring-breaking top-1 predictions are shown for the Ring test set. Top-1 accuracy and round-trip accuracy are shown for the General test set.

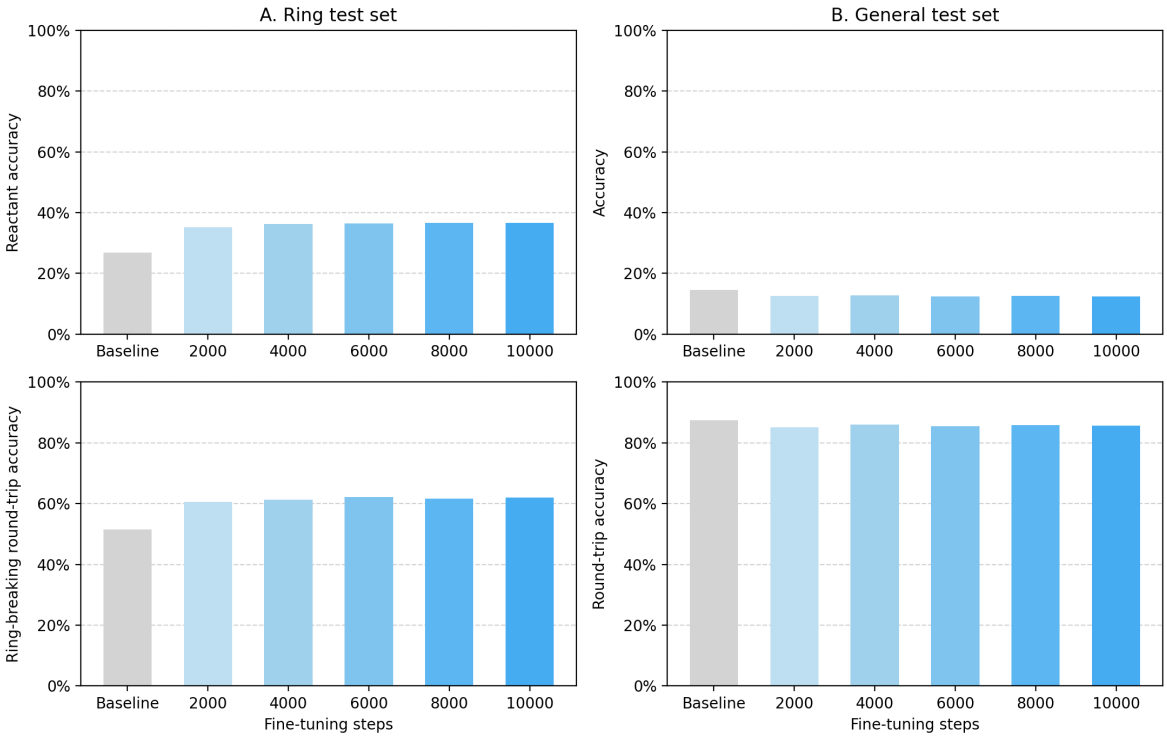


Figure A.2: Effect of the number of fine-tuning steps on model performance in mixed fine-tuning. Top-1 reactant accuracy and proportion of valid ring-breaking top-1 predictions are shown for the Ring test set. Top-1 accuracy and round-trip accuracy are shown for the General test set.

Appendix B

Recent reaction dataset

The reactions forming the *Recent* dataset were extracted from the reactant scope sections of the following articles:

Z.-H. Wang, L.-W. Shen, P. Yang, Y. You, J.-Q. Zhao, W.-C. Yuan, J. Org. Chem., 2022, 87, 5804-5816.

Z. Liu, S. Zhong, X. Ji, G.-J. Deng, H. Huang, Org. Lett., 2022, 24, 349-353.

Z. Li, K. Qiu, X. Yang, W. Zhou, Q. Cai, Org. Lett., 2022, 24, 2989-2992.

Z. Alikhani, A. G. Albertson, C. A. Walter, P. J. Masih, T. Kesharwani, J. Org. Chem., 2022, 87, 6312-6320.

Y.-H. Ma, F.-X. Meng, R.-N. Wang, Y.-X. Fan, Q.-Q. Su, J.-Y. Du, Synthesis, 2022, 54, 499-505.

Y.-C. Liu, P. Chen, X.-J. Li, B.-Q. Xiong, K.-W. Tang, P.-F. Huang, J. Org. Chem., 2022, 87, 4263-4272.

Y. Zheng, Y. Long, H. Gong, J. Xu, C. Zhang, H. Fu, X. Zheng, H. Chen, R. Li, Org. Lett., 2022, 24, 3878-3883.

Y. Yamaguchi, Y. Seino, A. Suzuki, Y. Kamei, T. Yoshino, M. Kojima, S. Matunaga, Org. Lett., 2022, 24, 2441-2445.

Y. V. Ostapiuk, M. Shehedyn, O. V. Barabash, B. Demydchuk, S. Batsyts, C.

Herzberger, A. Schmidt, Synthesis, 2022, 54, 732-740.

X. Jin, L. Xing, D. D. Deng, Y. Yan, Y. Fu, W. Dong, J. Org. Chem., 2022, 87, 1541-1544.

W. Li, R. Shi, S. Chen, X. Zhang, W. Peng, S. Chen, J. Li, X.-M. Xu, Y.-P. Zhu, X. Wang, J. Org. Chem., 2022, 87, 3014-3024.

T. Yang, H. Li, Z. Nie, M.-d. Su, W.-p. Luo, Q. Liu, C.-C. Guo, J. Org. Chem., 2022, 87, 2797-2808.

S. Zhang, Q. Zhang, M. Tang, J. Org. Chem., 2022, 87, 3845-3850. R. Zhang, M. Sun, Q. Yan, X. Lin, X. Li, X. Fang, H. H. Y.

Sung, J. D. Williams, J. Sun, Org. Lett., 2022, 24, 2359-2364.

Q. Guo, J. Chen, G. Shen, G. Lu, X. Yang, Y. Tang, Y. Zhu, S. Wu, B. Fan, J. Org. Chem., 2022, 87, 540-546.

M. Vadivelu, A. A. Raheem, J. P. Raj, J. Elangovan, K. Karthikeyan, C. Praveen, J. Sun, Org. Lett., 2022, 24, 2798-2803.

M. V. Il'in, A. A. Sysoeva, A. S. Novikov, D. S. Bolotin, J. Org. Chem., 2022, 87, 4569-4579.

M. Maji, I. Borthakur, S. Srivastava, S. Kundu, J. Org. Chem., 2022, 87, 5603-5616.

M. Baidya, S. Mallick, S. D. Sarkar, Org. Lett., 2022, 24, 1274-1279.

M. B. Reddy, K. Prasanth, R. Anandhan, Org. Lett., 2022, 24, 3674-3679.

M. A. Ansari, G. Kumar, M. S. Singh, Org. Lett., 2022, 24, 2815-2820.

L. Ren, J. Luo, L. Tan, Q. Tang, J. Org. Chem., 2022, 87, 3167-3176.

L. Liu, J. Lin, M. Pang, H. Jin, X. Yu, S. Wang, Org. Lett., 2022, 24, 1146-1151.

K. Sakai, K. Oisaki, M. Kanai, Org. Lett., 2022, 24, 3325-3330.

K. Ishihara, T. Shioiri, M. Matsugi, Synlett, 2022, 33, 781-784.

K. H. Min, N. Iqbal, E. J. Cho, Org. Lett., 2022, 24, 989-994.

K. Gnyawali, P. T. K. Arachchige, C. S. Yi, Org. Lett., 2022, 24, 218-222.

J.-L. Liu, W. Wang, X. Qi, X.-F. Wu, Org. Lett., 2022, 24, 2248-2252.

J. Ying, T. Liu, Y. Liu, J.-P. Wan, Org. Lett., 2022, 24, 2393-2398.

J. Talvitie, I. Alanko, E. Bulatov, J. Koivula, T. Pöllänen, J. Helaja, Org. Lett., 2022, 24, 274-278.

J. Li, Y. Liu, Z. Chen, J. Li, j. Li, X. Ji, L. Chen, Y. Huang, Q. Liu, Y. Li, J. Org. Chem., 2022, 87, 3555-3566.

J. K. Laha, M. K. Hunjan, J. Org. Chem., 2022, 87, 2315-2323.

J. Jiao, P. Wang, F. Xiao, Z. Zhang, Synlett, 2022, 33, 569-574.

J. Hou, G. Yang, Z. Chai, J. Org. Chem., 2022, 87, 453-463.

J. Dong, J. Hu, X. Liu, S. Sun, L. Bao, M. Jia, X. Xu, J. Org. Chem., 2022, 87, 2845-2852.

H. Shui, Y. Zhong, L. Ouyang, N. Luo, R. Luo, Synthesis, 2022, 54, 2876-2884.

H. Guo, L. Tian, Y. Liu, J.-P. Wan, Org. Lett., 2022, 24, 228-233.

G. E. Bell, J. W. B. Fyfe, E. M. Israel, A. M. Z. Slawin, M. Campbell, A. J. B. Watson, Org. Lett., 2022, 24, 3024-3027.

F. S. Movahed, S. W. Foo, S. Mori, S. Ogawa, S. Saito, J. Org. Chem., 2022, 87, 243-257.

F. Lu, Y. Chen, X. Song, C. Yu, T. Li, K. Zhang, C. Yao, J. Org. Chem., 2022, 87, 6902-6909.

D. Zhuang, T. Gatera, Z. An, R. Yan, Org. Lett., 2022, 24, 771-775.

D. Jankovič, M. Virant, M. Gazvoda, J. Org. Chem., 2022, 87, 4018-4028.

C. Shan, J. Xu, L. Cao, C. Liang, R. Cheng, X. Yao, M. Sun, J. Ye, Org. Lett., 2022, 24, 3205-3240.

B. Ramesh, M. Jeganmohan, J. Org. Chem., 2022, 87, 6902-6909.

B. Lin, Y. Yao, Y. Huang, Z. Weng, Org. Lett., 2022, 24, 2055-2058.

A. K. Guin, R. Mondal, G. Chakraborty, S. Pal, N. D. Paul, J. Org. Chem., 2022, 87, 7106-7123.

J. Zhang, Y. Zhang, J. Zhang, Q. Wu, H. Yang, Synlett, 2022, 33, 264-268.

Appendix C

Effect of dataset splitting on transfer learning

While the importance of splitting based on Tanimoto similarity of the reaction product has been demonstrated by Lee *et al.*[277] for forward reaction prediction, most retrosynthesis prediction models still use a random train/test split. In this work, we have used Tanimoto similarity to split the Ring dataset into train, validation and test sets when comparing the different domain adaptation methods. However, here we examine the effect of dataset splitting by comparing the performance of the *mixed fine-tuned* model trained on a randomly split dataset as shown in Figure C.1.

Our results show that the performance on the General test set is comparable to the model trained on Tanimoto split data, achieving an accuracy of 12.6% and round-trip accuracy of 85.7%. As the Ring test sets are different, a direct comparison between accuracies is not possible; however, we can assess the improvement between the *baseline* and *mixed fine-tuned* model. As expected, this improvement in reactant-accuracy is greater when the model is trained on randomly split data, increasing from 19.9% reactant-accuracy of the *baseline* model to 35.0% for the *mixed fine-tuned* model. In comparison, when the model is trained on Tanimoto split data, the

accuracy increases from 26.9% to 36.5%. This supports Lee's conclusion that using a random split overstates the model's predictive ability.

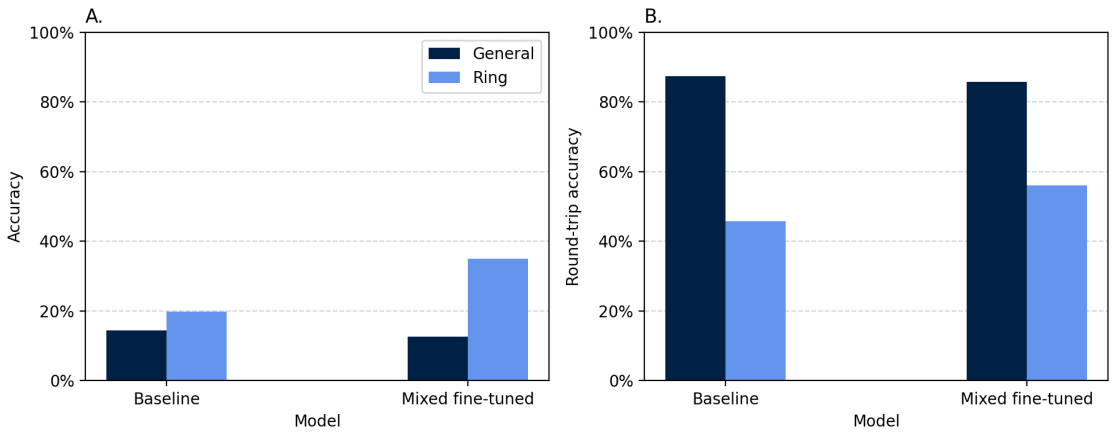


Figure C.1: Performance of the *mixed fine-tuned* model trained on randomly split Ring dataset compared to the *baseline* model: (A) Top-1 accuracy on the General test set and top-1 reactant accuracy on the randomly split Ring test set and (B) top-1 round-trip accuracy on the General test set and proportion of valid ring breaking top-1 predictions for the Ring test set.

Appendix D

Effect of initial random state on *retro-active* performance

For the benchmarking of acquisition and enumeration strategies of *retro-active*, six runs were performed for every combination, each run starting from a different random state. Figure D.1 shows the percentage of ground truth top-100 best scoring molecules that were recovered in every run after each active learning loop, together with the average of all runs. Those results are shown for all four combinations of explorative acquisition (EA) and greedy acquisition (GA) with hypercube enumeration (HE) and exhaustive enumeration (EE), optimising for both the ROCS and docking scores. The initial random state affects the outcome of *retro-active* runs more considerably when optimising for the ROCS score than the docking score.

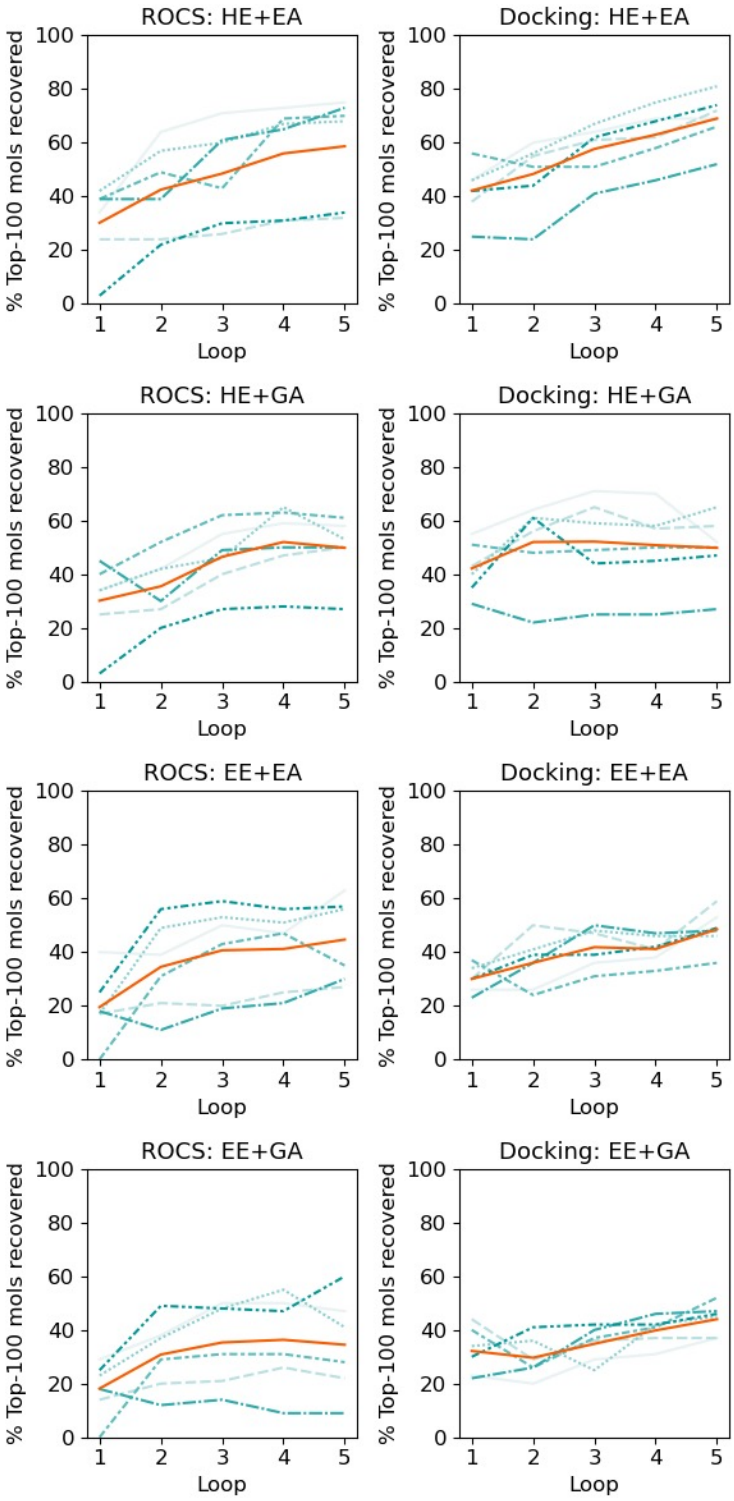


Figure D.1: Performance of *retro-active* across all runs for each combination of enumeration and acquisition methods scored with ROCS and docking scores. The proportion of ground truth top-100 best scoring molecules that would have been recovered after every loop is shown for each run (green) with the mean shown in orange.